Advances in Astronomy

Seeking for the Leading Actor on the Cosmic Stage: Galaxies versus Supermassive Black Holes

Guest Editors: Angela Bongiorno, Francesco Shankar, Francesca Civano, Isabelle Gavignaud, and Antonis Georgakakis

Seeking for the Leading Actor on the Cosmic Stage: Galaxies versus Supermassive Black Holes

Advances in Astronomy

Seeking for the Leading Actor on the Cosmic Stage: Galaxies versus Supermassive Black Holes

Guest Editors: Angela Bongiorno, Francesco Shankar, Francesca Civano, Isabelle Gavignaud, and Antonis Georgakakis

Editorial Board

Contents

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 625126, 3 pages
doi:10.1155/2012/625126

Editorial

Seeking for the Leading Actor on the Cosmic Stage: Galaxies versus Supermassive Black Holes

Angela Bongiorno,[1,2] Francesco Shankar,[3] Francesca Civano,[4] Isabelle Gavignaud,[5] and Antonis Georgakakis[1]

[1] *Max-Planck-Institut für Extraterrestrische Physik (MPE), Giessenbachstraße 1, Garching bei, 85748 München, Germany*
[2] *Osservatorio Astronomico di Roma, Via Frascati 33, Monteporzio Catone, 00040 Rome, Italy*
[3] *GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, 5 Place Jules Janssen, 92195 Meudon, France*
[4] *Smithsonian Astrophysical Observatory, 60 Garden Street, MS 67, Cambridge, MA 02138, USA*
[5] *Departamento de Ciencias Fisicas, Facultad de Ingenieria, Universidad Andres Bello, Avandia Republica 252 Santiago, Chile*

Correspondence should be addressed to Angela Bongiorno, angela.bongiorno@oa-roma.inaf.it

Received 31 May 2012; Accepted 31 May 2012

A major development in extragalactic astrophysics in recent years has been the realization that active galactic nuclei (AGN), which signpost accretion events onto supermassive black holes (SMBHs) [1, 2], may play a fundamental role in the formation and evolution of galaxies. Understanding the physics that drive the growth of SMBH across cosmic time is therefore important for having a complete picture of galaxy formation. The two processes of galaxy and BH evolution can no longer be regarded as separate, as was the case until about 10 years ago, but need to be studied in conjunction.

One of the first indications for a correlation between the formation of galaxies and the growth of SMBHs at their centres is the striking similarity between the redshift evolution of the accretion density and the star-formation rate density [3, 4] of the Universe. Both quantities show a rapid increase from the local Universe to $z \approx 1$, followed by a broad plateau at z ≈ 2–4 and a decline at $z \geq 4$, although the high redshift behavior of X-ray AGN remains controversial [5].

These similarities are more than a mere coincidence or a manifestation of the fact that the Universe was overall more active in the past. Observational evidence now shows that most, if not all, galaxies in the local universe posses a central SMBH at their center. Moreover, there is now emerging consensus from dynamical observations that SMBHs at the center of the local massive and bulge-dominated galaxies are tightly correlated with the velocity dispersion and masses of their stellar hosts, with an intrinsic scatter of a factor of two or even less [6–8]. Such strong correlations argue for a physical association between the two processes and suggest that BHs must have evolved, or coevolved, with their stellar hosts at some point in their past.

Analytical calculations identify AGN feedback as the process that can potentially link SMBH growth and star formation (e.g., [9–14]). In this picture, the energy released by the AGN is sufficient to either heat up or blow away the cold gas of galaxies, thereby irreversibly altering their evolution. Observations of powerful and/or nearby active SMBHs have recently started finding evidence for outflows, most likely associated with the AGN, in either the warm (e.g., [15]) or the cold (e.g., [16, 17]) gas component of their host galaxies.

Despite the increasing evidence for the importance of SMBH growth in galaxy formation, we have just started exploring the physical processes at play. The nature of AGN feedback, how it is related to the fueling mode of the SMBH, and ultimately what is its impact on kpc or even Mpc scales are still debated.

To shed light on these issues from both the observational and theoretical points of view, we initiated this special issue (http://www.hindawi.com/journals/aa/si/610485/) which collects reviews and new results, contributed by some of the key researchers in the field. We made an effort to

bring together observers and theoreticians to provide the community with a comprehensive state-of-the-art overview of our understanding of SMBH evolution across cosmic time. In this respect, the special issue deals with topics from the smallest to the largest scales including, for example, BH accretion mechanisms and properties, the influence on their host galaxies, and the connection to the dark matter halos. We hope this special issue will become a useful reference for all researchers in the field.

The special issue starts with a paper from S. Bianchi et al. who discuss the recent developments on the AGN unified models. They review the standard unified model to then move to an updated unification scenario that can better explain the complex phenomenology observed. This paper is followed by a series of papers which constitute a comprehensive description of the accretion history of the Universe. In particular, F. Fiore et al. present the $z > 3$ X-ray sources number counts in the 0.5–2 keV band and make predictions for new missions; E. Treister and C. M. Urry review what we know about the energy output of AGN describing the most common way to isolate these sources at different wavebands. They also summarize the cosmic history of black hole accretion, i.e., when in the history of the Universe supermassive black holes were obtaining most of their mass. B. C. Kelly and A. Merloni present the demographic of SMBHs in the local Universe discussing advantages and disadvantages of the different methods for estimating the black hole mass function, and A. Sesana reviews the current understanding of massive BH formation and evolution along the cosmic history highlighting which future observations will help to shed light on the cosmic history of SMBHs, paying particular attention to the upcoming gravitational wave window. A comprehensive discussion of current ideas of BHs fueling mechanisms is presented by G. Lodato, while T. Johannsen discusses possible new observational tests for confirming the existence of BH at the centers of galaxies.

However, BHs do not grow only via gas accretion but possibly also via BH-BH mergers. M. Dotti, A. Sesana et al. review the state of the art of mergers as an additional channel of BH growth including the expected observational signatures of massive binaries. In this context, S. Komossa presents the results of numerical relativity simulations which imply that after binary coalescence in a galaxy merger the newly formed single SMBH can receive kick velocities up to several 1000 km/s due to anisotropic emission of gravitational waves. Moreover, she presents the observational signatures of recoiling SMBHs and the properties of the first candidates which have emerged.

After that, we devoted part of the special issue to our current understanding of the connection between SMBHs and their host galaxies inferred from a variety of observational probes. The latest results on the AGN/Starburst connection are presented by E. Sani and E. Nardini who studied a peculiar ultraluminous infrared galaxy (IRAS 20551-4250) experiencing an intense starburst but hosting a highly obscured AGN. A. Constantin and A. C. Seth present the peculiar nature of the nucleus of M94, which is the least luminous broad-line (type 1) LINER with possibly the least luminous broad line region known. This is followed by a paper from R. Decarli et al. on the extension of the scaling relation between SMBHs and host galaxies to the smallest observed local SMBHs, showing that the relation holds over 2 dex in both $M_{\rm BH}$ and M_*. Moreover, N. Neumayer & C. J. Walcher and P. Erwin & D. A. Gadotti extended the work to the nuclear star clusters and their embedded BHs, which are the possible precursors of massive black holes in galaxy nuclei. N. Neumayer & C. J. Walcher studied the low mass end of the global-to-nucleus relations finding that the $M_{\rm BH}$-$M_{\rm bulge}$ relation may well flatten at low masses while the $M_{\rm BH}$-σ relation may steepen. Moreover, P. Erwin, & D. A. Gadotti analyze a sample of disk galaxies and performing a 2D bulge/disk/bar decompositions show that while SMBHs correlate with the stellar mass of the bulge component of galaxies, the masses of nuclear star clusters correlate much better with the total galaxy stellar mass.

We conclude the special issue by discussing what we can learn from studying the large-scale environment of AGN. R. Fassbender et al. studied the distribution of X-ray AGN in the large-scale structure environments of 22 X-ray luminous galaxy clusters in the redshift range $0.9 < z < 1.6$. They found two overdensities, one at r < 1 Mpc of predominantly low-luminosity AGN and another one of brighter soft-band detected AGN at cluster-centric distances of 2-3 Mpc (about three times the average cluster radius R_{200} of the systems). Their results support the idea of two different physical triggering mechanisms of X-ray AGN activity in dependence of the radially changing large-scale structure environment of the distant clusters. M. Gitti et al. draw a qualitative picture of the current knowledge of the effects of the AGN feedback on the intracluster medium by summarizing the recent results in this field. Finally, N. Cappelluti et al. review the now abundant data on AGN small- and large-scale clustering properties from different deep and large surveys and the full important information we can derive from them.

Angela Bongiorno
Francesco Shankar
Francesca Civano
Isabelle Gavignaud
Antonis Georgakakis

References

[1] A. Merloni and S. Heinz, "A synthesis model for AGN evolution: supermassive black holes growth and feedback modes," *Monthly Notices of the Royal Astronomical Society*, vol. 388, no. 3, pp. 1011–1030, 2008.

[2] F. Shankar, "The demography of supermassive black holes: growing monsters at the heart of galaxies," *New Astronomy Reviews*, vol. 53, no. 4–6, pp. 57–77, 2009.

[3] A. Merloni, "The anti-hierarchical growth of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 353, no. 4, pp. 1035–1047, 2004.

[4] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes, and galaxy spheroids," *The Astrophysical Journal Supplement Series*, vol. 163, no. 1, pp. 1–49, 2006.

[5] M. Brusa, A. Comastri, R. Gilli et al., "High-redshift quasars in the cosmos survey: the space density of $z > 3$ X-ray selected QSOs," *The Astrophysical Journal*, vol. 693, no. 1, article 8, 2009.

[6] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *The Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[7] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *The Astrophysical Journal*, vol. 539, pp. L9–L12, 2000.

[8] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *The Astrophysical Journal*, vol. 604, no. 2, pp. L89–L92, 2004.

[9] J. Silk and M. J. Rees, "Quasars and galaxy formation," *Astronomy & Astrophysics*, vol. 331, pp. L1–L4, 1998.

[10] A. C. Fabian, "The obscured growth of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 308, no. 4, pp. L39–L43, 1999.

[11] A. King, "Black holes, galaxy formation, and the $M_{\rm BH}$-σ relation," *The Astrophysical Journal Letters*, vol. 596, no. 1, article L27, 2003.

[12] G. L. Granato, G. de Zotti, L. Silva, A. Bressan, and L. Danese, "A physical model for the coevolution of QSOs and their spheroidal hosts," *The Astrophysical Journal*, vol. 600, no. 2, pp. 580–594, 2004.

[13] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[14] N. Menci, A. Fontana, E. Giallongo, A. Grazian, and S. Salimbeni, "The abundance of distant and extremely red galaxies: the role of AGN feedback in hierarchical models," *The Astrophysical Journal Letters*, vol. 647, no. 2, pp. 753–762, 2006.

[15] J. Holt, C. N. Tadhunter, R. Morganti, and B. H. C. Emonts, "The impact of the warm outflow in the young (GPS) radio source and ULIRG PKS 1345+12 (4C 12.50)," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 3, pp. 1527–1536, 2011.

[16] C. Feruglio, R. Maiolino, E. Piconcelli et al., "Quasar feedback revealed by giant molecular outflows," *Astronomy & Astrophysics*, vol. 518, no. 13, article L155, 2010.

[17] E. Sturm, E. González-Alfonso, S. Veilleux et al., "Massive molecular outflows and negative feedback in ULIRGs observed by HERSCHEL-pacs," *The Astrophysical Journal Letters*, vol. 733, no. 1, article L16, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 782030, 17 pages
doi:10.1155/2012/782030

Review Article

AGN Obscuration and the Unified Model

Stefano Bianchi,[1] Roberto Maiolino,[2,3] and Guido Risaliti[4,5]

[1] *Dipartimento di Fisica, Università degli Studi Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy*
[2] *INAF-Osservatorio Astronomico di Roma, Via di Frascati 33, 00040 Monte Porzio Catone, Italy*
[3] *Cavendish Laboratory, University of Cambridge, 19 J. J. Thomson Avenue, Cambridge CB3 0HE, UK*
[4] *INAF-Osservatorio Astrofisico di Arcetri, L.go E. Fermi 5, Firenze, Italy*
[5] *High Energy Astrophysics Division, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 0218, USA*

Correspondence should be addressed to Stefano Bianchi, bianchi@fis.uniroma3.it

Received 31 August 2011; Accepted 29 December 2011

Academic Editor: Isabelle Gavignaud

Unification Models of Active Galactic Nuclei postulate that all the observed differences between type 1 and type 2 objects are due to orientation effects with respect to the line of sight to the observer. The key ingredient of these models is the obscuring medium, historically envisaged as a toroidal structure on a parsec scale. However, many results obtained in the last few years are clearly showing the need for a more complex geometrical distribution of the absorbing media. In this paper, we review the various pieces of evidence for obscuring media on different scales, from the vicinity of the black hole to the host galaxy, in order to picture an updated unification scenario explaining the complex observed phenomenology. We conclude by mentioning some of the open issues.

1. Introduction: The Standard Unified Model

In this paper, we discuss the recent developments on the AGN unified models, specifically for what concern the geometry, location, and physics of the absorbing medium. Before discussing the more recent results, in this section we shortly review the early, classical arguments that historically led to the formulation of the standard Unified Model. An early review of the initial results was also given in Antonucci [1].

The first unification attempts have been focussed on polarization measurements. In particular Antonucci [2] found a perpendicular alignment of optical polarization relative to the radio axis in a sample of radio galaxies, which was interpreted as due to scattering of photons, whose direction before entering the line of sight was primarily in the vertical direction. Shortly after, additional evidence was found in low luminosity, local AGN, and specifically Seyfert galaxies [3]. Seyfert 1 galaxies are characterized by the presence of broad optical permitted lines (FWHM $>$ 1000 km/s), such as Hα and Hβ, that are not observed in Seyfert 2 galaxies. However, the presence of both strong high ionization and low ionization narrow (FWHM $<$ 1000 km/s) forbidden lines (such as [O III], [Ne III], [O II], [O I], [N II], [S II]), and several very high ionization coronal lines (such as [Fe X], [Fe XI], [Si IX], [Si X]) is common to both types of Seyfert galaxies and with similar line ratios. The latter finding suggested that all Seyfert galaxies are powered by the same intrinsic engine.

A strong observational evidence of a unification between type 2 and type 1 Seyfert nuclei has been the discovery of broad optical lines in the polarized spectrum of the archetypal Seyfert 2, NGC 1068, obtained by Antonucci and Miller [4]. This finding revealed the presence of a Broad Line Region (BLR) in this Seyfert 2 nucleus, which is hidden to our line of sight, but whose light is scattered in our direction from material (in the case of NGC 1068, likely free electrons in ionized gas) distributed on scales larger than the absorber. Such reflected light is very weak compared to the light of the galaxy, hence difficult to detect in the total spectrum, but it is highly polarized, and therefore detectable in the polarized spectrum. The basic idea of the Unified Model is that type 2 and type 1 AGN are intrinsically the same class of objects and their differences are only due to orientation effects relative to an obscuring medium.

The need for a "toroidal", axisymmetric structure of the absorber was initially inferred from the fact itself that the reflected broad lines are polarized and from the measured polarization angles. If the absorber was a simple cloud along the line of sight, then reflection should come from all directions, hence by averaging from all angles the total polarization should be zero. In order to break the symmetry of the polarization angles the absorber should prevent the nuclear light to be scattered in a significant range of angles, and a "torus" is the most natural configuration that can achieve this effect.

The size of the toroidal absorber was initially postulated to be on the parsec scale [5, 6]. Such typical size was simply inferred by the need for the absorber to be large enough to obscure the BLR, which in Seyfert nuclei has a size well below a parsec, based on reverberation studies (e.g., [7, 8]), but small enough not to obscure the Narrow Line Region (NLR), which is distributed on the 10–100 pc scale. However, as we will discuss extensively in this paper, there is clear evidence that the absorbing medium is also distributed on smaller and larger scales.

Since the seminal discovery of Antonucci and Miller [4], polarized broad lines have been discovered in several other Seyfert 2 nuclei [9, 10], contributing to the generalization of the unified model to the whole class of type 2 AGN. Further evidence for a unified theory between type 2 and type 1 AGN has been obtained from various multiwavelength studies. In a number of type 2 AGN the presence of an obscured BLR was inferred from the detection of a broad component of hydrogen recombination lines in the near infrared, such as Paβ, Paα, Brγ, where dust absorption can be several times lower than in the optical [11–15].

Hard X-ray observations of AGN have provided additional, unambiguous evidence in favor of the unified model, obtained already with the early hard X-ray satellites such as Exosat, Ginga, ASCA, and BeppoSAX [16–19]. Although generally weak or even undetected in the soft X-ray band (<2 keV), most type 2 AGN are detected in the hard X-rays (>2 keV) and are characterized by a power-law spectrum similar to Sy1s, favoring a common central engine, but affected by a photoelectric absorption cutoff directly demonstrating the presence of an absorbing medium along the line of sight, with a column density typically well in excess of 10^{22} cm^{-2}. In a number of Seyfert 2s the hard X-ray spectrum does not show the presence of a prominent absorption cutoff, but it is characterized by a very strong Fe Kα line at 6.4 keV, with an equivalent width larger than 500 eV (e.g., [19–21]). Such high equivalent widths of the Fe line can only be explained by assuming that the direct X-ray radiation is totally absorbed by a Compton thick medium ($N_H > 10^{24}$ cm^{-2}) and that the observed (weak) continuum and Fe Kα are due to reflection from the circumnuclear medium. A detailed discussion of these effects and on the distribution of the reflecting medium will be given in the next sections.

If the NLR in type 2 AGN is created by the photoionization of a nuclear UV/X source hidden from our line of sight by a toroidal absorber, then the expectation is that the NLR should have a (bi)conical morphology, due to the light cones defined by the nuclear absorber. High-resolution, narrow band imaging (or integral field spectroscopy), especially with the advent of HST, has indeed revealed such ionization cones on scales ranging from a few 10 pc up to several 100 pc, in many nearby AGN (e.g., [22–25]). The opening angle of the cones gives the fraction of the sky hidden to our line of sight, which is in reasonable agreement with what inferred from the relative fraction between type 1 and type 2 AGN in the local universe [26]. It is interesting that the axis of the ionization cones is often not aligned with the minor axis of the host galaxy, meaning that the circumnuclear absorber is not necessarily aligned with the gaseous disk of the host galaxy. Generally the orientation of the ionization cone axis is in the same direction as the radio jet, but often the two are not exactly aligned, meaning that even on small (parsec or subparsec) scales there is a slight misalignment between the dusty absorbing medium and the central engine (the accretion disc) [24, 27].

We conclude this introduction by shortly mentioning that several theoretical works have modelled the physics and the structure of the "toroidal" absorbing medium in the attempt of reproducing the observable properties. Initial models had assumed a simple toroidal structure with a uniform distribution of gas and dust with a parsec-scale radius (e.g., [6, 28]), while other models suggested more extended geometries, up to 100 pc, to explain the broad infrared spectral energy distribution observed in AGN [29]. One of the main problems of these models is their dynamical stability. Often radiation pressure from infrared photons within the torus is invoked as a solution to keep the torus geometrically thick [30]. Other authors ascribe the geometrical thickness of the torus to turbulence introduced by supernovae or stellar winds [31, 32]. It has also been proposed that a wide angle of obscuration does not necessarily require a geometrically thick torus, but can also be achieved with a warped or tilted disk [33–36].

Recently, the hypothesis of uniform gas and dust distribution has been abandoned by many models, by introducing a clumpy structure of the absorbing medium [37–40]. These models can account for several of the observational properties of AGN and, most importantly, are strongly supported by recent X-ray observations that directly reveal the clumpiness of the absorbing medium, as discussed in the following sections.

2. From Galactic to Sub-Pc Scale: Absorption at Different Scales

One of the most significant new aspects on the structure of AGN, as emerged in last few years, is that the standard, parsec-scale "torus" is not enough to explain all the complex absorption features discovered by many observations. While the unified picture remains valid in its more general sense (i.e., the presence of nonspherically symmetric absorbers at the origin of the type 1/type 2 dichotomy) several new observations and models, mostly in the X-ray and infrared domain, suggest that multiple absorbers are present around the central source, on quite different physical scales. In the

following we review the main observational evidence for each of them, together with some brief discussion on their physical interpretation.

2.1. Absorption within the Sublimation Radius. The evidence for gas absorption within the sublimation radius comes mostly from X-ray observations. The most direct way to probe the presence of such gas component is through absorption variability measurements.

X-ray absorption variability is a common feature in AGN. An analysis of a sample of nearby obscured AGN with multiple X-ray observations, performed a few years ago [41] revealed that column density (N_H) variations are almost ubiquitous in local Seyfert galaxies. More recent observations performed with XMM-*Newton*, *Chandra,* and *Suzaku* further confirmed this finding. The physical implications of these measurements are that the circumnuclear X-ray absorber (or, at least, one component of it) must be clumpy and located at subparsec distances from the central source.

The comparison between different observations, typically performed at time distances of months-years, only provides upper limits to the intrinsic time scales of N_H variations. An improvement of these estimates could only be obtained through observational campaigns within a few weeks/days, and/or through the search for N_H variations within single long observations. Such short time-scale studies have been performed for a handful sources: NGC 1365 [42–45], NGC 4388 [46], NGC 4151 [47], NGC 7582 [48], and Mrk 766 [49].

In particular, in the case of NGC 1365, *Chandra*, XMM-*Newton,* and *Suzaku* observations revealed extreme spectral changes, from Compton thin (N_H in the range $10^{23}\,cm^{-2}$) to reflection dominated ($N_H > 10^{24}\,cm^{-2}$) on time scales from a couple of days to ~10 hours (see Figure 1). Such rapid events imply that the absorption is due to clouds with velocity $v > 10^3\,km\,s^{-1}$, at distances from the BH of the order of 10^4 gravitational radii (assuming that they are moving with Keplerian velocity around the central black hole). The physical size and density of the clouds are estimated to be of the order of 10^{13} cm and 10^{10}-$10^{11}\,cm^{-3}$, respectively. All these physical parameters are typical of BLR clouds, strongly suggesting that the X-ray absorber and the clouds responsible for broad emission lines in the optical/UV are one and the same.

These results are obtained from the analysis of the observed absorption variations in the sources mentioned above, assuming a simple scheme, where the clouds are homogeneous, with a constant column density and are moving across the line of sight with Keplerian velocity. However, in a few cases with particularly high signal-to-noise ratio, the analysis of X-ray "eclipses" can provide further information on the geometrical and physical structure of the cloud.

In the case of NGC 1365, a careful analysis [45] of the spectral X-ray variability during two eclipses revealed a "cometary" shape of the obscuring cloud, consisting of a high-density head, and an elongated, lower-density tail. This structure is revealed by the time evolution of the two key observational parameters of the cloud (Figure 2): its covering factor to the X-ray source (suddenly increasing at the beginning of the occultation, then slowly increasing over a relatively long time interval), and its column density (highest at the beginning of the occultation, and then decreasing steadily).

Occultations have also been observed in the bright Narrow Line Seyfert 1 Mrk 766 [49], revealing that such events are possible (though rare) in on-average unobscured sources. While the data quality was not enough to perform a column density/covering factor deconvolution as the one described above for NGC 1365, additional information came from the detection of highly ionized iron absorption lines. These lines, due to Fe XXV and Fe XXVI ions, are present only in the same time intervals where the occultations due to the neutral cloud are also present. Furthermore, their energy clearly reveal that the absorbing gas is outflowing with velocities of several $10^3\,km\,s^{-1}$. The straightforward interpretation of these measurements is that the absorption is due to an outflowing cloud, with a high-density, low-ionization head, and a low-density, high-ionization tail.

The above examples remain unique among the available X-ray observations of AGN, but they suggest that with future, larger area X-ray observatories, X-ray absorption variability may become a powerful, relatively standard way to directly measure the physical properties of the absorbing BLR clouds.

One of the most direct consequences of the presence of gas inside the dust sublimation radius is a decrease of the expected dust-to-gas ratio, as measured from the ratio between optical/near-IR reddening, and X-ray column density. This is a well-known observational evidence in nearby Seyfert galaxies (e.g., [50, 51]), which is therefore naturally explained in the context of X-ray absorption by BLR clouds.

The presence and structure of "cometary" clouds poses challenges and possible solutions for long-standing problems in modeling the BLR in AGN. Maiolino et al. [45] estimate that the cloud head loses a significant fraction of its mass through the cometary tail, which is expected to cause the total cloud destruction within a few months. If these clouds are representative of most BLR clouds (or at least the high-ionization ones), this implies that the BLR region must be continuously replenished with gas clouds, possibly from the accretion disk. Such a "dynamical equilibrium" scenario would solve the problem of the long-term stability of BLR clouds, for which a convincing solution has not been found yet.

A further possible interesting aspect of the "cometary" structure (not yet explored in the literature) is the reduction of the number of clouds needed to explain the smoothness of the profiles of the broad absorption lines. High-quality observations of the profile of optical broad emission lines suggest that the minimum number of clouds needed to reproduce the observed smoothness is quite high: one of the most striking examples is that of NGC 4151, with an estimated number of at least 10^8, based on high S/N *Keck* spectra [52]. Considering that the estimated black hole mass of NGC 4151 is of the order of $10^7\,M_\odot$, and that the distance of BLR clouds is of the order of 10^4 gravitational radii, a huge

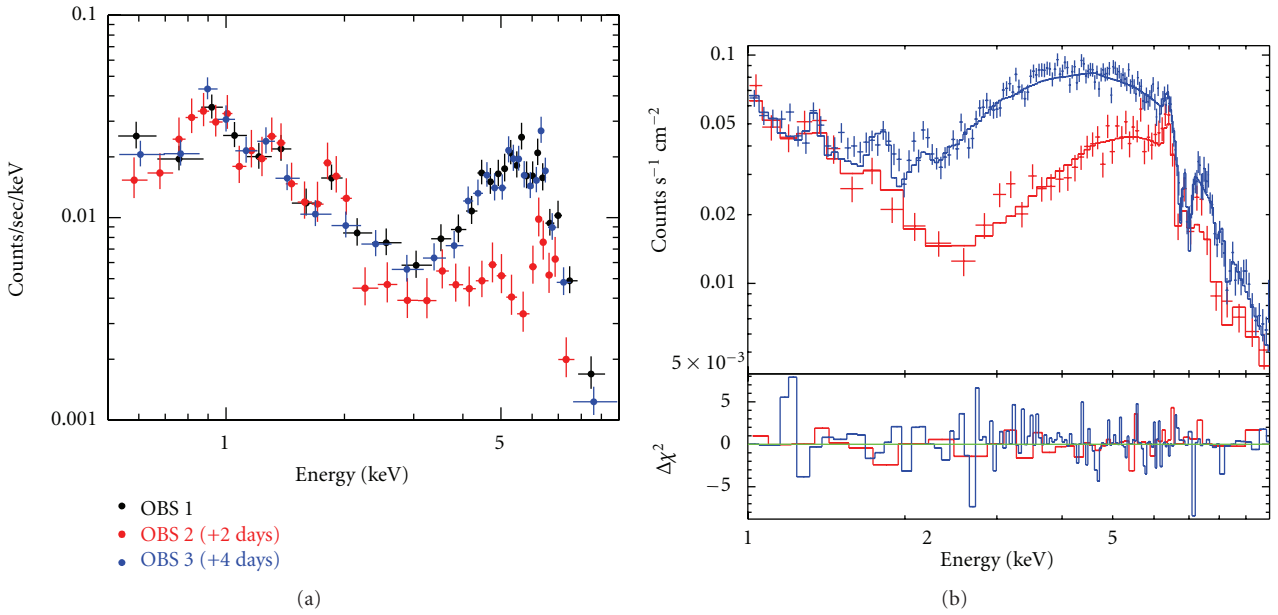


Figure 1: Spectra of NGC 1365, showing X-ray occultations due to BLR clouds: (a) spectra obtained from three Chandra snapshot observations performed every two days [43]. The reflection-dominated spectrum in the second observation is due to an occultation by a Compton-thick cloud (the only one ever detected in short time scales). (b) Spectra obtained from different intervals (length of about 10ks) of a single Suzaku observation, revealing absorption changes due to a cloud with a column density of the order of 10^{23} cm^{-2}.

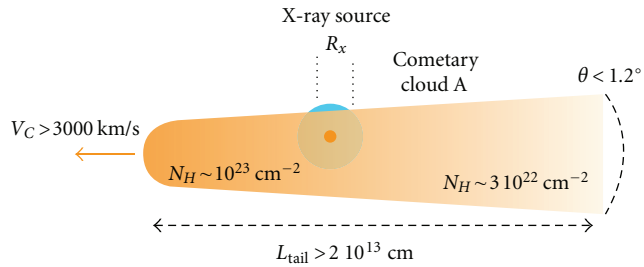


Figure 2: Structure of the absorbing cloud as obtained from a Suzaku observation of NGC 1365. The estimates are based on the hypothesis of Keplerian motion, and on a black hole mass of 2×10^6 M$_\odot$ [45]. The cloud size is not in the correct scale: the tail is much longer when compared with the source size, which is of the order of a few 10^{11} cm.

cloud density is obtained, with a nearly complete occupation of the available volume. If instead every single cloud has a cometary structure, analogous to that inferred for NGC 1365, each cloud would contribute to the observed profile with a small, but not null, width, which would greatly reduce the total number of clouds in order to reproduce the same smoothness.

Finally, if the covering factor and the optical depth of the BLR are large enough, a significant fraction of the iron Kα line should be produced there. In order to test this hypothesis, NGC 7213 represents a unique opportunity. The X-ray spectrum of this source has no evidence for Compton reflection, a unique result among bright Seyfert 1s [53–55]. The observed neutral iron Kα line, therefore, cannot be produced in a Compton-thick material, like the disc or the torus. Indeed, the iron line is resolved in a *Chandra* High-Energy Transmission Grating observation, with a measured FWHM which is in perfect agreement with the value measured for the broad component of the Hα in a simultaneous optical observation [56]. Moreover, the observed equivalent width (EW) of the iron line is in agreement with an origin in the BLR, under reasonable assumptions on the geometrical distribution of the clouds, their covering factor, and their column density [56]. Therefore, NGC 7213 is the only Seyfert 1 galaxy whose iron Kα line is unambiguously produced in the BLR. It is difficult to reach similar conclusions for other objects, because of the presence of the Compton reflection component never allows us to exclude the contribution, from the more extended torus, and because it is difficult to measure the FWHM of the iron line and compare it with the optical lines (see also next section). In the future, high-resolution X-ray spectroscopy with microcalorimeters and X-ray reverberation studies will be extremely powerful in tackling this issue.

2.2. Absorption form Pc-Scale Tori. Early evidence for a circumnuclear dusty medium on parsec, or subparsec scales, as initially invoked by the Unified Model, was obtained from near-IR studies, which revealed the presence of very hot dust, close to the sublimation temperature, in the nuclei of Sy1s [57–59]. The dust sublimation radius in Seyfert nuclei is on subparsec scales and on parsec scales at quasar luminosities. The subparsec location of the hot dust emitting in the near-IR has been confirmed by extensive reverberation observational campaigns [60], which also confirmed the expected $L^{1/2}$ dependence of the sublimation radius. The covering factor of the circumnuclear dusty medium, inferred from the near-IR observations, is very high in most of the nearby Sy1s (exceeding 0.8: [61, 62]), and generally in agreement with the observed type 2/type 1 ratio [26].

Radio *VLBI* observations were the first ones to effectively image the AGN circumunclear medium on parsec and subparsec scales. Greenhill et al. [63] obtained *VLBA* water maser images at subparsec resolution of NGC 1068, revealing a rotating warped disk structure. The warping of the maser disk orientation may indicate that the inclination of the nuclear molecular disk is likely responsible for the large covering faction, rather than a geometrically thick torus [36]; however, one must also take into account that the maser emission does not necessarily trace the global morphology of the circumnuclear molecular gas, but only the equatorial edge-on medium (within $\sim 15^\circ$ from the line of sight) where maser amplification is highest. *VLBI* water maser observations were subsequently obtained for other AGN, finding similar structures [64–67]. Gallimore et al. [68] obtained *VLBA* images of the nuclear continuum radio emission of NGC 1068. Beside the nuclear nonthermal emission, tracing the nuclear engine, they resolved two symmetric radio free-free blobs at a radius of ~ 0.3 pc, which were interpreted as the inner ionized edge of the obscuring torus.

The possibility to effectively "image" the dusty component of the parsec-scale torus has become possible in 2004, when mid-infrared interferometry allowed Jaffe et al. [69] to map for the first time the dust at parsec resolution in the Seyfert 2, NGC 1068. Their results, refined by following observations [70], are consistent with a two-component dust distribution: an inner (0.5 pc of thickness), rather elongated hot ($T > 800$ K) component, and a more extended (3-4 pc), less elongated colder ($T \simeq 300$ K) component. The compact component is coincident, in size and orientation, with the nuclear water maser. Most of the absorption appears to be located outside 1 pc. A similar result was found, with the same technique, for another Seyfert 2, Circinus: again two components, an inner and more compact (0.4 pc), and an outer (2 pc) component [71]. However, the temperature of the inner component in Circinus ($T = 330$ K) is significantly lower than in NGC 1068, and far from the sublimation temperature (see Figure 3). The first observation carried on a type 1 object, NGC 4151, led to results in agreement with the previous ones on Seyfert 2s [72].

When such interferometric studies were performed on a sizeable sample of objects in mid- and near-infrared, it was observed that no significant differences are found between type 1 and 2 sources and the size of the dusty emitter scales with the square root of the luminosity [73–75]. Comparisons with tori models suggest that, in principle, it would be possible to disentangle between face-on and edge-on distributions by comparing the compactness of the dusty structure to the AGN luminosity, but uncertainties on the observed measures are still too large.

The presence of Compton-thick neutral material with large covering factor in the environment of AGN is also supported by the ubiquitous presence of the iron Kα line and the Compton reflection component in the X-ray spectra of Seyfert galaxies (e.g., [54, 76, 77]). Although a component broadened by strong gravity effects arising in the accretion disk is observed in at least a third of the sources (see, e.g., [78, 79]), a narrow core of the iron line is a much more common feature. The line, typically unresolved (with upper limits of several thousand km/s for its FWHM), must be produced far from the nucleus, either in the BLR, the torus, or the NLR. Apart from single exceptional cases (like NGC 7213: see previous section), current X-ray satellites allow us to resolve its FWHM only in a few objects and with limited information, generally leading to inconclusive estimates on the location of the material producing the lines (see, e.g., [80, 81]). Future X-ray missions taking advantage of microcalorimeters will represent a breakthrough in this kind of analysis, allowing us to deconvolve all the components possibly present in the iron line, as routinely performed for the optical lines.

However, clues in favour of a parsec-scale distance of the material producing the narrow iron line and the Compton reflection component already come from the lack of variability of these features. X-ray spectra of Compton-thick sources (i.e., obscured by a column density larger than $\simeq 10^{24}$ cm^{-2}) are completely dominated by reflection features, and they typically do not show any variability even on long time scales. This is particular clear in sources where the central engine fades away for a long time interval (years), while the reflection component (including the Fe Kα line) remains stable over the same time scale [82, 83]. This suggests that the obscuration/reflection occurs on (at least) pc-scale, like the standard torus envisaged in the Unification Models and mapped by interferometry. In principle, the geometry and distance of the torus could be estimated by doing accurate X-ray reverberation analysis of the iron line and the Compton reflection component, in order to take into account in detail how the material reacts to the intrinsic variability of the central source. Unfortunately, such a study is extremely difficult and uncertain with current X-ray missions.

While X-ray absorption variability studies have delivered exquisite information on the structure of the absorbing medium on scales of the BLR, in particularly for what concerns the clumpiness of the absorber, the same kind of analysis is difficult to perform on the parsec scale absorber, both because the time scale for variation is much longer, and because it is totally overwhelmed by the variability introduced by the clouds in the BLR. However, as mentioned in the introduction, recent models have shown that even for the dusty pc-scale torus a clumpy structure (Figure 4)

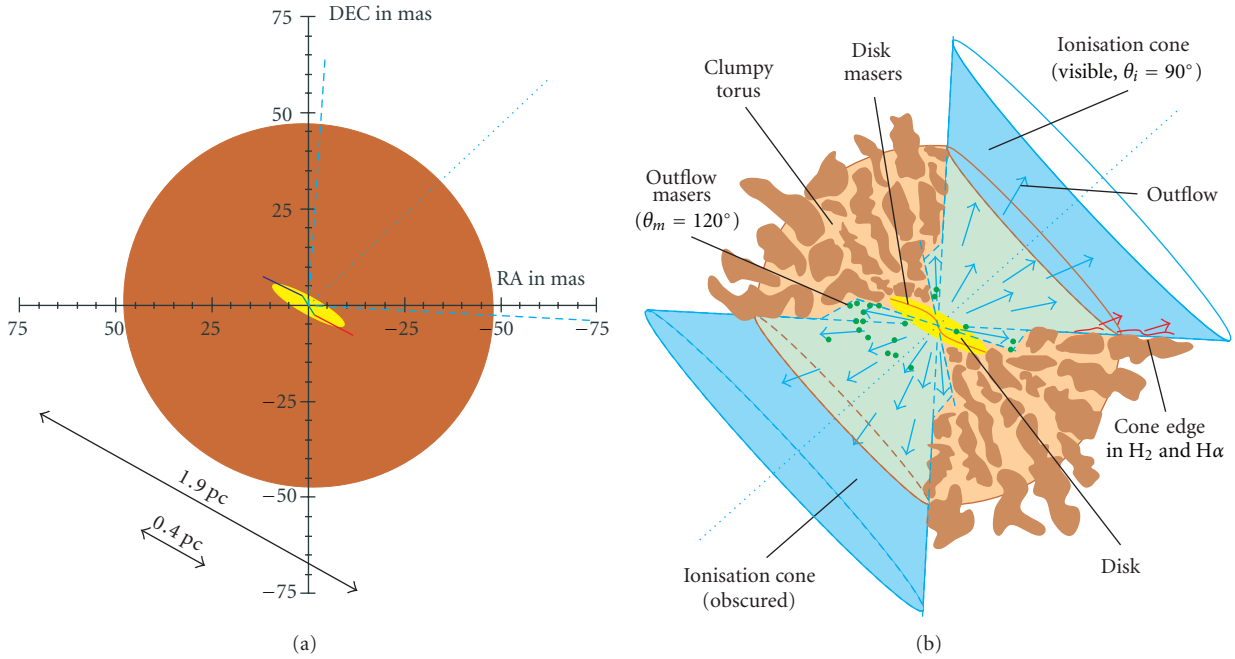


FIGURE 3: The Circinus Galaxy (from Tristram et al. [71]). (a): Sketch of the best-fit physical model where a warm ($T \simeq 330$ K) emission region (yellow) is surrounded by a cooler ($T \simeq 300$ K) one (brown). The water maser emitters are over-plotted (blue and red lines), together with the ionised cone (dashed and dotted light blue line). (b): the dusty torus as derived from the MIR interferometric observations.

can better account for the infrared observational properties [38, 84–86]. In particular, the very broad infrared Spectral Energy Distribution (SED) of AGN requires dust at multiple temperature, which is hardly achieved by models with a compact (pc-scale) uniform torus [28]. Large-scale (100 pc) dusty torii [87] can reproduce the broad IR SED, but can be hardly reconciled with the small sizes observed in mid-IR interferometric observations. A pc-scale, but clumpy torus can at the same time match the observed mid-IR size and reproduce the broad range of dust temperatures, since within each dense clump dust does span a wide range of temperatures. A crucial test of these models will be feasible with *ALMA*, which will allow us to image the cooler dust thermal emission of torus at sub-mm wavelengths at sub-pc resolution. Indeed, as discussed in Maiolino [88], a clear prediction of this model is that the morphology of the "torus" at far-IR/sub-mm wavelengths (tracing cold dust) should be very similar to the morphology observed in the mid-IR (tracing warm dust). An independent observational indication of the clumpiness of the dusty absorber is, as discussed in Shi et al. [89] and in Nikutta et al. [90], the large scatter of the depth of the 9.7 μm silicate absorption feature (directly tracing dust absorption at mid-IR wavelengths) as a function of the X-ray gaseous column density, as well as the finding that the same feature is observed in emission in some type 2 AGN.

2.3. Absorption by Gas in the Host Galaxy. Although clear evidence in favour of absorption from the BLR and the pc-scale "torus" has been presented in the previous sections, in some cases the lowest column densities are consistent with

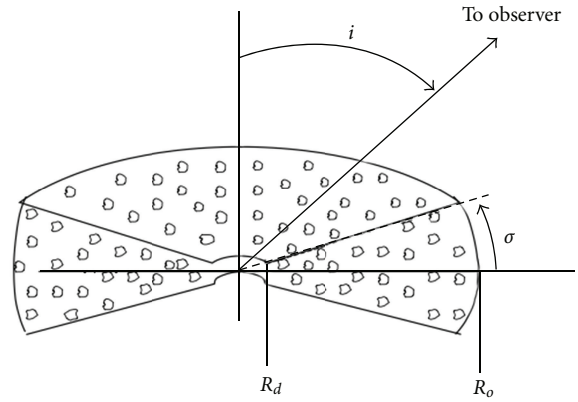


FIGURE 4: Sketch of the dusty torus as in the model of Nenkova et al. [38], which can reproduce the observed dust temperature distributions derived from infrared observations.

the optical reddening associated with the medium in the host galaxy, which is therefore an additional ingredient that must therefore be considered in a global Unification Model (see, e.g., [26, 91]). Early evidence of obscuration on 100 pc scales by the host galaxy gaseous disk came from the finding that optically selected AGN samples tend to avoid edge-on systems [26], a result which has been confirmed and refined with much higher statistics by using the *SDSS* survey [92]. Furthermore, it has been suggested that the gas in the host galaxy disk can partially obscure also the NLR.

Further direct evidence for obscuration on "large" scales was obtained through high-resolution *HST* images, showing

that dust lanes at distances of hundreds of parsecs are very common in Seyfert galaxies [93]. The presence of these structures is correlated with Compton-thin X-ray obscuration, even if not necessarily being directly responsible of the obscuration of the nucleus [94]. In some cases, the effect of dust lanes can be seen directly as X-ray obscuration towards the soft X-ray emission from the NLR (e.g., [95]) see Figure 5.

Interferometric maps of the molecular gas distribution have provided additional evidence for large amount of dense gas on the 100 pc scale surrounding AGN, which certainly contributes to the obscuration of the central engine along some line of sights (e.g., [96–98]). Certainly, within this context, the advent of ALMA is going to be a breakthrough by providing detailed maps of the molecular gas distribution in the circumnuclear region of many AGN.

As it will be discussed further below, there is evidence that the role of absorption from gas in the host galaxy becomes increasingly important in high-z AGN [99, 100], likely as a consequence of the higher gas content and higher star formation in high-z AGN hosts.

It is worth reminding that the obscuration occurring on such large scales is limited by dynamical mass constraints. Risaliti et al. [101] showed that Compton-thick gas must be contained on scales significantly smaller than 100 pc, in order not to exceed the dynamical mass in the same region and have a covering factor large enough to account for the high number of observed Compton-thick sources. This in turns means that the bulk of the ubiquitous Compton reflection component and narrow neutral iron Kα line must also come from a compact region. However, in NGC 1068 hard X-ray emission (mostly associated to the reflection component) and the neutral iron line are seen extending up to $\simeq$ 2 kpc from the nucleus [102].

3. Open Issues

3.1. "True" Seyfert 2s. As discussed in Section 1, one of the most convincing pieces of evidence in favor of the Unification Model is the detection of broad optical lines in the polarized spectra of type 2 AGN. However, about half of the brightest Seyfert 2 galaxies appear not to have hidden broad-line regions in their optical spectra, even when high-quality spectropolarimetric data are analysed [9, 10].

A number of these cases may be associated with nuclei where the mirror reflecting the broad lines either has very low scattering efficiency (either due to low covering factor or low column density) or is obscured [103]. Evidence was also obtained that the lack of polarized broad lines is associated with a stronger contribution/dilution from the host galaxy or from a circumnuclear starburst, making the detection of polarized broad lines harder [104, 105].

However, a number of Sy2s without polarized broad lines may be genuine type 2 Seyferts, in the sense that they intrinsically lack a BLR. Indeed, observational evidence suggests that Seyfert 2s with polarized broad lines are more easily associated with truly obscured Seyfert 1 nuclei, while Seyfert 2s without polarized broad lines preferentially host weak AGN, possibly incapable of generating a classical BLR [9, 10].

In a model proposed by Nicastro [106], the creation of the BLR is connected with disk instabilities occurring in proximity of a transition radius at which the accretion disk changes from gas-pressure dominated to radiation-pressure dominated. Since this transition radius becomes smaller than the innermost stable orbit for very low accretion rates (and therefore luminosities), very weak AGN are expected not to be surrounded by a BLR. More recently, the disappearance of the BLR in low-luminosity AGN has been predicted by Elitzur and Ho [107] and Trump et al. [108] within a disk-wind scenario, where the BLR is embedded in an outflowing wind, which is no longer supported by the disk below a critical value of the Eddington ratio.

If the BLR cannot form in weakly accreting AGN, we expect the existence of "true" Seyfert 2 galaxies, that is, optically classified Type 2 objects, without any evidence of obscuration of their nuclei. Such unabsorbed Seyfert 2 galaxies do exist, and the best examples (where the lack of the optical broad lines and of the X-ray obscuration are unambiguously found in simultaneous observations with high SNR) have low Eddington rates: NGC 3147 ($\log L_{\rm bol}/L_{\rm Edd} \simeq -4$: [109]), Q2131427 ($\log L_{\rm bol}/L_{\rm Edd} \simeq -2.6$: [110]), and NGC 3660 ($\log L_{\rm bol}/L_{\rm Edd} \simeq -2$: [111, 112]). When a sizeable catalogue of X-ray unobscured radio-quiet AGN is analysed, these sources are among those with the lowest accretion rates (e.g., [113]). It is interesting to note that, although lacking the BLR, these objects do appear to have a face-on Compton-thick torus, as type 1 sources, as evidenced by the presence of silicate emission features in their IR spectra [114], and of a neutral iron emission line [109].

The existence of a critical threshold in luminosity and Eddington rate has been confirmed on observational grounds (e.g., [113, 115–118]). Below this threshold (whose exact value depends on the adopted sample, and the methods to derive the bolometric luminosities), no broad lines are detected (either in total or polarized light). On the other hand, it is clear that above the threshold the BLR still cannot be detected in many sources (see Figure 6). If the scenario proposed by Nicastro [106] is correct, these sources should possess a BLR, so there must be something that prevents us from observing it. Indeed, all these sources are Compton-thick, so the nucleus is severely obscured by intervening absorbers. It was suggested that this could be explained within the framework of standard Unification Models, whereby more inclined sources (with respect to the line of sight) should intercept a larger column density of the torus and may obscure the medium responsible for the scattering of the BLR photons (e.g., [117, and references therein]).

3.2. Disk-Torus Alignment. Although not explicitly required by any Unification Model, the most natural assumption on the geometry of the circumnuclear matter in AGN is that it is coaxial with the spin of the BH. This expectation is based on an angular conservation argument: if the obscuring

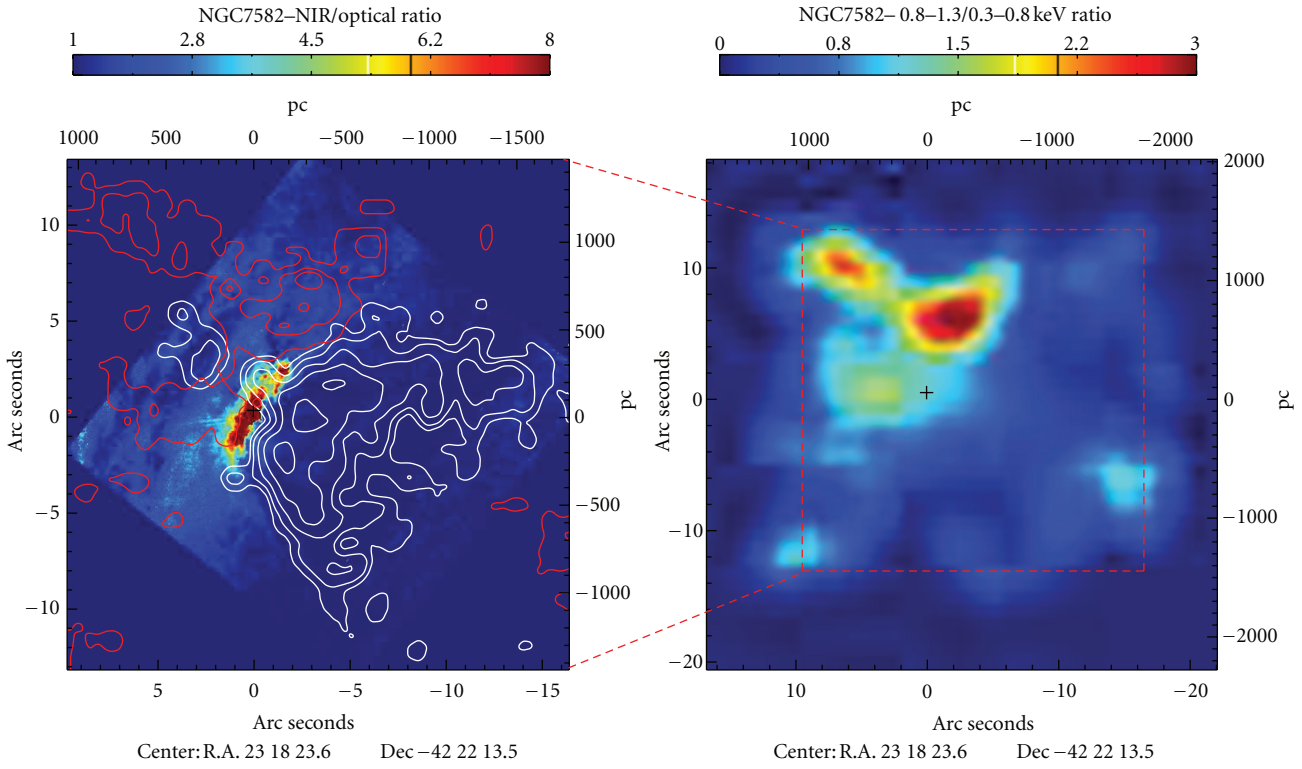


Figure 5: The Seyfert 2 galaxy NGC 7582 (from Bianchi et al. [95]). (a) *HST* NIR to optical ratio, mapping the amount of dust in the circumnuclear region of NGC 7582. The white contour plots refer to the *Chandra* emission below 0.8 keV, while the red ones the ratio shown in the right panel. (b) *Chandra* 0.8–1.3 to 0.3–0.8 keV ratio. The red contour plots shown in left panel refer to this image, but the scale is different, as outlined by the broken-line rectangle. In both panels, the black cross indicates the position of the nucleus.

torus is related to the inflowing material, it is natural to expect that the torus, the accretion disk, and the black hole rotation (mostly due to the angular momentum of the accreting material) share the same axis. This hypothesis, though reasonable, may not be verified in several more complex scenarios. For example, if the BH growth is due to multiple, unrelated accretion events, the actual BH spin may not reflect the rotation axis of the accretion disk. Another possibility is that the obscuring torus is not within the gravitational sphere of influence of the BH (e.g., a galactic dust lane). In this case, no obvious physical relation is expected between the torus axis and the BH spin. If the torus-BH spin alignment hypothesis is accepted, it implies that the accretion disk is aligned with the obscuring torus. Any radio jet should be aligned to the same axis. Finally, the orientation and the opening angle of the NLR ionization cones are collimated by the inner aperture of the torus, thus being themselves coaligned with the common disk/torus axis. These expectations are very difficult to test, due to the very small scales of the inner regions of AGN. When larger scales are imaged, at least the radio jet and the optical/X-ray NLR appear to be generally in agreement with this simple picture (e.g., [119, 120]).

The mid-infrared interferometric studies described in Section 2.2 allowed us for the first time to directly image the geometry of the torus with respect to the optical cones. Surprisingly, the results obtained by Raban et al. [70] strongly suggest that the two structures are misaligned in NGC 1068 (Figure 7). Moreover, the direction of the radio jet is also clearly tilted with respect to both the NLR and the torus. Some of the discrepancy can be solved by taking into account the detailed kinematics of the outflow, and by assuming a clumpy torus that could prevent the ionization of all the gas present in the geometrical opening angle of the ionization cones. However, these solutions are not in agreement with the results inferred from the appearance of the cones in the infrared [70, and references therein]. Similar analysis on other sources is clearly needed in order to shade some light on this issue. A promising, independent, method to test the torus/ionization cones misalignment is via X-ray polarimetry, but we will have to wait for a future X-ray mission equipped with a broadband polarimeter [121].

It is more difficult to estimate the inclination angle of the accretion disk. One possible method is via the relativistic profile of iron emission lines produced in the inner regions of the disk, which is expected to be strongly dependent on the inclination angle (e.g., [122]). When sizeable samples of AGN are systematically analysed, it appears that a simple relation between the inclination of the nuclear obscuring matter (as measured by the optical type) and that of the accreting matter should be ruled out, in contrast with the naive expectations from Unified Models [123]. However, as

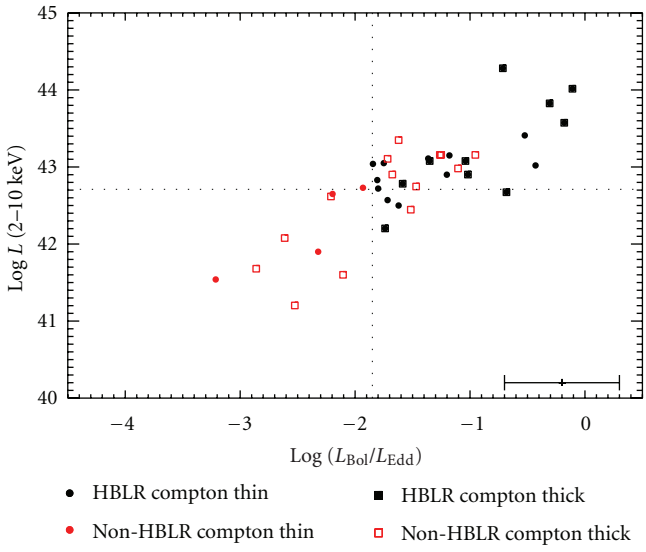


Figure 6: The separation between Seyfert 2 with detection of polarized broad lines (hidden BLR, HBLR) and Seyfert 2s without polarized broad lines (non-HBLR) is clear both in X-ray luminosity and Eddington rate, when only Compton-thin sources are taken into account. However, a number of Compton-thick Non-HBLR sources have high accretion rates and luminosity, suggesting that obscuration also plays an important role in the detection of the HBLR (from Marinucci et al. [113]).

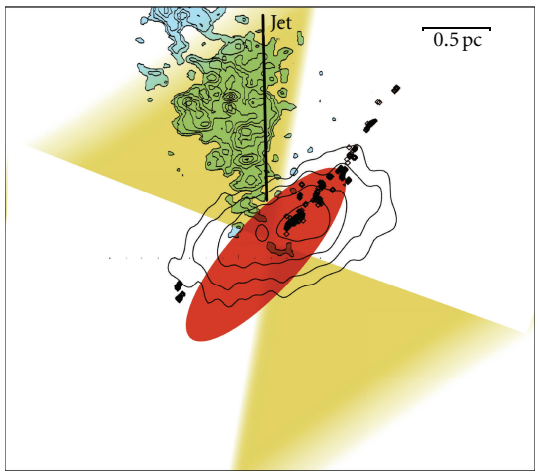


Figure 7: Sketch of the geometry of the circumnuclear matter in NGC 1068: compact dust inferred from mid-infrared interferometry (red), ionization cones as mapped by the [O III] emission (blue, reduced in scale by a factor ~100) and as inferred from spectroscopy (yellow). From Raban et al. [70].

noted by the same authors, the inclination angles derived from the profile of a relativistically broadened line are still affected by large systematic uncertainties. Another method was recently suggested by Risaliti et al. [124], who analysed the distribution of the equivalent widths of the [O III] emission line in a large sample of AGN. Their results are again not compatible with the presence of a torus coaligned with the accretion disk, unless the torus covering factor is extremely small.

Finally, evidence of a misalignment between disk and radio jet for a few individual sources comes from the analysis of the jet-disk relation in objects where the disk inclination can be inferred from maser emission. A prominent example of this kind of studies is that of NGC 3079 [67].

3.3. Luminosity and Redshift Dependence of the Covering Factor. An interesting finding of the recent years has been the evidence for a dependence of the obscuring medium covering factor on the AGN luminosity. More specifically, the covering factor of the obscuring medium appears to decrease significantly with luminosity. This effect has been shown quite clearly by various hard X-ray studies [125–130] and by optical surveys [131], which have measured the relative fraction of obscured and unobscured AGN as a function of the bolometric luminosity. Clearly these works are affected by various uncertainties and caveats, primarily related to possible incompleteness effects and biases that may prevent the identification of obscured AGN in more distant galaxies (hence more luminous, as a consequence of the Malmquist bias). Indeed, these results have been questioned by some authors [132–134]. However, more recent extensive hard X-ray surveys have further confirmed a clear trend for a decreasing fraction of obscured AGN at high luminosity [135–138]. Figure 8 shows the results of some of these studies [135], where black and blue solid symbols show the fraction of obscured AGN as a function of X-ray luminosity, while red-dotted symbols illustrate the same trend from optical surveys [131].

The luminosity dependence of the covering factor has been recently questioned by Lawrence and Elvis [36], showing that, at least in optical and IR-selected sample, the luminosity dependence is largely a consequence of the inclusion of low excitation AGN (which may have an intrinsically different engine) and on the definition of "obscured" AGN (i.e., whether including mildly obscured AGN or not). However, the luminosity dependence of the covering factor still persists in the X-ray-selected samples, regardless of the classification scheme.

An alternative method to investigate the covering factor of the obscuring medium is by means of the dust reprocessing by the obscuring medium. More specifically, the ratio between the hot dust emission observed in the near/mid-IR and the primary AGN bolometric emission (optical/UV/X-ray), responsible for heating the dust, is proportional to the covering factor of the obscuring medium. This method can only be applied to type 1 AGN, where the primary optical/UV radiation is directly detectable. By using *ISO* and *Spitzer* data to trace the dust emission, various studies have confirmed that the covering factor of the absorbing medium decreases as a function of luminosity, as shown in Figure 8 [62, 139, 141, 142] (but see also [143]). These IR results obtain a covering factor significantly higher than X-ray studies, but this is likely due to the population of Compton-thick AGN that is mostly lost by X-ray surveys.

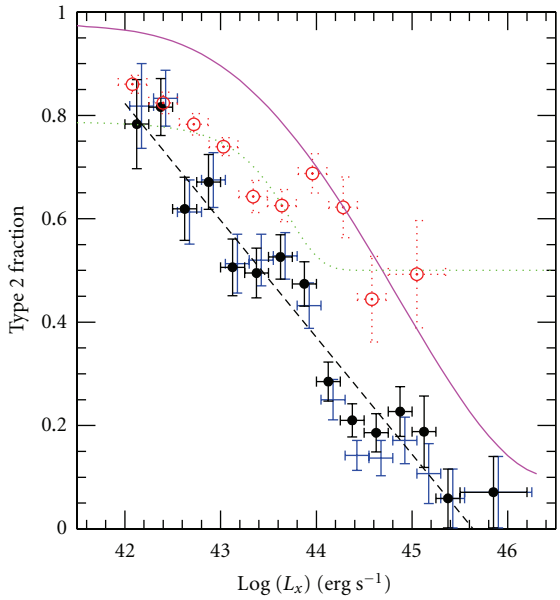


Figure 8: Fraction of type 2 AGN as a function of X-ray luminosity (from Hasinger [135]). The solid symbols show hard X-ray-selected AGN (blue and black symbols illustrate different subsamples in terms of redshift completeness). The red-dotted symbols show optically selected AGN (from Simpson [131]). The magenta line shows the type 2 fraction inferred from the dust covering factor obtained through near/mid-IR observations (from Maiolino et al. [139], note that these include also Compton-thick AGN, which are not present in X-ray samples). The green-dotted line shows the expected fraction of Compton-thin type 2 AGN according to the X-ray background synthesis models given in Gilli et al. [140].

In a similar manner, the equivalent width of the (narrow) Fe Kα line has been used as a proxy of the covering factor of the circumnuclear material responsible for producing the Fe Kα line relative to the primary X-ray continuum. The EW of the Fe Kα has also been found to anticorrelate with luminosity ("X-ray Baldwin effect" or "Iwasawa-Taniguchi effect"), a trend which was generally interpreted in terms of decreasing covering factor of the circumnuclear absorbing medium as a function of luminosity [144–148].

Therefore, although partly questioned, several independent observational studies strongly favor a luminosity dependence of the covering factor of the circumnuclear absorbing medium. The origin of the anticorrelation between luminosity and covering factor is unclear. The "receding torus" scenario [149] is often invoked to explain this trend: higher luminosities imply a larger dust sublimation radius and, if the torus has a constant height as a function of radius, this results into a smaller covering factor of the dusty medium. However, this scenario cannot explain the results on the decreasing covering factor inferred from X-ray studies, which do not trace the dusty component of the absorber.

In the model proposed by Lamastra [150], the covering factor-luminosity anticorrelation naturally arises if the X-ray obscuration in Compton-thin sources is due to interstellar gas, distributed in a rotationally supported disk with an extension of a few hundred pc. The covering factor of this disc diminishes as the gravitational pull from the central SMBH and the bulge increases with the BH mass (and, therefore, the luminosity), producing the observed anticorrelation. However, this model can explain the anticorrelation only for Compton-thin sources.

Another possible scenario is that the lower covering factor in luminous AGN (quasars) is simply a consequence of the stronger AGN radiation pressure impinging onto the circumnuclear medium and expelling larger fractions of material. In support of this scenario growing evidence for massive outflows in luminous AGN has been reported in the recent years [151–155].

Some studies have also claimed an evolution of the covering factor with redshift. More specifically, by using large samples of X-ray selected objects, La Franca et al. [127] and Hasinger [135] have found evidence for a strong increase of the fraction of obscured AGN from $z = 0$ to $z = 2$. This result is however more debated, due to the larger uncertainties in disentangling luminosity and redshift effects in flux-limited samples and also more critically subject to selection effects in high-z objects. Indeed, other authors did not find significant redshift dependence [125, 140], while Ballantyne et al. [156] and Treister and Urry [157] found a shallow redshift dependence of the fraction of obscured AGN. Here we only mention that an increase of the fraction of obscured AGN at high redshift is naturally expected by the larger gas content found in high-z galaxies [158, 159] and associated with the enhanced star formation rates in high-z galaxies, also observed in high-z AGN host galaxies [160, 161].

A possible additional complication to the simple version of the Unified Model is that the covering factor of the obscuring medium may not only depend on the AGN luminosity, and possibly on redshift, but may also be intrinsically different between type 2 and type 1 AGN. Ricci et al. [162] have shown that the X-ray cold reflection "hump" in type 2 AGN is stronger than in type 1 AGN, suggesting a larger covering factor of the circumnuclear medium in the former class. This result is, however, at odds with those obtained by mid-IR interferometry studies, as discussed above. More data are certainly required to tackle the discrepancy.

3.4. Unusual Geometries. While the general picture discussed in this paper can account for most of the observational properties observed in AGN, some subsamples require different geometries. This is the case for a number of nuclei that are unambiguously hosting a relatively powerful AGN, based on their X-ray or mid-IR properties, but which do not show any evidence for classical NLR tracers in their optical spectra [163–173]. A possibility is that in these objects the NLR, although extended, is heavily obscured by absorbing medium distributed on large scales in the host galaxy. Indeed, in a number of objects *Spitzer* spectroscopic observations have revealed mid-IR high excitation lines (e.g., [NeV] 14.3 μm) typical of the NLR [174–176].

However, a number of active nuclei do not show NLR-like lines even in the mid-IR [177]. In these cases a likely scenario

is that the nuclear engine is obscured in all directions (4π obscuration) on small (~pc) scales, so that UV photons cannot escape to produce an NLR. Further observational indications in favour of 4π obscuration is the very small amount of reflection component at energies below 10 keV in some Compton thick AGN e.g. [178, 179]. The most likely explanation in this case is that the Compton thick absorbing medium totally covers the nuclear sources, including the Compton reflecting medium.

If totally "buried" AGN may be one extreme of the population, on the other extreme there is growing evidence for a population of (type 1) AGN missing the circumnuclear absorber. Jiang et al. [180] discovered two quasars at $z \sim 6$ showing no indication of hot dust emission typically observed in AGN. They interpreted this result as evidence for young quasars that had no time yet to form dust at such early epochs. However, evidence for AGN with little or no circumnuclear hot dust emission has been found also at lower redshifts [142, 181, 182], in evolved quasars, where time scale for dust formation is likely not an issue. The nature of this class of objects is still unclear. Possible scenarios invoke dust destruction (dynamically or by radiation) or the AGN is not centered onto the supermassive black hole, as it may happen in the case of BH recoils in merging events, which may result in the off-nuclear AGN to be still surrounded by an accretion disk and a BLR, but not from pc-scale dusty medium [182].

4. Summary and Conclusions

The Unified Model for AGN has been tested in many different ways in the past few years, through a large set of new imaging, spectral, and timing observations. Overall, the fundamental aspect of the model, that is, that nonspherically symmetric absorption plays a major role in explaining the differences in the observed features among AGN, has been confirmed, and even reinforced by the most recent observations.

However, some more complexity has been added. In particular, the existence of the standard "torus", in the generic sense of an axisymmetric, rather than spherically symmetric, circumnuclear absorber has been confirmed, but its physical and geometrical structure has been proven to be far from homogeneous among the AGN family. There is now strong evidence of at least three absorption components on very different scales:

(i) on scales of hundreds of parsecs, or even larger (e.g., galactic dust lanes), circumnuclear tori have been imaged, with different techniques, and are clearly responsible of the "type 2" (in optical/UV) or "absorbed" (in X-rays) classification of a significant fraction of AGN;

(ii) on the parsec scale, and down to the dust sublimation radius, the "standard" torus, as initially postulated in the earliest works on AGN unification, has been now directly imaged in a few sources with interferometric techniques, and its presence is suggested by X-ray reflection properties, and by dust reverberation mapping in the near-IR;

(iii) on the 0.01 pc scale, the presence of dust-free gas along the line of sight has been demonstrated through X-ray absorption variability in several AGN, thus suggesting that part of the observed X-ray absorption is due to Broad Line Region clouds.

From the latter study, and from models of dust reemission in the infrared, it is also clear that the absorber has a clumpy, rather than homogeneous, structure.

Several "open issues" have been also briefly discussed in this paper, showing that not all AGN fit in a simple unification scheme, and that even within the general boundaries of the model, more work and new observations are needed to fully understand the dependence of all the observed properties on the physical parameters of the central source and of its environment.

Acknowledgments

The authors would like to thank the anonymous referees for useful suggestions, and R. Antonucci, M. Elvis, G. Matt, K. Tristram for feedback on the first version of the paper.

References

[1] R. Antonucci, "Unified models for active galactic nuclei and quasars," *Annual Review of Astronomy and Astrophysics*, vol. 31, no. 1, pp. 473–521, 1993.

[2] R. R. J. Antonucci, "Optical spectropolarimetry of radio galaxies," *The Astrophysical Journal*, vol. 278, pp. 499–520, 1984.

[3] C. K. Seyfert, "Nuclear emission in spiral nebulae," *The Astrophysical Journal*, vol. 97, p. 28, 1943.

[4] R. R. J. Antonucci and J. S.. Miller, "Diffuse pulmonary hemorrhage: a review and classification," *The Astrophysical Journal*, vol. 297, pp. 621–632, 1985.

[5] J. H. Krolik and M. C. Begelman, "An X-ray heated wind in NGC 1068," *The Astrophysical Journal*, vol. 308, pp. L55–L58, 1986.

[6] J. H. Krolik and M. C. Begelman, "Molecular tori in seyfert galaxies—feeding the monster and hiding it," *The Astrophysical Journal*, vol. 329, pp. 702–711, 1988.

[7] S. Kaspi, D. Maoz, H. Netzer, B. M. Peterson, M. Vestergaard, and B. T. Jannuzi, "The relationship between luminosity and broad-line region size in active galactic nuclei," *The Astrophysical Journal*, vol. 629, no. 1, pp. 61–71, 2005.

[8] B. M. Peterson, "Reverberation mapping of active galactic nuclei," *Publications of the Astronomical Society of the Pacific*, vol. 105, pp. 247–268, 1993.

[9] H. D. Tran, "Hidden broad-line Seyfert 2 galaxies in the CfA and 12 μm samples," *Astrophysical Journal Letters*, vol. 554, no. 1, pp. L19–L23, 2001.

[10] S. Veilleux, R.-W. Goodrich, and G. J. Hill, "Infrared spectroscopy of seyfert 2 galaxies: a look through the obscuring torus? II," *The Astrophysical Journal*, vol. 477, article 631, 1997.

[11] S. Veilleux, R.-W. Goodrich, and G. J. Hill, "Infrared spectroscopy of seyfert 2 galaxies: a look through the obscuring

torus? II," *The Astrophysical Journal*, vol. 477, no. 2, pp. 631–660, 1997.

[12] N. M. Nagar, E. Oliva, A. Marconi, and R. Maiolino, "NGC 5506 unmasked as a narrow line Seyfert 1: a direct view of the broad line region using near-IR spectroscopy," *Astronomy and Astrophysics*, vol. 391, no. 2, pp. L21–L24, 2002.

[13] D. Lutz, R. Maiolino, A. F. M. Moorwood et al., "Infrared spectroscopy around 4 mu m of Seyfert 2 galaxies: Obscured broad line regions and coronal lines," *Astronomy and Astrophysics*, vol. 396, pp. 439–448, 2002.

[14] R. Riffel, A. Rodríguez-Ardila, and M. G. Pastoriza, "A 0.8-2.4 μm spectral atlas of active galactic nuclei," *Astronomy and Astrophysics*, vol. 457, no. 1, pp. 61–70, 2006.

[15] H.-B. Cai, X.-W. Shu, Z.-Y. Zheng, and J. X. Wang, "The study of Seyfert 2 galaxies with and without infrared broad lines," *Research in Astronomy and Astrophysics*, vol. 10, no. 5, pp. 427–437, 2010.

[16] D. A. Smith and C. Done, "Unified theories of active galactic nuclei: a hard X-ray sample of Seyfert 2 galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 280, no. 2, pp. 355–377, 1996.

[17] T. J. Turner, I. M. George, K. Nandra, and R. F. Mushotzky, "ASCA observations of type 2 seyfert galaxies. I. Data analysis results," *The Astrophysical Journal Supplement Series*, vol. 113, no. 1, article 23, 1997.

[18] R. Maiolino, M. Salvati, L. Bassani et al., "Heavy obscuration in X-ray weak AGNs," *Astronomy and Astrophysics*, vol. 338, no. 3, pp. 781–794, 1998.

[19] L. Bassani, M. Dadina, R. Maiolino et al., "A three-dimensional diagnostic diagram for Seyfert 2 galaxies: probing X-ray absorption and compton thickness," *Astrophysical Journal, Supplement Series*, vol. 121, no. 2, pp. 473–482, 1999.

[20] G. Matt, M. Guainazzi, F. Frontera et al., "Hard X-ray detection of NGC 1068 with BeppoSAX," *Astronomy and Astrophysics*, vol. 325, no. 2, pp. L13–L16, 1997.

[21] G. Matt, M. Guainazzi, R. Maiolino et al., "One more surprise from the circinus galaxy: beppoSAX discovery of a transmission component in hard X-rays," *Astronomy and Astrophysics*, vol. 341, no. 2, pp. L39–L42, 1999.

[22] R. W. Pogge, "Extended ionized gas in the Seyfert 2 galaxy NGC 4388," *The Astrophysical Journal*, vol. 332, pp. 702–710, 1988.

[23] I. N. Evans, H. C. Ford, A. L. Kinney, R. R. J. Antonucci, L. Armus, and S. Caganoff, "HST imaging of the inner 3 arcseconds of NGC 1068 in the light of [O III] λ5007," *The Astrophysical Journal*, vol. 369, no. 2, pp. L27–L30, 1991.

[24] A. S. Wilson and Z. I. Tsvetanov, "Ionization cones and radio ejecta in active galaxies," *Astronomical Journal*, vol. 107, no. 4, pp. 1227–1234, 1994.

[25] F. K. B. Barbosa, T. Storchi-Bergmann, R. C. Fernandes, C. Winge, and H. Schmitt, "Gemini/GMOS IFU gas velocity 'tomography' of the narrow line region of nearby active galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 1, pp. 2–18, 2009.

[26] R. Maiolino and G. H. Rieke, "Low-luminosity and obscured seyfert nuclei in nearby galaxies," *The Astrophysical Journal*, vol. 454, no. 1, pp. 95–105, 1995.

[27] N. M. Nagar, A. S. Wilson, J. S. Mulchaey, and J. F. Gallimore, "Radio structures of Seyfert galaxies. VIII. A distance- and magnitude-limited sample of early-type galaxies," *Astrophysical Journal, Supplement Series*, vol. 120, no. 2, pp. 209–245, 1999.

[28] E. A. Pier and J. H. Krolik, "Radiation-pressure-supported obscuring tori around active galactic nuclei," *The Astrophysical Journal*, vol. 399, no. 1, pp. L23–L26, 1992.

[29] G. L. Granato, L. Danese, and A. Franceschini, "Thick tori around active galactic nuclei: the case for extended tori and consequences for their X-ray and infrared emission," *The Astrophysical Journal*, vol. 486, no. 1, pp. 147–159, 1997.

[30] J. H. Krolik, "AGN obscuring tori supported by infrared radiation pressure," *The Astrophysical Journal*, vol. 661, no. 1, pp. 52–59, 2007.

[31] K. Wada and K. Tomisaka, "Molecular gas structure around an active galactic nucleus with nuclear starburst: three-dimensional non-LTE calculations of CO lines," *The Astrophysical Journal*, vol. 619, no. 1 I, pp. 93–104, 2005.

[32] Y. Watabe and M. Umemura, "Obscuration of active galactic nuclei by circumnuclear starbursts," *The Astrophysical Journal*, vol. 618, no. 2 I, pp. 649–656, 2005.

[33] S. Nayakshin, "Warped accretion discs and the unification of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 359, no. 2, pp. 545–550, 2005.

[34] A. Caproni, Z. Abraham, and H. J. Mosquera Cuesta, "Bardeen-petterson effect and the disk structure of the seyfert galaxy NGC 1068," *The Astrophysical Journal*, vol. 638, no. 1 I, pp. 120–124, 2006.

[35] A. Caproni, M. Livio, Z. Abraham, and H. J. M. Cuesta, "Warping and precession in galactic and extragalactic accretion disks," *The Astrophysical Journal*, vol. 653, no. 1 I, pp. 112–126, 2006.

[36] A. Lawrence and M. Elvis, "Misaligned disks as obscurers in active galaxies," *The Astrophysical Journal*, vol. 714, no. 1, pp. 561–570, 2010.

[37] M. Elitzur and I. Shlosman, "The AGN-obscuring torus: the end of the "doughnut" paradigm?" *The Astrophysical Journal*, vol. 648, no. 2, pp. L101–L104, 2006.

[38] M. Nenkova, M. M. Sirocky, R. Nikutta, Z. Ivezi, and M. Elitzur, "AGN dusty Tori. II. observational implications of clumpiness," *The Astrophysical Journal*, vol. 685, no. 1, pp. 160–180, 2008.

[39] S. F. Hönig and T. Beckert, "Active galactic nuclei dust tori at low and high luminosities," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 3, pp. 1172–1176, 2007.

[40] S. F. Hönig and M. Kishimoto, "The dusty heart of nearby active galaxies: II. from clumpy torus models to physical properties of dust around AGN," *Astronomy and Astrophysics*, vol. 523, no. 2, article no. A27, 2010.

[41] G. Risaliti, "The BeppoSAX view of bright compton-thin Seyfert 2 galaxies," *Astronomy and Astrophysics*, vol. 386, no. 2, pp. 379–398, 2002.

[42] G. Risaliti, M. Elvis, G. Fabbiano, A. Baldi, and A. Zezas, "Rapid compton-thick/compton-thin transitions in the seyfert 2 galaxy NGC 1365," *The Astrophysical Journal*, vol. 623, no. 2, pp. L93–L96, 2005.

[43] G. Risaliti, M. Elvis, G. Fabbiano, A. Baldi, A. Zezas, and M. Salvati, "Occultation measurement of the size of the X-ray-emitting region in the active galactic nucleus of NGC 1365," *The Astrophysical Journal*, vol. 659, no. 2, pp. L111–L114, 2007.

[44] G. Risaliti, M. Salvati, M. Elvis et al., "The XMM-Newton long look of NGC 1365: uncovering of the obscured X-ray source," *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 1, pp. L1–L5, 2009.

[45] R. Maiolino, G. Risaliti, M. Salvati et al., "Comets orbiting a black hole," *Astronomy and Astrophysics*, vol. 517, no. 6, article A47, 2010.

[46] M. Elvis, G. Risaliti, F. Nicastro, J. M. Miller, F. Fiore, and S. Puccetti, "An unveiling event in the type 2 active galactic nucleus NGC 4388: a challenge for a parsec-scale absorber," *The Astrophysical Journal*, vol. 615, no. 1, pp. L25–L28, 2004.

[47] S. Puccetti, F. Fiore, G. Risaliti, M. Capalbi, M. Elvis, and F. Nicastro, "Rapid NH changes in NGC 4151," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 2, pp. 607–616, 2007.

[48] S. Bianchi, E. Piconcelli, M. Chiaberge, E. J. Bailon, G. Matt, and F. Fiore, "How complex is the obscuration in active galactic nuclei? New clues from the suzaku monitoring of the X-ray absorbers in NGC 7582," *The Astrophysical Journal*, vol. 695, no. 1, pp. 781–787, 2009.

[49] G. Risaliti, E. Nardini, M. Salvati et al., "X-ray absorption by broad-line region clouds in Mrk 766," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 2, pp. 1027–1035, 2011.

[50] R. Maiolino, A. Marconi, M. Salvati et al., "Dust in active nuclei I. Evidence for "anomalous" properties," *Astronomy and Astrophysics*, vol. 365, no. 2, pp. 28–36, 2001.

[51] T. Maccacaro, G. C. Perola, and M. Elvis, "X-ray observations of emission line galaxies with the einstein observatory," *Space Science Reviews*, vol. 30, no. 1–4, pp. 61–65, 1981.

[52] T. Maccacaro, G. C. Perola, and M. Elvis, "X-ray observations with the Einstein Observatory of emission-line galaxies," *The Astrophysical Journal*, vol. 257, p. 47, 1982.

[53] S. Bianchi, G. Matt, I. Balestra, and G. C. Perola, "The origin of the iron lines in NGC 7213," *Astronomy and Astrophysics*, vol. 407, no. 2, pp. L21–L24, 2003.

[54] S. Bianchi, G. Matt, I. Balestra, M. Guainazzi, and G. C. Perola, "X-ray reprocessing in Seyfert galaxies: simultaneous XMM-Newton/BeppoSAX observations," *Astronomy and Astrophysics*, vol. 422, no. 1, pp. 65–76, 2004.

[55] A. P. Lobban, J. N. Reeves, D. Porquet et al., "Evidence for a truncated accretion disc in the low-luminosity Seyfert galaxy, NGC 7213?" *Monthly Notices of the Royal Astronomical Society*, vol. 408, no. 1, pp. 551–564, 2010.

[56] S. Bianchi, F. La Franca, G. Matt et al., "A broad-line region origin for the iron Kα line in NGC 7213," *Monthly Notices of the Royal Astronomical Society*, vol. 389, no. 1, pp. L52–L56, 2008.

[57] T. Storchi-Bergmann, J. S. Mulchaey, and A. S. Wilson, "Infrared emission in Seyfert 2 galaxies: reprocessed radiation from a dusty torus?" *The Astrophysical Journal*, vol. 395, no. 2, pp. L73–L77, 1992.

[58] A. Alonso-Herrero, A. C. Quillen, C. Simpson, A. Efstathiou, and M. J. Ward, "The nonstellar infrared continuum of seyfert galaxies," *Astronomical Journal*, vol. 121, no. 3, pp. 1369–1384, 2001.

[59] E. Oliva, L. Origlia, R. Maiolino, and A. F. M. Moorwood, "Starbursts in active galaxy nuclei: observational constraints from IR stellar absorption lines," *Astronomy and Astrophysics*, vol. 350, no. 1, pp. 9–16, 1999.

[60] M. Suganuma, Y. Yoshii, Y. Kobayashi et al., "Reverberation measurements of the inner radius of the dust torus in nearby seyfert 1 galaxies," *The Astrophysical Journal*, vol. 639, no. 1, pp. 46–63, 2006.

[61] R. Maiolino, O. Shemmer, M. Imanishi et al., "Dust covering factor, silicate emission, and star formation in luminous QSOs," *Astronomy and Astrophysics*, vol. 468, no. 3, pp. 979–992, 2007.

[62] E. Treister, J. H. Krolik, and C. Dullemond, "Measuring the fraction of obscured quasars by the infrared luminosity of unobscured quasars," *The Astrophysical Journal*, vol. 679, no. 1, pp. 140–148, 2008.

[63] L. J. Greenhill, C. R. Gwinn, R. Antonucci, and R. Barvainis, "VLBI imaging of water maser emission from the nuclear torus of NGC 1068," *The Astrophysical Journal*, vol. 472, no. 1, pp. L21–L24, 1996.

[64] L. J. Greenhill, J. M. Moran, and J. R. Herrnstein, "The distribution of H_2O maser emission in the nucleus of NGC 4945," *The Astrophysical Journal*, vol. 481, no. 1, pp. L23–L26, 1997.

[65] L. J. Greenhill, P. T. Kondratko, J. E. J. Lovell et al., "The discovery of H_2O maser emission in seven active galactic nuclei and at high velocities in the Circinus galaxy," *The Astrophysical Journal*, vol. 582, no. 1, pp. L11–L14, 2003.

[66] P. T. Kondratko, L. J. Greenhill, and J. M. Moran, "The parsec-scale accretion disk in NGC 3393," *The Astrophysical Journal*, vol. 678, no. 1, pp. 87–95, 2008.

[67] P. T. Kondratko, L. J. Greenhill, and J. M. Moran, "Evidence for a geometrically thick self-gravitating accretion disk in NGC 3079," *The Astrophysical Journal*, vol. 618, no. 2, pp. 618–634, 2005.

[68] J. F. Gallimore, S. A. Baum, and C. P. O'Dea, "A direct image of the obscuring disk surrounding an active galactic nucleus," *Nature*, vol. 388, no. 6645, pp. 852–854, 1997.

[69] W. Jaffe, K. Meisenheimer, H. J. A. Röttgering et al., "The central dusty torus in the active nucleus of NGC 1068," *Nature*, vol. 429, no. 6987, pp. 47–49, 2004.

[70] D. Raban, W. Jaffe, H. Röttgering, K. Meisenheimer, and K. R. W. Tristram, "Resolving the obscuring torus in NGC 1068 with the power of infrared interferometry: revealing the inner funnel of dust," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 3, pp. 1325–1337, 2009.

[71] K. R. W. Tristram, K. Meisenheimer, W. Jaffe et al., "Resolving the complex structure of the dust torus in the active nucleus of the Circinus galaxy," *Astronomy and Astrophysics*, vol. 474, no. 3, pp. 837–850, 2007.

[72] L. Burtscher, W. Jaffe, D. Raban, K. Meisenheimer, K. R. W. Tristram, and H. Röttgering, "Dust emission from a parsec-scale structure in the seyfert 1 nucleus of NGC 4151," *The Astrophysical Journal*, vol. 705, no. 1, pp. L53–L57, 2009.

[73] K. R. W. Tristram, D. Raban, K. Meisenheimer et al., "Parsec-scale dust distributions in Seyfert galaxies Results of the MIDI AGN snapshot survey," *Astronomy and Astrophysics*, vol. 502, no. 1, pp. 67–84, 2009.

[74] K. R. W. Tristram and M. Schartmann, "On the size-luminosity relation of AGN dust tori in the mid-infrared," *Astronomy and Astrophysics*, vol. 531, article A99, 2011.

[75] M. Kishimoto, S. F. Honig, R. Antonucci et al., "The innermost dusty structure in active galactic nuclei as probed by the Keck interferometer," *Astronomy & Astrophysics*, vol. 527, article A121, 2011.

[76] G. C. Perola, G. Matt, M. Cappi et al., "Compton reflection and iron fluorescence in BeppoSAX observations of Seyfert type 1 galaxies," *Astronomy and Astrophysics*, vol. 389, no. 3, pp. 802–811, 2002.

[77] S. Bianchi, M. Guainazzi, G. Matt, N. Fonseca Bonilla, and G. Ponti, "CAIXA: a catalogue of AGN in the XMM-Newton archive: I. Spectral analysis," *Astronomy and Astrophysics*, vol. 495, no. 2, pp. 421–430, 2009.

[78] I. De La Calle Pérez, A. L. Longinotti, M. Guainazzi et al., "FERO: finding extreme relativistic objects: I. Statistics of relativistic Fe Kα lines in radio-quiet Type 1 AGN," *Astronomy and Astrophysics*, vol. 524, no. 2, article no. A50, 2010.

[79] K. Nandra, P. M. O'Neill, I. M. George, and J. N. Reeves, "An XMM-Newton survey of broad iron lines in Seyfert galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 1, pp. 194–228, 2007.

[80] K. Nandra, "On the origin of the iron Ko line cores in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 368, pp. L62–L66, 2006.

[81] X. W. Shu, T. Yaqoob, and J. X. Wang, "Chandra High-energy Grating Observations of the Fe Kα Line Core in Type II Seyfert Galaxies: A Comparison with Type I Nuclei," *The Astrophysical Journal*, vol. 738, p. 147, 2011.

[82] R. Gilli, R. Maiolino, A. Marconi et al., "The variability of the Seyfert galaxy NGC 2992: the case for a revived AGN," *Astronomy and Astrophysics*, vol. 355, no. 2, pp. 485–498, 2000.

[83] G. Matt, S. Bianchi, M. Guainazzi et al., "Chandra and XMM-Newton observations of Tololo 0109-383," *Astronomy and Astrophysics*, vol. 399, no. 2, pp. 519–523, 2003.

[84] M. Nenkova, Z. Ivezi, and M. Elitzur, "Dust emission from active galactic nuclei," *The Astrophysical Journal*, vol. 570, no. 1, pp. L9–L12, 2002.

[85] M. Elitzur and I. Shlosman, "The AGN-obscuring torus: the end of the "doughnut" paradigm?" *The Astrophysical Journal*, vol. 648, no. 2, pp. L101–L104, 2006.

[86] S. F. Hönig, T. Beckert, K. Ohnaka, and G. Weigelt, "Radiative transfer modeling of three-dimensional clumpy AGN tori and its application to NGC 1068," *Astronomy and Astrophysics*, vol. 452, no. 2, pp. 459–471, 2006.

[87] G. L. Granato and L. Danese, "Thick tori around active galactic nuclei—a comparison of model predictions with observations of the infrared continuum and silicate features," *Monthly Notices of the Royal Astronomical Society*, vol. 268, no. 1, p. 235, 1994.

[88] R. Maiolino, "Prospects for AGN studies with ALMA," *New Astronomy Reviews*, vol. 52, no. 6, pp. 339–357, 2008.

[89] Y. Shi, G. H. Rieke, D. C. Hines et al., "9.7 μm silicate features in active galactic nuclei: new insights into unification models," *The Astrophysical Journal*, vol. 653, no. 1 I, pp. 127–136, 2006.

[90] R. Nikutta, M. Elitzur, and M. Lacy, "On the 10 μm silicate feature in active galactic nuclei," *The Astrophysical Journal*, vol. 707, no. 2, pp. 1550–1559, 2009.

[91] G. Matt, "Dust lanes, thick absorbers, and the unification model for Seyfert galaxies," *Astronomy and Astrophysics*, vol. 355, no. 2, pp. L31–L33, 2000.

[92] C. D. P. Lagos, N. D. Padilla, M. A. Strauss, S. A. Cora, and L. Hao, "Host galaxy-active galactic nucleus alignments in the sloan digital sky survey data release 7," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 3, pp. 2148–2162, 2011.

[93] M. A. Malkan, V. Gorjian, and R. Tam, "A hubble space telescope 1 imaging survey of nearby active galactic nuclei," *Astrophysical Journal, Supplement Series*, vol. 117, no. 1, pp. 25–88, 1998.

[94] M. Guainazzi, G. Matt, and G. C. Perola, "X-ray obscuration and obscured AGN in the local universe," *Astronomy and Astrophysics*, vol. 444, no. 1, pp. 119–132, 2005.

[95] S. Bianchi, M. Chiaberge, E. Piconcelli, and M. Guainazzi, "A multiwavelength map of the nuclear region of NGC 7582," *Monthly Notices of the Royal Astronomical Society*, vol. 374, no. 2, pp. 697–702, 2007.

[96] E. Schinnerer, A. Eckart, L. J. Tacconi, R. Genzel, and D. Downes, "Bars and warps traced by the molecular gas in the Seyfert 2 galaxy NGC 1068," *The Astrophysical Journal*, vol. 533, no. 2, pp. 850–868, 2000.

[97] F. Boone, S. García-Burillo, F. Combes et al., "High-resolution mapping of the physical conditions in two nearby active galaxies based on 12CO(1-0), (2-1), and (3-2) lines," *Astronomy and Astrophysics*, vol. 525, no. 1, article A18, 2010.

[98] M. Krips, S. Martin, and A. Eckart, "SMA/PdBI multiple line observations of the nearby Seyfert2 galaxy NGC 1068: shock related gas kinematics and heating in the central 100pc?" *The Astrophysical Journal*, vol. 736, article 37, 2011.

[99] M. Polletta, D. Weedman, S. Hönig, C. J. Lonsdale, H. E. Smith, and J. Houck, "Obscuration in extremely luminous quasars," *The Astrophysical Journal*, vol. 675, no. 2, pp. 960–984, 2008.

[100] A. Martínez-Sansigre, S. Rawlings, M. Lacy et al., "A population of high-redshift type 2 quasars - I. Selection criteria and optical spectra," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 3, pp. 1479–1498, 2006.

[101] G. Risaliti, R. Maiolino, and M. Salvati, "The distribution of absorbing column densities among Seyfert 2 galaxies," *The Astrophysical Journal*, vol. 522, no. 1, pp. 157–164, 1999.

[102] A. J. Young, A. S. Wilson, and P. L. Shopbell, "A Chandra X-ray study of NGC 1068. I. Observations of extended emission," *The Astrophysical Journal*, vol. 556, no. 1, pp. 6–23, 2001.

[103] C. A. Heisler, S. L. Lumsden, and J. A. Bailey, "Visibility of scattered broad-line emission in Seyfert 2 galaxies," *Nature*, vol. 385, no. 6618, pp. 700–702, 1997.

[104] D. M. Alexander, "The unified model and the Seyfert 2 infrared dichotomy," *Monthly Notices of the Royal Astronomical Society*, vol. 320, no. 1, pp. L15–L19, 2001.

[105] Q. Gu, R. Maiolino, and D. Dultzin-Hacyan, "Nuclear obscuration and scattering in Seyfert 2 galaxies," *Astronomy and Astrophysics*, vol. 366, no. 3, pp. 765–770, 2001.

[106] F. Nicastro, "Broad emission line regions in active galactic nuclei: the link with the accretion power," *The Astrophysical Journal Letters*, vol. 530, no. 2, article L65, 2000.

[107] M. Elitzur and L. C. Ho, "On the disappearance of the broad-line region in low-luminosity active galactic nuclei," *The Astrophysical Journal Letters*, vol. 701, no. 2, article L91, 2009.

[108] J. R. Trump, C. D. Impey, B. C. Kelly et al., "Accretion rate and the physical nature of unobscured active galaxies," *The Astrophysical Journal*, vol. 733, no. 1, article 60, 2011.

[109] S. Bianchi, A. Corral, F. Panessa et al., "NGC 3147: a "true" type 2 Seyfert galaxy without the broad-line region," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 1, pp. 195–199, 2008.

[110] F. Panessa, F. J. Carrera, S. Bianchi et al., "Unabsorbed Seyfert 2 galaxies: the case of "naked" AGN," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 4, pp. 1951–1960, 2009.

[111] M. Brightman and K. Nandra, "On the nature of unabsorbed Seyfert 2 galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 3, pp. 1241–1249, 2008.

[112] S. Bianchi, in prep., 2012.

[113] A. Marinucci, S. Bianchi, F. Nicastro, G. Matt, and A. D. Goulding, "The link between the hidden broad line region andthe accretion rate in seyfert 2 galaxies," *The Astrophysical Journal*. In press.

[114] Y. Shi, G. H. Rieke, P. Smith et al., "Unobscured type 2 active galactic nuclei," *The Astrophysical Journal*, vol. 714, no. 1, pp. 115–129, 2010.

[115] F. Nicastro, A. Martocchia, and G. Matt, "The lack of broad-line regions in low accretion rate active galactic nuclei

as evidence of their origin in the accretion disk," *The Astrophysical Journal*, vol. 589, no. 1, pp. L13–L16, 2003.

[116] W. Bian and Q. Gu, "The eddington ratios in seyfert 2 galaxies with and without hidden broad-line regions," *The Astrophysical Journal*, vol. 657, no. 1, pp. 159–166, 2007.

[117] X. W. Shu, J. X. Wang, P. Jiang, L. L. Fan, and T. G. Wang, "Investigating the nuclear obscuration in two types of Seyfert 2 galaxies," *The Astrophysical Journal*, vol. 657, no. 1 I, pp. 167–176, 2007.

[118] Y. Wu, E. Zhang, Y. Liang, C. Zhang, and Y. Zhao, "The different nature of seyfert 2 galaxies with and without hidden broad-line regions," *The Astrophysical Journal*, vol. 730, no. 2, p. 121, 2011.

[119] A. Capetti, D. J. Axon, F. Macchetto, W. B. Sparks, and A. Boksenberg, "Radio outflows and the origin of the narrow-line region in seyfert galaxies," *The Astrophysical Journal*, vol. 469, no. 2, pp. 554–563, 1996.

[120] S. Bianchi, M. Chiaberge, D. A. Evans et al., "High-resolution X-ray spectroscopy and imaging of Mrk 573," *Monthly Notices of the Royal Astronomical Society*, vol. 405, no. 1, pp. 553–563, 2010.

[121] R. W. Goosmann and G. Matt, "Spotting the misaligned outflows in NGC 1068 using X-ray polarimetry," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 4, pp. 3119–3128, 2011.

[122] A. C. Fabian, K. Iwasawa, C. S. Reynolds, and A. J. Young, "Broad iron lines in active galactic nuclei," *Publications of the Astronomical Society of the Pacific*, vol. 112, no. 775, pp. 1145–1161, 2000.

[123] M. Guainazzi, S. Bianchi, I. de La Calle Perez, M. Dovciak, and A. L. Longinotti, "On the driver of relativistic effect strength in Seyfert galaxies," *Astronomy & Astrophysics*, vol. 531, article A131, 2011.

[124] G. Risaliti, M. Salvati, and A. Marconi, "[O iii] equivalent width and orientation effects in quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 4, pp. 2223–2229, 2011.

[125] Y. Ueda, M. Akiyama, K. Ohta, and T. Miyaji, "Cosmological evolution of the hard X-ray active galactic nucleus luminosity function and the origin of the hard X-ray background," *The Astrophysical Journal*, vol. 598, no. 2 I, pp. 886–908, 2003.

[126] A. T. Steffen, A. J. Barger, L. L. Cowie, R. F. Mushotzky, and Y. Yang, "The changing active galactic nucleus population," *The Astrophysical Journal*, vol. 596, no. 1, pp. L23–L26, 2003.

[127] F. La Franca, F. Fiore, A. Comastri et al., "The HELLAS2XMM survey. VII. The hard X-ray luminosity function of AGNs up to $z = 4$: more absorbed AGNs at low luminosities and high redshifts," *The Astrophysical Journal*, vol. 635, no. 2 I, pp. 864–879, 2005.

[128] A. Akylas, I. Georgantopoulos, A. Georgakakis, S. Kitsionas, and E. Hatziminaoglou, "XMM-Newton and Chandra measurements of the AGN intrinsic absorption: dependence on luminosity and redshift," *Astronomy and Astrophysics*, vol. 459, no. 3, pp. 693–701, 2006.

[129] A. J. Barger and L. L. Cowie, "The number density of intermediate- and high-luminosity active galactic nuclei at $z \sim 2 - 3$," *The Astrophysical Journal*, vol. 635, no. 1 I, pp. 115–122, 2005.

[130] P. Tozzi, R. Gilli, V. Mainieri et al., "X-ray spectral properties of active galactic nuclei in the Chandra Deep Field South," *Astronomy and Astrophysics*, vol. 451, no. 2, pp. 457–474, 2006.

[131] C. Simpson, "The luminosity dependence of the type 1 active galactic nucleus fraction," *Monthly Notices of the Royal Astronomical Society*, vol. 360, no. 2, pp. 565–572, 2005.

[132] T. Dwelly and M. J. Page, "The distribution of absorption in AGN detected in the XMM-Newton observations of the CDFS," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 4, pp. 1755–1775, 2006.

[133] E. Treister and C. M. Urry, "Active galactic nuclei unification and the X-ray background," *The Astrophysical Journal*, vol. 630, no. 1 I, pp. 115–121, 2005.

[134] J.-M. Wang and E.-P. Zhang, "The unified model of active galactic nuclei. II. Evolutionary connection," *The Astrophysical Journal*, vol. 660, no. 2 I, pp. 1072–1092, 2007.

[135] G. Hasinger, "Absorption properties and evolution of active galactic nuclei," *Astronomy and Astrophysics*, vol. 490, no. 3, pp. 905–922, 2008.

[136] R. Della Ceca, A. Caccianiga, P. Severgnini et al., "The cosmological properties of AGN in the XMM-Newton Hard Bright Survey," *Astronomy and Astrophysics*, vol. 487, no. 1, pp. 119–130, 2008.

[137] E. Treister, S. Virani, E. Gawiser et al., "Optical spectroscopy of x-ray sources in the extended chandra deep field south," *The Astrophysical Journal*, vol. 693, no. 2, p. 1713, 2009.

[138] M. Brusa, F. Civano, A. Comastri et al., "The XMM-newton wide-field survey in the cosmos field (XMM-COSMOS): demography and multiwavelength properties of obscured and unobscured luminous active galactic nuclei," *The Astrophysical Journal*, vol. 716, no. 1, pp. 348–369, 2010.

[139] R. Maiolino, O. Shemmer, M. Imanishi et al., "Dust covering factor, silicate emission, and star formation in luminous QSOs," *Astronomy and Astrophysics*, vol. 468, no. 3, pp. 979–992, 2007.

[140] R. Gilli, A. Comastri, and G. Hasinger, "The synthesis of the cosmic X-ray background in the Chandra and XMM-Newton era," *Astronomy and Astrophysics*, vol. 463, no. 1, pp. 79–96, 2007.

[141] J.-M. Wang, E.-P. Zhang, and B. Luo, "Evolutionary consequences of dusty tori in active galactic nuclei," *The Astrophysical Journal*, vol. 627, no. 1, article L5, 2005.

[142] R. Mor and B. Trakhtenbrot, "Hot-dust clouds with pure-graphite composition around type-i active galactic nuclei," *The Astrophysical Journal Letters*, vol. 737, no. 2, p. L36, 2011.

[143] M. Rowan-Robinson, I. Valtchanov, and K. Nandra, "AGN dust tori: the X-ray-infrared connection," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 3, pp. 1326–1337, 2009.

[144] K. Iwasawa and Y. Taniguchi, "The X-ray baldwin effect," *The Astrophysical Journal*, vol. 413, no. 1, pp. L15–L18, 1993.

[145] K. L. Page, P. T. O'Brien, J. N. Reeves, and M. J. L. Turner, "An X-ray Baldwin effect for the narrow Fe Kα lines observed in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 347, no. 1, pp. 316–322, 2004.

[146] P. Jiang, J. X. Wang, and T. G. Wang, "On the X-ray Baldwin effect for narrow Fe Ka emission lines," *The Astrophysical Journal*, vol. 644, no. 2, pp. 725–732, 2006.

[147] M. Guainazzi, S. Bianchi, and M. Dovčiak, "Statistics of relativistically broadened Fe Kα lines in AGN," *Astronomische Nachrichten*, vol. 327, no. 10, pp. 1032–1038, 2006.

[148] S. Bianchi, M. Guainazzi, G. Matt, and N. F. Bonilla, "On the Iwasawa-Taniguchi effect of radio-quiet AGN," *Astronomy and Astrophysics*, vol. 467, no. 1, pp. L19–L22, 2007.

[149] A. Lawrence, "The relative frequency of broad-lined and narrow-lined active galactic nuclei—implications for unified

schemes," *Monthly Notices of the Royal Astronomical Society*, vol. 252, pp. 586–592, 1991.

[150] A. Lamastra, G. C. Perola, and G. Matt, "A model for the X-ray absorption in compton-thin AGN," *Astronomy and Astrophysics*, vol. 449, no. 2, pp. 551–558, 2006.

[151] J. Fischer, E. Sturm, E. González-Alfonso et al., "Herschel-PACS spectroscopic diagnostics of local ULIRGs: conditions and kinematics in Markarian 231," *Astronomy and Astrophysics*, vol. 518, no. 14676, article L36, 2010.

[152] C. Feruglio, R. Maiolino, E. Piconcelli et al., "Quasar feedback revealed by giant molecular outflows," *Astronomy and Astrophysics*, vol. 518, no. 13, article L155, 2010.

[153] D. S. N. Rupke and S. Veilleux, "Warm molecular hydrogen in the galactic wind of M82," *The Astrophysical Journal*, vol. 729, no. 2, article L27, 2011.

[154] E. Sturm, E. Gonzalez-Alfonso, S. Veilleux et al., "Massive molecular outflows and negative feedback in ULIRGs observed by Herschel-PACS," *The Astrophysical Journal*, vol. 733, no. 1, p. L16, 2011.

[155] M. Cano-Diaz, R. Maiolino, A. Marconi, H. Netzer, O. Shemmer, and G. Cresci, "Observational evidence of quasar feedback quenching star formation at high redshift," *Astronomy & Astrophysics*, vol. 537, article L8, 2012.

[156] D. R. Ballantyne, Y. Shi, G. H. Rieke, J. L. Donley, C. Papovich, and J. R. Rigby, "Does the AGN unified model evolve with redshift? Using the X-ray background to predict the mid-infrared emission of AGNs," *The Astrophysical Journal*, vol. 653, no. 2, pp. 1070–1088, 2006.

[157] E. Treister and C. M. Urry, "The evolution of obscuration in active galactic nuclei," *The Astrophysical Journal*, vol. 652, pp. L79–L82, 2006.

[158] L. J. Tacconi, R. Genzel, R. Neri et al., "High molecular gas fractions in normal massive star-forming galaxies in the young Universe," *Nature*, vol. 463, no. 7282, pp. 781–784, 2010.

[159] E. Daddi, F. Bournaud, F. Walter et al., "Very high gas fractions and extended gas reservoirs in z = 1.5 disk galaxies," *The Astrophysical Journal*, vol. 713, no. 1, pp. 686–707, 2010.

[160] L. Shao, D. Lutz, R. Nordon et al., "Star formation in AGN hosts in GOODS-N," *Astronomy and Astrophysics*, vol. 518, no. 1, article L26, 2010.

[161] Y. Shi, G. H. Rieke, P. Ogle, L. Jiang, and A. M. Diamond-Stanic, "Cosmic evolution of star formation in type-1 quasar hosts since z = 1," *The Astrophysical Journal*, vol. 703, no. 1, pp. 1107–1122, 2009.

[162] C. Ricci, R. Walter, T. J.-L. Courvoisier, and S. Paltani, "Reflection in Seyfert galaxies and the unified model of AGN," *Astronomy & Astrophysics*, vol. 532, article A102, 2011.

[163] A. Marconi, E. Oliva, P. P. van der Werf et al., "The elusive active nucleus of NGC 4945," *Astronomy and Astrophysics*, vol. 357, no. 1, pp. 24–36, 2000.

[164] R. Della Ceca, L. Ballo, F. Tavecchio et al., "An enshrouded active galactic nucleus in the merging starburst system Arp 299 revealed by BeppoSAX," *The Astrophysical Journal*, vol. 581, no. 1, pp. L9–L13, 2002.

[165] R. Maiolino, A. Comastri, R. Gilli et al., "Elusive active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 344, no. 4, pp. L59–L64, 2003.

[166] L. Ballo, V. Braito, R. Della Ceca, L. Maraschi, F. Tavecchio, and M. Dadina, "ARP 299: a second merging system with two active nuclei?" *The Astrophysical Journal*, vol. 600, no. 2, pp. 634–639, 2004.

[167] A. Caccianiga, P. Severgnini, R. Della Ceca, T. Maccacaro, F. J. Carrera, and M. J. Page, "Elusive AGN in the XMM-Newton bright serendipitous survey," *Astronomy and Astrophysics*, vol. 470, no. 2, pp. 557–570, 2007.

[168] M. Imanishi, R. Maiolino, and T. Nakagawa, "Spitzer infrared low-resolution spectroscopic study of buried active galactic nuclei in a complete sample of nearby ultraluminous infrared galaxies," *The Astrophysical Journal*, vol. 709, no. 2, pp. 801–815, 2010.

[169] M. Imanishi, C. C. Dudley, R. Maiolino, P. R. Maloney, T. Nakagawa, and G. Risaliti, "A Spitzer IRS low-resolution spectroscopic search for buried AGNs in nearby ultraluminous infrared galaxies: a constraint on geometry between energy sources and dust," *Astrophysical Journal, Supplement Series*, vol. 171, no. 1, pp. 77–100, 2007.

[170] M. Imanishi, "Luminous buried active galactic nuclei as a function of galaxy infrared luminosity revealed through spitzer low-resolution infrared spectroscopy," *The Astrophysical Journal*, vol. 694, no. 2, pp. 751–764, 2009.

[171] G. Risaliti, M. Imanishi, and E. Sani, "A quantitative determination of the AGN content in local ULIRGs through L-band spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 1, pp. 197–203, 2010.

[172] E. Nardini, G. Risaliti, M. Salvati et al., "Exploring the active galactic nucleus and starburst content of local ultraluminous infrared galaxies through 5-8 μm spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 3, pp. 1373–1402, 2009.

[173] E. Nardini, G. Risaliti, M. Salvati et al., "Spectral decomposition of starbursts and active galactic nuclei in 5—8 μm Spitzer-IRS spectra of local ultraluminous infrared galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 1, pp. L130–L134, 2008.

[174] D. A. Evans, W.-F. Fong, M. J. Hardcastle et al., "A Radio through X-Ray Study of the Jet/Companion-Galaxy Interaction in 3C 321," *The Astrophysical Journal*, vol. 675, pp. 1057–1066.

[175] A. D. Goulding and D. M. Alexander, "Towards a complete census of AGN in nearby Galaxies: a large population of optically unidentified AGN," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 3, pp. 1165–1193, 2009.

[176] J. P. Perez-Beaupuits, H. W. W. Spoon, M. Spaans, and J. D. Smith, "The deeply obscured AGN of NGC 4945. I. Spitzer-IRS maps of [Ne V], [Ne II], H_2 0-0 S(1), S(2), and other tracers," *Astronomy & Astrophysics*, vol. 533, article A56, 2011.

[177] L. Armus, V. Charmandaris, J. Bernard-Salas et al., "Observations of ultraluminous infrared galaxies with the infrared spectrograph on the spitzer space telescope. II. The iras bright galaxy sample," *The Astrophysical Journal*, vol. 656, no. 1, pp. 148–167, 2007.

[178] Y. Ueda, S. Eguchi, Y. Terashima et al., "Suzaku observations of active galactic nuclei detected in the swift BAT survey: discovery of a "New Type" of buried supermassive black holes," *The Astrophysical Journal*, vol. 664, pp. L79–L82, 2007.

[179] K. Noguchi, Y. Terashima, and H. Awaki, "A new sample of buried active galactic nuclei selected from the second xmm-newton serendipitous source catalogue," *The Astrophysical Journal*, vol. 705, no. 1, pp. 454–467, 2009.

[180] L. Jiang, X. Fan, W. N. Brandt et al., "Dust-free quasars in the early Universe," *Nature*, vol. 464, no. 7287, pp. 380–383, 2010.

[181] H. Hao, M. Elvis, F. Civano et al., "Hot-dust-poor type 1 active galactic nuclei in the COSMOS survey," *The Astrophysical Journal*, vol. 724, no. 1, pp. L59–L63, 2010.

[182] H. Hao, M. Elvis, F. Civano, and A. Lawrence, "Hot-dust-poor quasars in mid-infrared and optically selected samples," *The Astrophysical Journal*, vol. 733, no. 2, article 108, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 271502, 7 pages
doi:10.1155/2012/271502

Research Article

Demography of High-Redshift AGN

Fabrizio Fiore,[1] Simonetta Puccetti,[2] and Smita Mathur[3]

[1] *Fabrizio Fiore INAF-Osservatorio Astronomico di Roma, via Frascati 33, I00040 Monteporzio Catone, Italy*
[2] *Simonetta Puccetti ASI Science Data Center, via Galileo Galilei, 00044 Frascati, Italy*
[3] *Smita Mathur Astronomy Department, Ohio State University, Columbus, OH 43210, USA*

Correspondence should be addressed to Fabrizio Fiore, fiore@oa-roma.inaf.it

Received 21 September 2011; Accepted 9 November 2011

Academic Editor: Angela Bongiorno

High-redshift AGN holds the key to understanding early structure formation and to probe the Universe during its infancy. We review the latest searches for high-z AGN in the deepest X-ray field so far, the Chandra Deep Field South (CDFS) 4 Msecond exposure. We do not confirm the positive detection of a signal in the stacked Chandra images at the position of $z \sim 6$ galaxies recently reported by Treister and collaborators (2011). We present $z > 3$ X-ray sources number counts in the 0.5–2 keV band, obtained joining CDFS faint detections (see Fiore et al. (2011)), with Chandra-COSMOS and XMM-COSMOS detections. We use these number counts to make predictions for surveys with three mission concepts: Athena, WFXT, and a Super-Chandra.

1. Introduction

The study of high-redshift AGN holds the key to understanding early structure formation and probing the Universe during its infancy. There are fundamental issues that can be tackled by studying high-z AGN (1) the evolution of the correlations between the black hole mass and the galaxy properties (see, e.g., [1] and references therein); (2) the AGN contribution to the reionization and to the heating of the Intergalactic medium and its effect on structure formation (e.g., [2, 3] and references therein); (3) scenarios for the formation of the black hole (BH) seeds that will eventually grow up to form the supermassive black holes (SMBHs) seen in most galaxy bulges (e.g., [4]); (4) we can investigate the physics of accretion at high-z. One open question is whether BH growth is mainly due to relatively few accretion episodes, as predicted in hierarchical scenarios (see, e.g., [5] and references therein) or by the so-called chaotic accretion (hundreds to thousands of small accretion episodes, [6]). The two scenarios predict different BH spin distributions, and thus different distributions for the radiative efficiency; (5) since BHs are the structures with the fastest (exponential) growth rate, they can be used to constrain both the expansion rate of the Universe and the growth rate of the primordial perturbation at high-z, that is, competing cosmological scenarios [7, 8]; (6) since the slopes of both the high-z AGN luminosity function and the SMBH mass function strongly depend on the AGN duty cycle, their measurements can constrain this critical parameter. In turn, the AGN duty cycle holds information on the AGN triggering mechanisms. The evaluation of the evolution of the AGN duty cycle can thus help us to distinguish the competing scenarios for AGN triggering and feeding [9].

Large area optical and near infrared surveys such as the SDSS, the CFHQS, the NOAO DWFS/DLS, and the UKIDSS surveys have already been able to discover large samples of $z > 4.5$ QSOs (e.g., [10, 11]) and about 50 QSOs at $z > 5.8$ (e.g., [12–14]). The majority of these high-z AGN are broad line, unobscured, high UV rest-frame luminosity (thus high bolometric luminosity) AGN. They are likely the tips of the iceberg of the high-z AGN population. Lower bolometric luminosity and/or moderately obscured AGN can, in principle, be detected directly in current and future X-ray surveys. Dedicated searches for high-z AGN using both deep and wide area X-ray surveys and a multiband selection of suitable candidates can increase the number of high-z AGN by a factor >10. In particular, it should be possible to find hundreds rare high-z, high luminosity QSOs, in both the all sky and deep eROSITA surveys (the 0.5–2 keV flux limit of the all sky survey being the order of 10^{-14} erg/cm^2/s, while that of the deep survey, covering hundreds deg^2, should be 2-3 times deeper [15]) with a selection function much less biased

than optical surveys. To constrain the faint end of the high-z AGN luminosity function and therefore the shape of the luminosity function and of the SMBH mass function, we need to best exploit current and future deep surveys. The Chandra Deep Field South is today the *premiere* field, with its 4 Msec and 3 Msec exposures obtained by Chandra and XMM, respectively, since 1999 [16, 17]. Three different approaches have been so far applied to this field: (a) direct detection of sources in X-ray maps (e.g., [16]); (b) search for X-ray emission at the position of candidate high-z galaxies selected in the red and near infrared bands [9]; (c) stacking of X-ray counts at the position of candidate high-z galaxies [18]. Here, we review all three methods and give state of the art number counts of high-z AGN at faint fluxes. We use these number counts to predict the number of high-z AGN in possible future deep X-ray surveys. A $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ cosmology is adopted throughout.

2. Stacking Analysis of Candidate High-z Galaxies

Recently, Treister et al. [18] published a positive detection of X-ray counts in stacked Chandra images obtained adding together the X-ray counts at the position of 197 candidate high-z galaxies at $z \sim 6$ in the CDFS and CDFN [19]. They find 5σ and 6.8σ detections in the soft 0.5–2 keV and hard 2–8 keV bands. Since the 2–8 keV flux they detect is about 9 times the 0.5–2 keV flux, they infer that the majority of these faint high-z galaxies host highly obscured, Compton thick AGN. The total rest frame 2–10 keV luminosity density implied by the Treister [18] result is 1.6×10^{46} ergs/s/deg^2 at $z \sim 6$. In contrast, Fiore et al. [9] analyzed X-ray counts at the position of the same Bouwens et al. [19] $z \sim 6$ galaxies in the CDFS finding just one marginal detection. Fiore et al. [9] find that the $z \sim 6$ luminosity function can be modeled using the standard double power law shape

$$\frac{\mathrm{d}\Phi(L_\mathrm{X})}{\mathrm{d}\,\mathrm{Log}\,L_\mathrm{X}} = A\left[\left(\frac{L_\mathrm{X}}{L_*}\right)^{\gamma 1} + \left(\frac{L_\mathrm{X}}{L_*}\right)^{\gamma 2}\right]^{-1}, \quad (1)$$

with $L^* = 2 \times 10^{44}$ ergs/s, $\gamma 1 = 0.8$, and $\gamma 2 = 3.4$ (the faint end slope is not truly constrained). By integrating, this luminosity function above a luminosity of 10^{42} erg/s, one obtains a total rest frame 2–10 keV luminosity density at $z \sim 6$ of 5.6×10^{45} ergs/s/deg^2, a value ~3 times smaller than that reported by Treister et al. [18]. We investigated this discrepancy between the Treister [18] and Fiore [9] results. Once again, we considered the Bouwens et al. [19] sample of 371 candidate $z \sim 6$ galaxies in the CDFS. Some of these galaxies lies close to bright X-ray sources, identified with galaxies at a different redshift, and must therefore be excluded from the analysis. We considered two exclusion radii, one similar to that used by Treister et al. [18], that is, 22 arcsec, and another, less conservative, of 10 arcsec. In both cases, we used the new Xue et al. [16] catalog of 740 directly detected X-ray sources (it is not possible to ascertain which X-ray catalog was used by [18]). We considered sources at an offaxis angle <8 arcmin, to avoid the inclusion of sources observed with a too broad PSF. We considered only one galaxy when we find 2 or more within 2 arcsec, not to count twice the contribution from each single object. We finally excluded from the samples $z \sim 6$ galaxies closer than 2 arcsec from lower redshift galaxies brighter than zmag =25, which may emit X-rays and hence contaminate the high redshift stacks. The final samples that we consider include 210 galaxies (10 arcsec exclusion radii) and 77 galaxies (22 arcsec exclusion radii). We performed stacks of Chandra counts at the position of these galaxies in four energy bands: 0.5–2 keV, 2–7 keV, 0.8–4 keV, and 1.5–5.5 keV. The total exposure times for the two sample are $\sim 2.3 \times 10^8$ seconds (77 galaxies) and $\sim 6.3 \times 10^8$ seconds (210 galaxies). Figure 1 shows the stacked images for the two samples in the four energy bands. We do not find a significant signal at the position of the galaxies in any of these images. Table 1 gives the PSF-corrected 3σ count rate upper limits from the counts collected in boxes of 5 arcsec side (area of 100 original pixels). As a comparison, Treister et al. [18] report a count rate $3.4 \pm 0.7 \times 10^{-7}$ counts/s in the 0.5–2 keV band and $8.8 \pm 1.3 \times 10^{-7}$ counts/s in the 2–8 keV band. Our more stringent upper limits are obtained for the 210 galaxy sample in the 0.5–2 keV and 2–7 keV bands. These are, respectively, comparable and 1.5 times lower than the Treister [18] claimed detections.

We can convert our count rate upper limits to a limit to the rest frame 2–10 keV luminosity density following [18]. We find a 3σ limit of $\sim 10^{46}$ ergs/s/deg^2, lower than the [18] feature, but about twice the luminosity density estimated by Fiore et al. [9].

We recall that our analysis applies to the CDFS field alone, while the [18] result applies to the joined CDFS and CDFN area. In principle, part of the discrepancy between Treister et al. [18] and our analysis could be due to cosmic variance. However, this seems unlikely because most of the Treister sources are in the CDFS, which has an exposure twice that of the CDFN. We also recall that, for the sake of robustness, our stacking analysis is the simplest possible. First, counts at the position of galaxies are added together. Second, aperture photometry is performed on the stacked images, without any optimization for off-axis-dependent PSF. Third, background is estimated in nearby regions, and, unlike [18], no removal of positive fluctuations is performed. While this simple technique does not probably push the detection to the limit, it nevertheless produced robust result. In particular, it produced positive, valuable results in the past, when applied to samples of candidate, faint, Compton thick AGN [20, 21]. The result of our Chandra analysis of the Bouwens candidate $z > 6$ galaxies has been recently confirmed by Willott [22].

3. High-z AGN Number Counts

The analyses on the CDFS, CDFN, EGS, and COSMOS fields provide samples of individual sources detected; hence, X-ray number counts of faint high-z sources can be easily computed from these samples. Figure 2 shows $z > 3$ number counts from a compilation of surveys: the Fiore et al. [9] survey of the ERS and GOODS fields, and the Brusa et al. and Civano et al. [23, 24] XMM and Chandra surveys of the COSMOS field. Figure 2 also shows the $z \sim 6$ point from [18] and the upper limit we evaluated in the previous section from

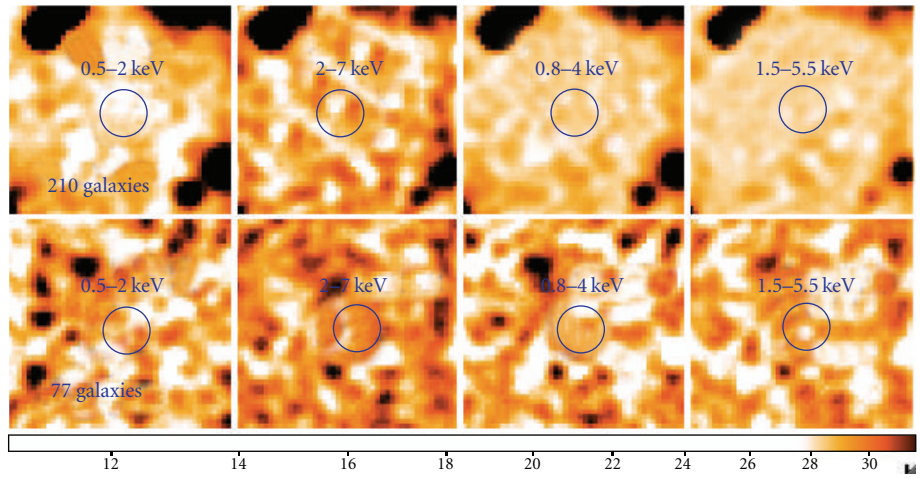


Figure 1: Stacks of Chandra images at the position of 210 and 77 Bouwens et al. [19] candidate $z \sim$ 6-7 galaxies in four energy bands: 0.5–2 keV, 2–7 keV, 0.8–4 keV, and 1.5–5.5 keV.

a stacking analysis. Black solid lines are model number counts obtained by converting the [9] luminosity functions. The model reproduces reasonably well the number counts at $z > 3$ in the full flux range probed by observations. It is consistent with the CDFS points at low fluxes in the other redshift bins, while it is slightly above the Chandra-COSMOS points at intermediate fluxes for the redshift bins $z > 4$ and $z > 5$. However, the number of sources in this survey at these redshifts is low, and the corresponding error due to both statistics and cosmic variance is large. Deeper exposures of fields with area of 1-2 deg^2 are needed to better explore this region of the redshift-flux parameter space. At the flux limits reached by the deepest Chandra exposure (4 Mseconds), there are >1000 $z > 3$ AGN/deg^2, several hundreds $z > 4$ AGN/deg^2, >100 $z > 5$ AGN/deg^2, 20–100 $z > 5.8$ AGN/deg^2 (the uncertainty on the latter number is that large, because the measure is based on just 2 candidate $z > 5.8$ galaxies detected by Chandra in the small ERS field). It is clear that, to obtain a more robust demography of the $z > 6$ AGN, a search in a much wider area, such as the CANDELS area [25], is mandatory, and requires spectroscopic confirmation of the X-ray emitting, candidate $z > 6$ galaxies. The CANDELS deep and wide surveys cover a total of 130 arcmin^2 and 670 arcmin^2 to a depth of $H = 27.8$ and $H \sim 26.5$, respectively. As a comparison, the ERS survey covers an area of 50 arcmin^2 to a depth of $H \sim 27$. The two candidate $z > 6$ ERS galaxies detected by Chandra in the ERS field are faint, $H = 26.6$ and $H = 27$ sources. The GOODS source with $z > 7$ in the Luo et al. [26] catalog has $H = 27.6$. The other $z > 6$ ERS galaxy with a marginal X-ray detection is brighter, $H = 23.8$. In summary, we expect 1–5 $z > 6$ AGN in the CANDELS deep survey and 4–20 $z > 6$ AGN in the CANDELS wide survey. We note that a fraction of these sources will be at the limit, or below, the H band sensitivity threshold of the wide survey. As of today, Chandra has spent of the order of 8 Mseconds on the CANDELS fields, most of them on the CANDELS deep fields. To reach the sensitivity to detect the faint $z > 6$ AGN in the wide area, additional 5-6 Mseconds are needed. This is within reach of the Chandra observatory in the next few years. To make further progresses with Chandra, that is, quantitatively probe the first generation of accreting SMBH, which would allow putting stringent constraints on SMBH formation models [27–30] and accretion scenarios [4–6, 31], would require to at least triple the exposure times, that is, 30–40 Mseconds dedicated to deep surveys. While this is certainly extremely expensive, it is not technically unfeasible.

Table 1: 3σ count rates upper limits.

Sample	0.5–2 keV 10^{-7} cts/s	2–7 keV 10^{-7} cts/s	0.8–4 keV 10^{-7} cts/s	1.5–5.5 keV 10^{-7} cts/s
210 galaxies	3.4	5.8	6.2	6.2
77 galaxies	5.9	9.7	9.6	10.1

4. Predictions for Future Deep Surveys

The Chandra limiting problem is that its sensitivity is very good on axis but degrades quickly at off-axis angles higher than a few arcmin, making difficult and expensive in terms of exposure time to cover with good sensitivity an area larger than a few hundred arcmin^2. A significant leap forward in the field would then be obtained by an instrument capable of reaching the Chandra Msecond on axis sensitivity (i.e., flux limit of 2–4 $\times 10^{-17}$ erg/cm^2/s), but on a factor of >10 wider field of view (FOV). We consider here three possible mission concepts, and we make predictions on the number of $z > 4$, $z > 5$ and $z > 6$ faint X-ray sources based on our best knowledge of the X-ray the number counts at high redshift (Figure 2) and of the X-ray AGN luminosity functions [9].

(1) The first mission concept that we consider is that of Athena. This is a proposal for L class mission in the framework of the ESA Cosmic Vision program. The baseline mission concept foresees an effective area for

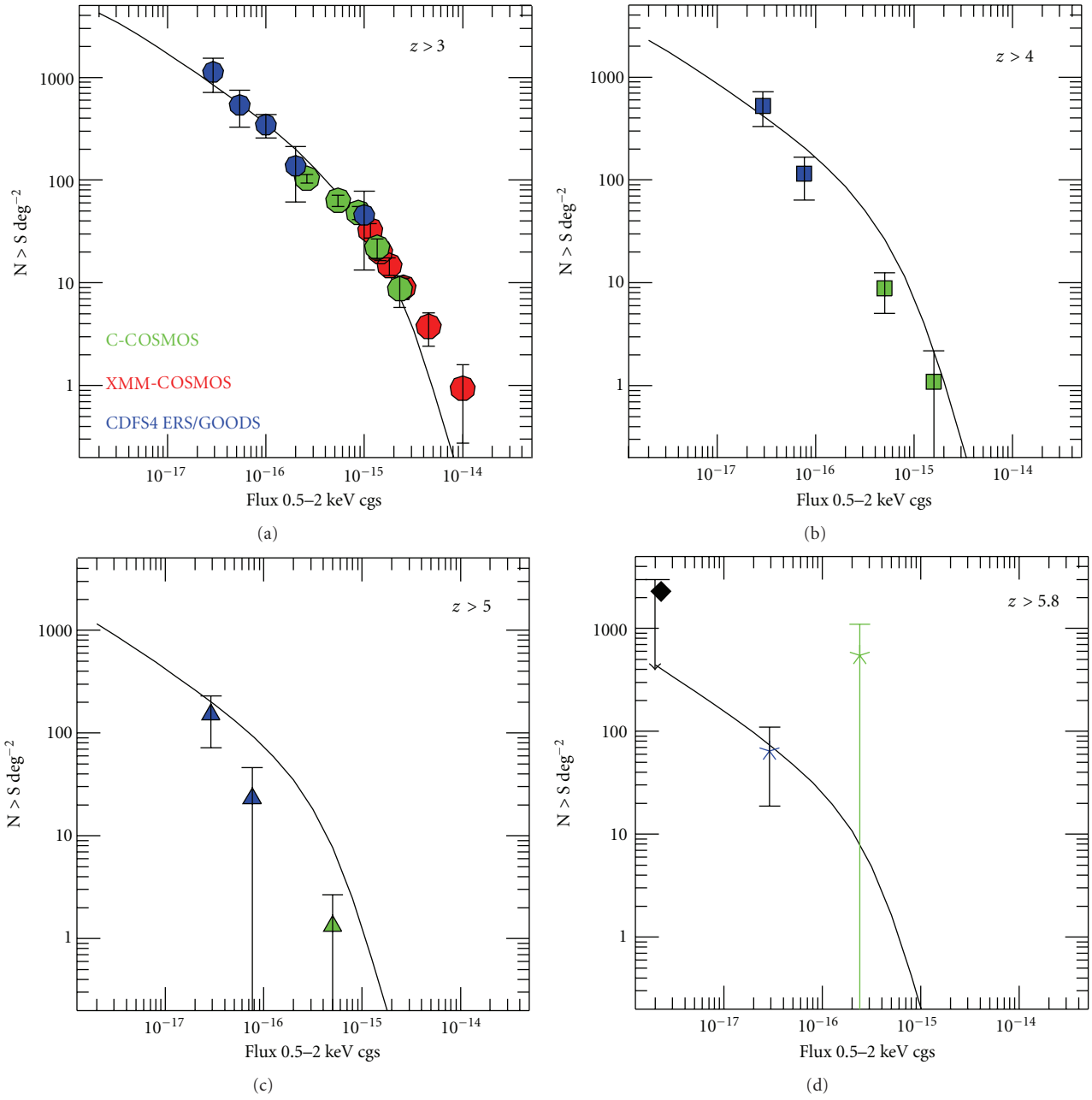


Figure 2: Faint X-ray sources number counts in the 0.5–2 keV band and in four redshift ranges. Blue points are from X-ray detections at the position of ERS and GOODS-MUSIC $z > 3$ galaxies [9]; green points are from Chandra-COSMOS [23]; red points are from XMM-COSMOS [24]. Circles = $z > 3$ sources, squares = $z > 4$ sources, triangles = $z > 5$, and star = $z > 5.8$ sources. Diamond = Treister et al. determination [18]. The upper limit in the lower right panel is from Table 1. The thin solid curves are model number counts based on the best fit high-z luminosity functions presented in [9].

imaging of the order of half square meter at 1-2 keV, a mirror PSF with half power diameter HPD ~ 10 arcsec (requirement, 5 arcsec goal), focal length 11 m, FOV = 0.17 deg^2 (25 × 25 arcmin). The observatory should be launched on a high earth orbit (HEO) or a L2 orbit, and therefore a rather high internal background is predicted, similar to the internal background measured by the instruments on board XMM and Chandra, which are flying on HEO.

(2) The second mission concept that we consider is a wide field X-ray telescope, WFXT, fully dedicated to X-ray surveys. This idea is quite old, the first proposal dating mid 90′, and the mission concept evolved considerably over the years. We assume a configuration similar to that in [32], that is an effective area of >half square meter at 1-2 keV, split in three mirror units, with HPD = 10 arcsec (requirement, goal 5 arcsec) and 5.5 m focal length. Each mirror unit is feeding

Table 2: Predicted number of faint, high-z X-ray sources.

Mission concept	PSF HPD Arcsec	Mosaics	Total FOV deg^2	z = 4-5 L_X (z = 5)	z = 5–5.8 L_X (z = 6)	z > 5.8 L_X (z = 7)
Athena*	10	60 × 0.2 Msec	10	940 > 43.3	480 > 43.5	250 > 43.6
Athena*	5	6 × 2 Msec	1.0	360 > 42.5	210 > 42.6	125 > 42.8
Athena	5	40 × 0.3 Msec	7.0	1100 > 43	650 > 43.1	360 > 43.2
WFXT*	10	24 × 0.5 Msec	24	2300 > 43.2	1300 > 43.4	600 > 43.5
WFXT*	5	4 × 3 Msec	4	1200 > 42.5	700 > 42.6	400 > 42.8
WFXT	5	60 × 0.2 Msec	60	6000 > 43.35	3200 > 43.4	1600 > 43.5
Super-Chandra	2	6 × 2 Msec	0.6	310 > 42.2	185 > 42.3	110 > 42.5
Super-Chandra	2	24 × 0.5 Msec	2.4	390 > 43	220 > 43.1	125 > 43.2
Super-Chandra	1	2 × 6 Msec	0.2	175 > 41.8	100 > 42.0	65 > 42
Super-Chandra	1	6 × 2 Msec	0.6	350 > 42.1	210 > 42.2	120 > 42.4

*Close to confusion limit (40 beams per source).

a focal plane camera with FOV ~ 1 deg^2. We assume that the observatory is in a low earth orbit (LEO), ensuring a low internal background (similar to that of the instrument on board Swift and Suzaku, which are flying on a LEO).

(3) We finally consider the concept of a Super-Chandra. This is a straw-man design for a mission with imaging capabilities comparable to Chandra (i.e., arcsec HPD), but using high throughput, light-weight mirrors (a concept pioneered by Elvis and Fabbiano some 15 years ago [33]). Good imaging capabilities using thin glass or nickel shells may be obtained by correcting the shell shape with actuators. Studies of active X-ray mirrors have been performed in the past with good results (see http://www.mssl.ucl.ac.uk/sxoptics/). A SPIE conference had been devoted to active X-ray mirrors in 2010 (proceedings of SPIE 7803). Active X-ray optics have been foreseen for extremely large throughput, subarcsec future missions like Generetion-X [34, 35] or, more recently, for a square meter, subarcsec mission (Vikhlinin et al. 2011, HEAD meeting). Here, we assume more modest throughput (~3000 cm^2 at 1-2 keV) and PSF (1-2 arcsec HPD). We also assume a limited FOV (0.1 deg^2) and a LEO, which ensures a low internal background.

We computed on axis sensitivities as a function of the observing time using the above parameters and assuming a signal to noise ratio of 3 for source detection. We assume that the effective area decreases linearly from the center to the limit of the FOV by 50% to make realistic predictions for the number of detected high-z AGN in each FOV. The background includes particle-induced internal background, as measured on HEO and LEO by Swift, Suzaku, XMM, and Chandra, cosmic X-ray background (CXB), and low temperature thermal X-ray background due to the local superbubble. To simulate the background expected for the three mission concepts, we have modified background scripts and files prepared to produce simulations for the IXO and NHXM missions. We find that the internal background dominates over the X-ray background (CXB and the local superbubble) above 0.5 keV in a HEO. Conversely, in LEO, the local superbubble dominates below 1 keV. We finally assumed a total net observing time of 12 Mseconds devoted to surveys, split in several shorther observations, to cope with source confusion and optimize the detection of $z > 5$ sources with 2–10 keV luminosity $\gtrsim 10^{42}$ ergs/s. The standard criterion for source confusion (40 beams per source) translates in a flux limit for source confusion of $\sim 6 \times 10^{-17}$ erg/cm^2/s in the 0.5–2 keV band for PSF HPD = 10 arcsec and just above 10^{-17} erg/cm^2/s for HPD = 5 arcsec. Source confusion is not an issue for any realistic exposure time for a PSF with HPD = 1 arcsec or below.

To estimate the faint X-ray sources number density, we used the model number counts in Figure 2, based on the luminosity functions presented in [9]. We conservatively assumed a flat faint-end slope of the X-ray luminosity functions, $\gamma 1 = 0.6$. Table 2 gives the predicted number of z = 4-5, z = 5–5.8, and $z > 5.8$ sources, along with their minimum 2–10 keV luminosity, for several indicative mosaics for the three mission concepts briefly described above.

It must be noted that the uncertainties on the number of sources in Table 2 is large. It is at least a factor of two at z = 4–6 and even larger at $z > 6$ (factor of 3 lower limit and a factor of 2 upper limit). The obvious message of Table 2 is that a wide field greatly helps in searching for faint high-z AGN (also see [36]). This is probably the only solution to collect samples of thousands X-ray AGN at $z > 4$. However, even a PSF as good as 5 arcsec HPD does not allow searching for sources fainter than 10^{43} ergs/s at z = 5–7. This means that only a mission with Chandra-like PSF but much higher throughput (>5 × Chandra effective area at 1-2 keV) would be able to target normal star-forming galaxies and miniquasars at z = 6-7. On one hand, a 2–10 keV luminosity of 10^{42} ergs/s at $z = 7$, reachable by extradeep exposures with a 1 arcsec PSF Super-Chandra, would be produced by a 7×10^5 SMBH emitting at its Eddington luminosity (assuming a bolometric correction of 10). Even smaller masses may be probed, if the accretion is super-Eddington. A Super-Chandra would then be able to directly search for the first generation of SMBH produced by monolithic collapse of $\gtrsim 10^5$ M$\odot$

gas clouds to BH [28–30, 37]. On the other hand, $L(2–10) = 10^{42}$ ergs/s is also produced by galaxies that form stars at a rate of $\gtrsim 200\,M_\odot$/yr ([38]). Since, at such high redshifts, the X-ray emission should mainly be due to high mass X-ray binaries, X-ray high-z galaxies could then be used to constrain the initial mass function at the epoch of galaxy formation. A Super-Chandra would then be able to open two brand-new fields in structure formation. Of course, it is not casual that the considered configuration for a Super-Chandra is able to reach these goals. Going back from scientific requirements to mission parameters, the goal of detecting sources with a 2–10 keV luminosity of $\sim 10^{42}$ at $z \sim 7$, in feasible exposure times, requires an effective area $\sim 3000\,cm^2$, given a PSF HPD $\sim$ 1 arcsec and assuming a LEO internal background.

Unfortunately a Super-Chandra is beyond the horizon of the present decade, both because of technological and programmatic issues. Furthermore, it is not clear whether a WFXT is truly feasible with such huge 1 deg^2 FOV and large throughput, and, in any case, it does not appear to be a priority in the latest US Decadal Survey (http://sites.nationalacademies.org/bpa/BPA_049810), nor in the ESA Cosmic Vision program. Conversely, Athena is a study mission for an L class mission in the framework of the ESA Cosmic Vision program. A decision on Cosmic Vision L class missions should be taken in February 2012. If positive, Athena could be implemented for the first years of the next decade. Although not reaching exquisite, Chandra-like image quality, nor extralarge field of view, Athena would be able to give a substantial contribution to the knowledge of the high-z Universe, with hundreds to a thousand $z > 4$ faint X-ray AGN (an improvement by a factor 10–100 with respect to today situation) and tens to hundreds $z > 5.8$ faint X-ray selected AGN (today, there are only 3-4 candidate $z > 6$ X-ray selected AGN in the literature [9, 26, 39], none of them spectroscopically confirmed so far).

Acknowledgments

This work was supported by ASI/INAF contracts I/024/05/0 and I/009/10/0. This work is based on observations made with NASA X-ray observatory Chandra. The authors thank the Chandra Director's office for allocating the time for these observations. X-ray data were obtained from the archive of the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory.

References

[1] A. Lamastra, N. Menci, R. Maiolino, F. Fiore, and A. Merloni, "The building up of the black hole-stellar mass relation," *Monthly Notices of the Royal Astronomical Society*, vol. 405, no. 1, pp. 29–40, 2010.

[2] K. Boutsia, A. Grazian, E. Giallongo et al., "A low escape fraction of ionizing photons of $L > L^*$ Lyman break galaxies at $z = 3.3$," *The Astrophysical Journal*, vol. 736, no. 1, article 41, 2011.

[3] S. Mitra, T. R. Choudhury, and A. Ferrara, "Joint quasar-cosmic microwave background constraints on reionization history," *Monthly Notices of the Royal Astronomical Society*, vol. 419, no. 2, pp. 1480–1488, 2012.

[4] M. Volonteri and M. J. Rees, "Rapid growth of high-redshift black holes," *The Astrophysical Journal*, vol. 633, no. 2, pp. 624–629, 2005.

[5] M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruszkowski, and F. Haardt, "Dual black holes in merger remnants—II. Spin evolution and gravitational recoil," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 682–690, 2010.

[6] A. R. King, J. E. Pringle, and J. A. Hofmann, "The evolution of black hole mass and spin in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 3, pp. 1621–1627, 2008.

[7] F. Fiore, "Multiwavelength perspective of AGN evolution," in *Proceedings of the International Conference on X-Ray Astronomy*, vol. 1248, pp. 373–380, 2010.

[8] A. Lamastra, N. Menci, F. Fiore, C. Di Porto, and L. Amendola, "Constraining dynamical dark energy models through the abundance of high-redshift supermassive black holes," submitted to *Monthly Notices of the Royal Astronomical Society*, http://arxiv.org/abs/1111.3800.

[9] F. Fiore, S. Puccetti, A. Grazian et al., "Faint high-redshift AGN in the *Chandra* deep field south: the evolution of the AGN luminosity function and black hole demography," *Astronomy and Astrophysics*, vol. 537, article A16, 22 pages, 2011.

[10] G. T. Richards, M. Lacy, L. J. Storrie-Lombardi et al., "Spectral energy distributions and multiwavelength selection of type 1 quasars," *The Astrophysical Journal*, vol. 166, no. 2, pp. 470–497, 2006.

[11] E. Glikman, S. G. Djorgovski, D. Stern, A. Dey, B. T. Jannuzi, and K.-S. Lee, "The faint end of the quasar luminosity function at $z \sim 4$: implications for ionization of the intergalactic medium and cosmic downsizing," *The Astrophysical Journal Letters*, vol. 728, no. 2, article L26, 2011.

[12] L. Jiang, X. Fan, F. Bian et al., "A survey of z ~ 6 quasars in the sloan digital sky survey deep stripe. II. Discovery of six quasars at $z_{AB} > 21$," *Astronomical Journal*, vol. 138, no. 1, pp. 305–311, 2009.

[13] C. J. Willott, P. Delorme, C. Reylé et al., "The Canada-France high-z quasar survey: nine new quasars and the luminosity function at redshift 6," *Astronomical Journal*, vol. 139, no. 3, pp. 906–918, 2010.

[14] D. J. Mortlock, S. J. Warren, B. P. Venemans et al., "A luminous quasar at a redshift of $z = 7.085$," *Nature*, vol. 474, no. 7353, pp. 616–619, 2011.

[15] P. Predehl, H. Böhringer, H. Brunner et al., "eROSITA on SRG," in *Proceedings of the International Conference on X-ray Astronomy*, vol. 1248, pp. 543–548, September 2009.

[16] Y. Q. Xue, B. Luo, W. N. Brandt et al., "The *Chandra* deep field-south survey: 4 Ms source catalogs," *Astrophysical Journal Supplement Series*, vol. 195, no. 1, article 10, 2011.

[17] A. Comastri, P. Ranalli, K. Iwasawa et al., "The XMM Deep survey in the CDF-S: I. First results on heavily obscured AGN," *Astronomy and Astrophysics*, vol. 526, no. 12, article L9, 2011.

[18] E. Treister, K. Schawinski, M. Volonteri, P. Natarajan, and E. Gawiser, "Black hole growth in the early Universe is self-regulated and largely hidden from view," *Nature*, vol. 474, no. 7351, pp. 356–358, 2011.

[19] R. J. Bouwens, G. D. Illingworth, J. P. Blakeslee, and M. Franx, "Galaxies at z ~ 6: the UV luminosity function and luminosity density from 506 HUDF, HUDF parallel ACS field, and goods i-dropouts," *The Astrophysical Journal*, vol. 653, no. 1 I, pp. 53–85, 2006.

[20] F. Fiore, A. Grazian, P. Santini et al., "Unveiling obscured accretion in the *Chandra* deep field-south," *The Astrophysical Journal*, vol. 672, no. 1, pp. 94–101, 2008.

[21] F. Fiore, S. Puccetti, M. Brusa et al., "Chasing highly obscured QSOs in the COSMOS field," *The Astrophysical Journal*, vol. 693, no. 1, pp. 447–462, 2009.

[22] C. J. Willott, "No evidence of obscured, accreting black holes in most $z = 6$ star-forming galaxies," *Astrophysical Journal Letters*, vol. 742, no. 1, article L8, 2011.

[23] F. Civano, M. Brusa, A. Comastri et al., "The population of high-redshift active galactic nuclei in the *Chandra*-cosmos survey," *The Astrophysical Journal*, vol. 741, no. 2, article 91, 2011.

[24] M. Brusa, A. Comastri, R. Gilli et al., "High-redshift quasars in the cosmos survey: the space density of Z > 3 X-ray selected qsos," *The Astrophysical Journal*, vol. 693, no. 1, article 8, 2009.

[25] N. A. Grogin, D. D. Kocevski, S. M. Faber et al., "Candels: the cosmic assembly near-infrared deep extragalactic legacy survey," *Astrophysical Journal Supplement Series*, vol. 197, no. 2, article 35, 2011.

[26] B. Luo, W. N. Brandt, Y. Q. Xue et al., "Identifications and photometric redshifts of the 2 Ms *Chandra* deep field-south sources," *The Astrophysical Journal*, vol. 187, no. 2, pp. 560–580, 2010.

[27] P. Madau and M. J. Rees, "Massive black holes as population III remnants," *The Astrophysical Journal*, vol. 551, no. 1, pp. L27–L30, 2001.

[28] G. Lodato and P. Natarajan, "Supermassive black hole formation during the assembly of pre-galactic discs," *Monthly Notices of the Royal Astronomical Society*, vol. 371, no. 4, pp. 1813–1823, 2006.

[29] M. Volonteri and M. C. Begelman, "Quasi-stars and the cosmic evolution of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 409, no. 3, pp. 1022–1032, 2010.

[30] M. C. Begelman, "Evolution of supermassive stars as a pathway to black hole formation," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 673–681, 2010.

[31] N. Fanidakis, C. M. Baugh, A. J. Benson et al., "Grand unification of AGN activity in the ΛCDM cosmology," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 1, pp. 53–74, 2011.

[32] S. Murray, R. Gilli, P. Tozzi et al., "The growth and evolution of super massive black holes," submitted to. *The Astrophysical Journal*, http://eprintweb.org/s/article/arxiv/0903.5272/.

[33] M. Elvis and G. Fabbiano, "Science driven arguments for a 10 sq.meter, 1 arcsecond X-ray telescope," *SPIE*. In press, http://arxiv.org/abs/astro-ph/9611178/.

[34] M. Elvis, R. J. Brissenden, G. Fabbiano et al., "Active X-ray optics for generation-X, the nxt high resolution X-ray observatory," *SPIE*. In press, http://arxiv.org/abs/astro-ph/0608533.

[35] S. L. O'Dell, R. J. Brissenden, W. N. Davis et al., "High-resolution x-ray telescopes," submitted to*Proceedings of the International Conference X-Ray Astronomy*, http://arxiv.org/abs/1010.4892v1/.

[36] R. Gilli, P. Tozzi, and P. Rosati, "Demography of obscured and unobscured AGN: prospects for a wide field X-ray telescope," *MemSAIT*. In press, http://arxiv.org/abs/1010.6024/.

[37] G. Lodato and P. Natarajan, "The mass function of high-redshift seed black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L64–L68, 2007.

[38] P. Ranalli, A. Comastri, and G. Setti, "The 2-10 keV luminosity as a star formation rate indicator," *Astronomy and Astrophysics*, vol. 399, no. 1, pp. 39–50, 2003.

[39] M. Salvato, O. Ilbert, G. Hasinger et al., "Dissecting photometric redshift for active galactic nucleus using *XMM*- and *Chandra*-cosmos samples," *Astrophysical Journal*, vol. 742, no. 2, article 61, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 516193, 21 pages
doi:10.1155/2012/516193

Review Article

The Cosmic History of Black Hole Growth from Deep Multiwavelength Surveys

Ezequiel Treister[1,2] and C. Megan Urry[3,4,5]

[1] *Departamento de Astronomía, Universidad de Concepción, Casilla 160-C, Concepción 4070386, Chile*
[2] *Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[3] *Yale Center for Astronomy and Astrophysics (YCAA), P.O. Box 208121, New Haven, CT 06520, USA*
[4] *Department of Physics, Yale University, P.O. Box 208121, New Haven, CT 06520, USA*
[5] *Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520, USA*

Correspondence should be addressed to Ezequiel Treister, etreiste@astroudec.cl

Received 21 September 2011; Accepted 30 November 2011

Academic Editor: Francesca Civano

Significant progress has been made in the last few years on understanding how supermassive black holes form and grow. In this paper, we begin by reviewing the spectral signatures of active galactic nuclei (AGN) ranging from radio to hard X-ray wavelengths. We then describe the most commonly used methods to find these sources, including optical/UV, radio, infrared, and X-ray emission, and optical emission lines. We then describe the main observational properties of the obscured and unobscured AGN population. Finally, we summarize the cosmic history of black hole accretion, that is, when in the history of the universe supermassive black holes were getting most of their mass. We finish with a summary of open questions and a description of planned and future observatories that are going to help answer them.

1. Introduction

Astrophysical black holes come in a wide range of masses, from $\gtrsim 3M_\odot$ for stellar mass black holes [1] to $\sim 10^{10}M_\odot$ for so-called supermassive black holes [2, 3]. The best evidence for the existence of a supermassive black hole can be found in the center of the Milky Way galaxy, where from dynamical studies, the mass of the Sgr A* source was established to be $\sim 4.4 \times 10^6 M_\odot$ [4, 5].

Evidence for the existence of supermassive black holes has also been found in other massive nearby galaxies [6], mostly from resolved stellar and gas kinematics. For active galaxies, it has been possible to use the technique known as reverberation mapping [7–9]. From these observations, a clear correlation has been established between the mass of the central black hole and properties of the host galaxy such as stellar mass in the spheroidal component [10], luminosity [11], velocity dispersion [12, 13] and mass of the dark matter halo [14]. The fact that such correlations exist, even though these components have very different spatial scales, suggests a fundamental relationship between black hole formation and galaxy evolution. Furthermore, it is now well established by simulations [15] that the energy output from the growing central black hole can play a significant role in the star formation history of the host galaxy. In particular, theory suggests that nuclear activity regulates star formation either by removing all the gas [16, 17] or by heating it [18]. It is therefore obvious that a complete study of galaxy evolution requires a comprehensive understanding of black hole growth.

Most current black hole formation models tell us that the first black hole seeds formed at $z \gtrsim 15$. While the exact mechanism for the formation of the first black holes is not currently known, there are several prevailing theories (see the comprehensive reviews by Rees [19] and Volonteri [20] for more details). One of the most popular possibilities is that the first black hole seeds are the remnants of the first generation of stars, the so-called population III stars, formed out of primordial ultralow metallicity gas. These black holes formed at $z \sim 20$ and have typical masses ~ 100–$1,000 M_\odot$.

This scenario has problems explaining the very high masses, of $\sim 10^9 M_\odot$, estimated for supermassive black holes in $z \sim 6$ opticallyselected quasars [21]. Alternatively, the first black holes could have formed directly as the result of gas-dynamical processes. It is possible for metal-free gas clouds with $T_{\rm vir} \gtrsim 10^4$ K and suppressed H_2 formation to collapse very efficiently [22], possibly forming massive black hole seeds with $M \sim 10^4$-$10^5 M_\odot$ as early as $z \sim 10$–15. If instead the UV background is not enough to suppress the formation of H_2, the gas will fragment and form "normal" stars in a very compact star cluster. In that case, star collisions can lead to the formation of a very massive star, that will then collapse and form a massive black hole seed with mass $\sim 10^2$–$10^4 M_\odot$ [23].

Given the current masses of $10^{6-9} M_\odot$, most black hole growth happens in the active galactic nuclei (AGN) phase [2, 24]. With typical bolometric luminosities $\sim 10^{45-48}$ erg s^{-1}, AGN are amongst the most luminous emitters in the universe, particularly at high energies and radio wavelengths. These luminosities are a significant fraction of the Eddington luminosity—the maximum luminosity for spherical accretion beyond which radiation pressure prevents further growth—for a $10^{8\text{-}9} M_\odot$ central black hole. A significant fraction of the total black hole growth, ~60% [25], happens in the most luminous AGN, quasars, which are likely triggered by the major merger of two massive galaxies [26]. In an AGN phase, which lasts $\sim 10^8$ years, the central supermassive black hole can gain up to $\sim 10^7$-$10^8 M_\odot$, so even the most massive galaxies will have only a few of these events over their lifetime. Further black hole growth, mostly in low-luminosity (low Eddington rate) AGN, is likely due to stochastic accretion of cold gas, mostly in spiral galaxies [27].

According to the AGN unification paradigm [28, 29], a large fraction of these sources, ~75% locally, are heavily obscured by optically and geometrically thick axisymmetric material, which explains many of the observed differences among different types of active galaxies. In addition, luminosity [30] and cosmic epoch [31] play a significant role. One constraint on the fraction of obscured AGN and its evolution comes from the spectral shape of the extragalactic X-ray "background" (XRB). Thanks to deep X-ray observations at $E \lesssim 10$ keV performed by *Chandra* and *XMM-Newton*, a very large fraction of the X-ray background, ~80%, has been resolved into point sources [32], the vast majority of them AGN [33]. Several studies, the first of them ~20 years ago [34], have used a combination of obscured and unobscured AGN to explain the spectral shape and normalization of the X-ray background with overall good results. The latest AGN population synthesis models [35, 36] assume an average ratio of obscured to unobscured AGN of ~3 : 1 locally, increasing towards lower luminosities and higher redshifts, as well as a fraction of Compton-thick sources (CT; $N_H > 10^{24}$ cm^{-2}) of ~5–10%, consistent with the value observed at higher energies, $E = 10$–100 keV, of ~5% by INTEGRAL in the local universe [35, 37], lower by factors of ~3 than expectations of previous population synthesis models [38, 39].

In this paper, we review multiwavelength methods used to trace the growth of SMBHs (Section 2), the known properties of unobscured and obscured AGN (Sections 3 and 4, resp.), the cosmic history of black hole accretion (Section 5), and prospects for future observations (Section 6). Throughout this paper, we assume a ΛCDM cosmology with $h_0 = 0.7$, $\Omega_m = 0.27$, and $\Omega_\Lambda = 0.73$, in agreement with the most recent cosmological observations [40].

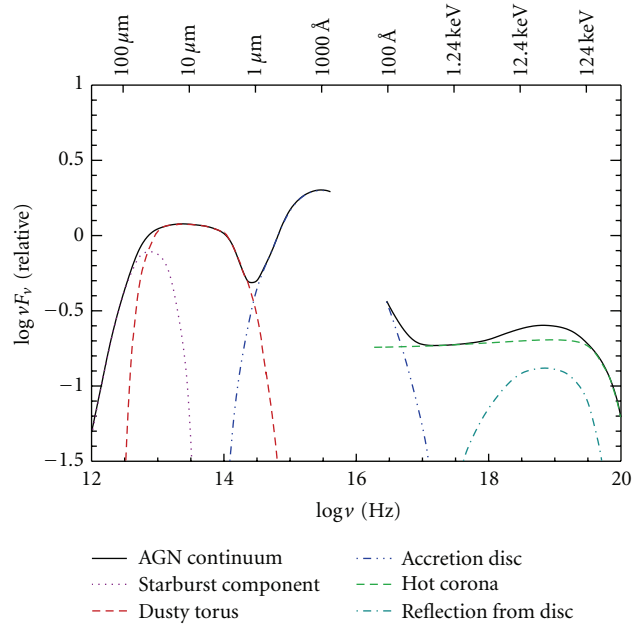


FIGURE 1: Average unobscured radio-quiet AGN spectrum from far-infrared to hard X-rays, as compiled by Manners (2002; [41]). The contributions from each component are shown separately together with the total AGN emission (*solid line*). Most of the AGN radiation appears in three regions: infrared (reemission from dust; *dashed red line*), UV and optical (accretion disk; *dot-dot-dot-dashed blue line*), and X-rays (hot corona and reflection from the accretion disk; *dashed green and dot-dashed cyan lines*). At longer wavelengths, >100 μm, a starburst component associated with the host galaxy dominates (*dotted magenta line*).

2. How to Trace SMBH Growth?

One of the main features of the AGN emission is that it covers a very wide range of wavelengths, from radio to Gamma-rays (Figure 1). While unobscured sources are easily detectable and identified by their UV and soft X-ray continuum and their broad optical emission lines, obscured AGN can only be found at longer, mid-IR, wavelengths or in hard X-rays. Of course, selections at different wavelengths have different biases. For example, while radio surveys are not particularly affected by obscuration, they are most likely to detect radio-loud sources, which are only ~10% of the total AGN population at bright fluxes [42]. However, combining different multiwavelength techniques gets us closer to a complete picture. Below, we briefly describe the most popular AGN selection methods, their advantages, and drawbacks.

2.1. Optical/UV Continuum. Rest-frame optical/UV selection of AGN, in particular for high-luminosity unobscured quasars, is particularly efficient because the spectral shapes of

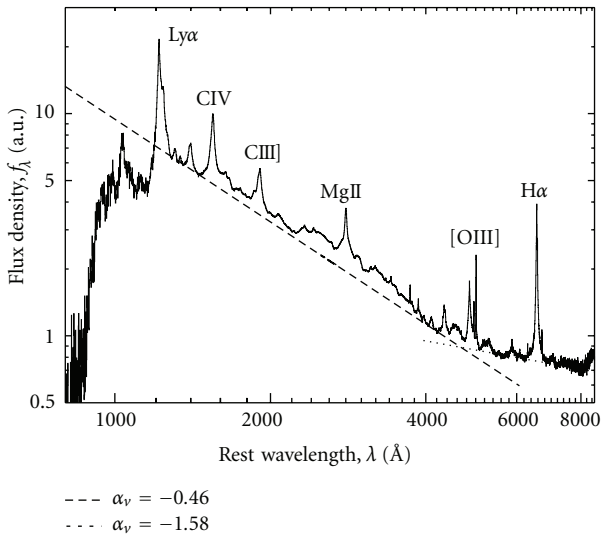


FIGURE 2: Composite rest-frame optical/UV spectrum for the optically selected quasars in the Sloan Digital Sky Survey, from the work of Vanden Berk et al. [49]. The *dashed* and *dotted* lines show power-law fits to the continuum emission.

stars and quasars at these wavelengths produce very different broadband colors due to the presence of the "big blue bump" [43] in quasar spectra from ∼100 Å to ∼1 μm (Figure 2). This emission is often attributed to the thermal radiation with temperatures ∼30,000 K originating in the accretion disk [44]. This unique spectral shape has been used in the past to identify quasars with great success by optical surveys such as the Palomar-Green (PG) survey [45], the 2 degree field QSO redshift survey [46], or more recently the Sloan Digital Sky Survey (SDSS; [47]), which has found now more than 1 million quasars [48].

However, samples selected in the optical are far from complete, as emission at these wavelengths is strongly affected by reddening or extinction, and most AGN are obscured along our line of sight. Furthermore, for lower-luminosity sources, the optical light from the host galaxy can easily outshine the nuclear emission. This is particularly important for ground-based observations and high-redshift sources, for which it is very hard to separate the nonthermal and stellar components spatially.

2.2. Radio. Historically, identification of AGN based on their radio emission has been very important. In fact, the first discovered quasar, 3C 273, was originally classified as a radio source [71]. In spite of this, radio selection can be very problematic. In radio-loud sources (defined as $f_{5\,\mathrm{GHz}}/f_B > 10$ [72]) radio emission is associated with a strong, nonthermal, component, probably originating in a beamed collimated relativistic jet [73]. In radio-quiet sources, which are typically ∼3 orders of magnitude fainter at these wavelengths [74], the radio emission corresponds to the long-wavelength tail of the far-infrared dust emission. As a consequence, radio-selected samples are necessarily biased towards radio-loud sources, which represent only ∼10% of the overall AGN population.

2.3. Optical Emission Lines. As first reported by Baldwin et al. [81], the photoionizing spectrum of a power law continuum source, such as an AGN, produces very different emission line intensity ratios when compared with that of typical star-forming regions (mostly due to O and B stars). Hence, emission lines can be used to identify the presence of AGN even in galaxies in which the optical/continuum does not show any direct AGN signature, due to obscuration and/or low luminosity. Because the AGN ionizing emission reaches material even a few kiloparsecs away from the nuclear region, this selection technique is less sensitive to circumnuclear obscuration and thus provides a more complete AGN view when compared with, for example, optical/UV continuum selection. This technique was used successfully in the SDSS [50, 51] to extend the low-redshift AGN sample to lower luminosities. Emission line ratios and diagnostic regions can be seen in Figure 3. Emission-line selection can also be used at higher redshifts, as shown by the DEEP2 galaxy redshift survey, which selected a sample of 247 AGN at $z \sim 1$ from optical spectroscopy using the DEIMOS spectrograph at the Keck observatory [82].

While this is an efficient AGN selection technique, optical spectroscopy is very expensive in telescope time and is only feasible for relatively bright emission line regions. This selection may be incomplete at the low luminosity end, if the host galaxy can outshine the high-ionization emission lines. It is currently very difficult to extend this selection beyond $z \sim 1$, as the relevant emission lines move to observed-frame near-IR wavelengths, where current-generation spectrographs are significantly affected by atmospheric emission, do not cover wide field of views, and have limited multiobject capabilities.

2.4. X-Rays. As was found more than 30 years ago, AGN are ubiquitous X-ray emitters [89]. Their X-ray emission extends from ∼0.1 keV to ∼300 keV and is attributed to inverse-Compton scattering due to high-energy electrons in a hot corona, surrounding the accretion disk. The high-energy cutoff at ∼100–300 keV is presumably due to a cutoff in the energy distribution of the electrons in the hot corona. AGN are typically ∼1–5 orders of magnitude more luminous in X-rays than normal galaxies, which makes them the dominant extragalactic population at these wavelengths. Most AGN are obscured by photoelectric absorption by gas along the line of sight, which preferentially affects emission at the lower energies. This is often parametrized by the amount of neutral hydrogen column density in the line of sight. Figure 4 shows typical AGN X-ray spectra including the power law component and high energy cutoff, for different levels of photoelectric absorption.

Deep X-ray surveys with *Chandra* and *XMM-Newton* have found the largest AGN densities, ∼7,000 deg^{-2} [90], ∼10–20 times higher than even the deepest optical surveys. Still, X-ray selected AGN samples are still biased against the most obscured sources. In fact, even the deepest *Chandra* surveys can miss more than half of the total AGN due to a

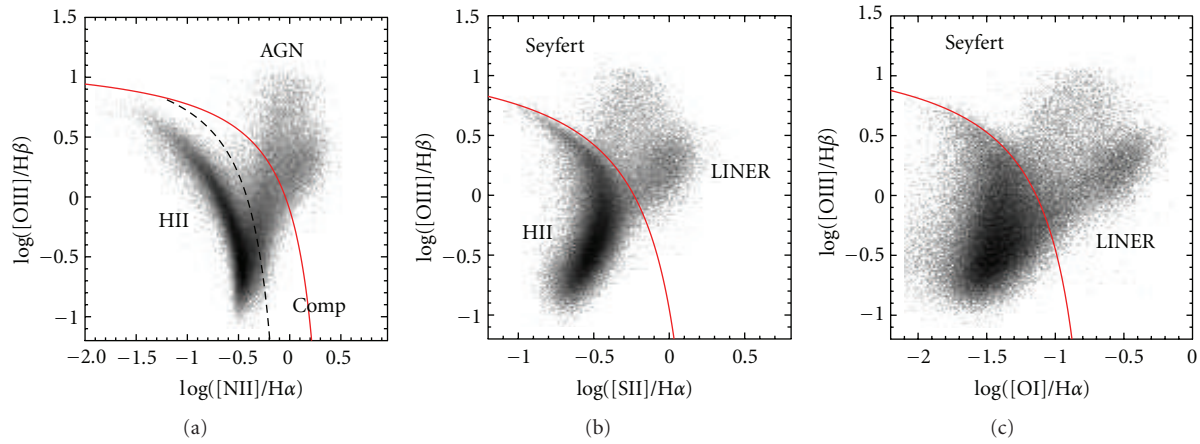


Figure 3: AGN selection diagrams based on optical emission line ratios. Figure obtained from Kewley et al. [50]. Divisions between star-forming galaxies and AGN are shown by the *dashed* [51] and *solid* [52] lines; bona-fide AGN sit at the upper right in these distributions.

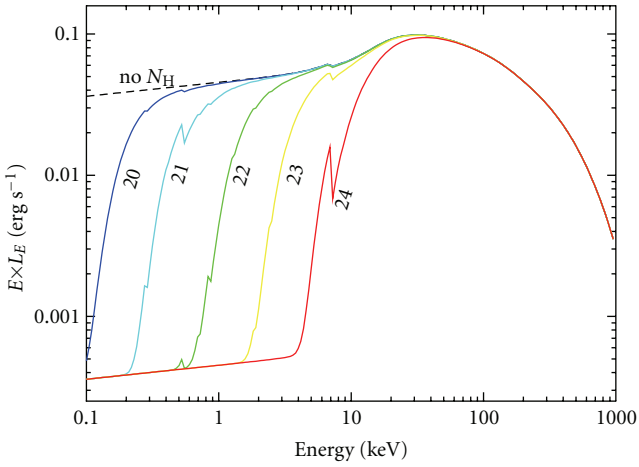


Figure 4: Typical AGN X-ray spectrum from 0.1 to 100 keV. Each spectrum includes a power law with photon index $\Gamma = 1.9$, a Compton reflection hump, peaking at ~30 keV, a high-energy cutoff at 300 keV, and photoelectric absorption. Numbers for each curve indicate the amount of absorption, given as $\log(N_H)$, with N_H in units of atoms cm^{-2}.

combination of obscuration and/or low luminosity [91]. Observing at higher energies helps to alleviate the effects of obscuration. Wide-area surveys performed by the *Swift* [66] and *INTEGRAL* [92] satellites have detected a large number of AGN in the local universe at $E = 10$–100 keV, where all but the most obscured, Compton-thick, sources emit strongly. Unfortunately, due to their relatively high flux limits (~2-3 order of magnitudes shallower than the deeper *Chandra* observations), surveys at these energies are so far limited to low redshifts only.

2.5. Infrared. Intrinsic AGN emission is not particularly strong at near/mid-IR wavelengths. Radiation coming from the accretion disk, often characterized as a power-law, while very strong at UV and optical wavelengths decreases rapidly beyond ~1 μm [93]. However, as was originally found by *IRAS* [74] and latter confirmed by *ISO* [94] and *Spitzer* [95], AGN are luminous IR sources. This is commonly attributed to reemission of absorbed energy by dust. This radiation can be found starting at ~2-3 μm, which corresponds to the dust sublimation temperature, about ~1,000-2,000 K [96], and extends to ~100 μm, at the tail of the black body spectrum for a ~100–1,000 K distribution [97], where the emission from the host galaxy typically starts to dominate. Typical AGN infrared luminosities are $10^{44–46}$ erg s^{-1} and thus they represent a significant fraction, ~30% on average [98], of the bolometric luminosity.

One clear advantage of infrared AGN selection is that this reemission is mostly isotropic, and hence both obscured and unobscured sources have similar detection probabilities. Therefore, it provides a complementary approach to the most common selection techniques described above, which are less sensitive to obscured sources. However, star forming galaxies are very luminous at these wavelengths as well, so that the host galaxy can easily outshine the central emission, in particular for low-luminosity sources, hence yielding a low efficiency in detecting AGN [99]. Furthermore, the selection function in infrared studies is more complicated, as the probability of detecting the AGN depends on the properties of the host galaxy, such as the amount of dusty star-formation, which is in principle independent of the nuclear luminosity.

3. Unobscured AGN

Because quasars are the most luminous, and thus easily detectable, members of the AGN family, the luminosity function of optical quasars has been well determined for years. In particular, it was found that the number of quasars evolve strongly [100] and peak at $z \simeq 2$ [101]. This evolution has been modeled either as a pure luminosity evolution (PLE), in which the characteristic luminosity changes with redshift while the shape of the luminosity function remains the same [102], or a pure density evolution (PDE), so that only the normalization of the luminosity function depends on redshift [100]; however, it was quickly discovered that

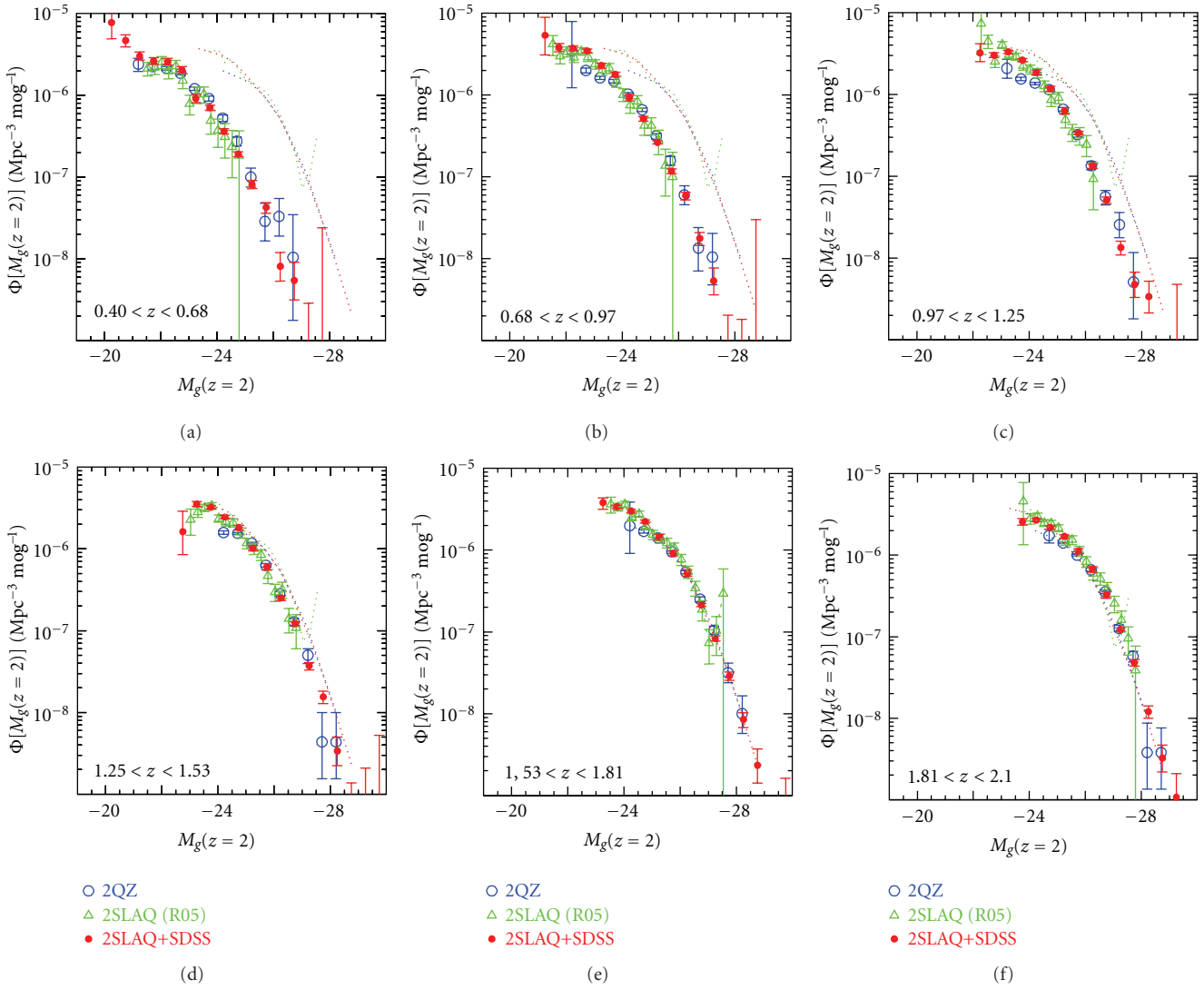


FIGURE 5: Binned optical quasar luminosity function from the work of Croom et al. ([53]; their Figure 11) in six redshift intervals, from $z = 0.4$ to $z = 2.1$. Measurements where obtained from the 2dF QSO Redshift Survey (2QZ; [54] *blue*), the 2dF-SDSS LRG and QSO survey (2SLAQ; [55]; *green*), and combining the latter with the general SDSS QSO sample ([56]; *red*). The dotted line shows the measured luminosity function in the $1.53 < z < 1.81$ range for reference. The strong evolution of the QSO luminosity function is clearly seen in this figure. This evolution is best fitted by a LEDE model, as described in the text.

at least in the PG quasar survey, the shape of luminosity function also evolves with redshift and thus neither a PLE nor a PDE provides a good description [45].

In Figure 5 we show one of the latest measurements of the quasar luminosity function at $0.4 < z < 2.6$, from Croom et al. [53]. They conclude that a luminosity-dependent density evolution provides a better fit to the optical quasar luminosity function, than either a PLE or PDE. Similar conclusions were reached by studying X-ray selected sources using soft X-ray (0.5–2 keV) observations [103] or hard X-ray (2–10 keV) data [79, 104] which includes both obscured and unobscured AGN. However, taking advantage of the large number of sources in their sample, ~10,000, Croom et al. [53] concluded that the best fit to the observed luminosity function is obtained by using a model based on a luminosity evolution + density evolution (LEDE). The most important difference between a LEDE and a PLE fit is a change in amplitude and bright-end slope at high redshift.

An important conclusion obtained from observations of the quasar luminosity function is the evidence for "cosmic downsizing" [105], that is, that the most massive black holes get most of their mass at high redshift, while at low redshift only low mass black holes are still growing. This is observed both in optical [53] and hard X-ray luminosity functions [104, 105], thus indicating that this result is independent of obscuration. Recent deep optical surveys such as the Great Observatories Origins Deep Survey (GOODS; [106]), the Cosmic Evolution Survey (COSMOS; [59]), the NOAO Deep Wide Field Survey (NDWFS), and the Deep Lens Survey (DLS; [58]) have produced significant advances in extending

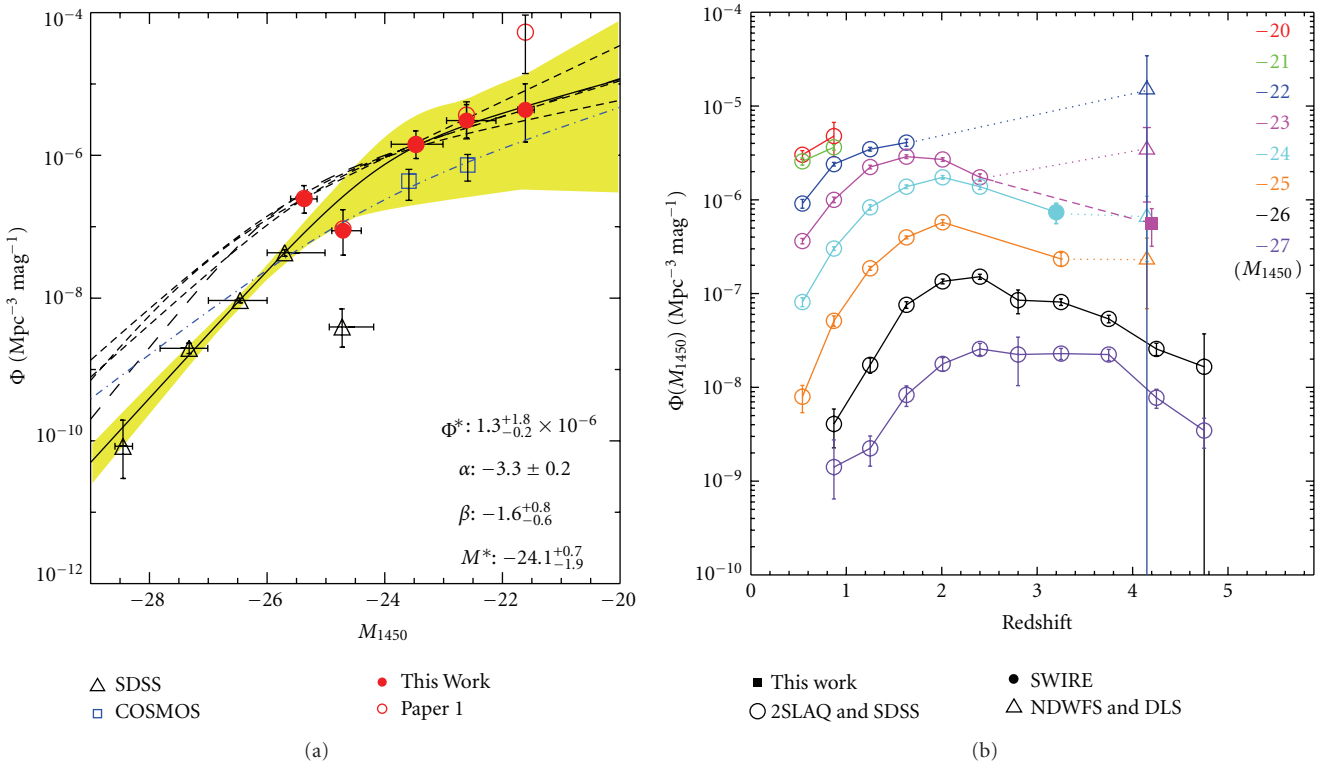


FIGURE 6: (a) $z \sim 4$ quasar luminosity function, from the work of Glikman et al. [57]. The *filled red circles* show the measurements presented on that work, while the *open red circles* were reported previously by the same group [58] and show the change with increased spectroscopic completeness. The *triangles* were obtained from the SDSS quasar sample at $z = 4.25$ [56]. The *blue squares* are the space densities of $z \sim 4$ QSOs from Ikeda et al. [59] and the *dot-dashed line* is their best-fit double power law. The lower right-hand legend lists the best-fit parameters to a double power-law (*solid line*) shaded region represents the 1σ uncertainties. Dashed and dotted lines show the $z \sim 3$ QLF [60], representing different fits to the observed QLF. (b) Quasar space density as a function of redshift, from the work of Ikeda et al. [59]. *Dotted lines* used the combined 2SLAQ, SDSS, SWIRE, NDFWS and DLS samples, while the *dashed lines* combine the COMOS and 2SLAQ sources. While AGN downsizing is clearly visible at $z < 2.5$, at higher redshifts the situation is more uncertain.

the quasar luminosity function to higher redshifts, $z > 3$. Figure 6 shows the quasar luminosity function at $z \sim 4$ and the redshift dependence of the quasar spatial density [58, 106]. While the presence of downsizing is clear up to $z \sim 2.5$, at higher redshifts it is less convincing, most likely due to poor statistics and incompleteness. As argued by Glikman et al. [57], the slope of the faint end of the quasar luminosity function is critical in determining the contribution of these sources to the ionization of the intergalactic medium. Based on current measurements, quasars contribute ~60% of the ionizing photons at $z \sim 4$ and thus are the dominant source at this redshift.

At even higher redshifts, $z \sim$ 5-6, current deep surveys do not cover enough area to detect a significant number of sources. However, wide-area survey such as SDSS [107] and the Canada-France High-z Quasar Survey [108] have been able to find a sizable sample, ~40 high-luminosity quasars, at these high redshifts. According to these samples, there is a large decrease in the number density of high-redshift quasars, when compared to $z \sim 2$, suggesting that the peak of the quasar activity is at $z \sim 2.5$ [56]. The decline towards higher redshifts is given by $10^{-0.47z}$ [107] from $z = 3$ to $z = 6$. A similar trend is also observed for high-luminosity sources in X-ray surveys [109], which are also dominated by unobscured sources. This indicates that due to their very low spatial density, unobscured quasars do not contribute significantly to the early hydrogen reionization of the intergalactic medium at $z \sim 6$ [107, 108], in contrast to the situation at $z \sim 4$.

4. Obscured Accretion

The space density and evolution of the unobscured AGN population has been well studied, mostly from optical and soft X-ray surveys. However, we know that a large fraction of the SMBH growth happens in heavily obscured systems. Observations of the nearest AGN suggest that the local ratio of obscured to unobscured sources is ~4 : 1 [110]. A similarly high fraction of obscured AGN has been used to explain the spectrum and normalization of the extragalactic XRB, as shown by the latest AGN population synthesis models [35, 38, 39, 111]. The XRB gives an integral constraint to the AGN population and its evolution; the most recent deep surveys show that ~90% of the observed 2–8 keV XRB radiation can be attributed to resolved AGN [32]. In Figure 7, we show the latest AGN population synthesis models for

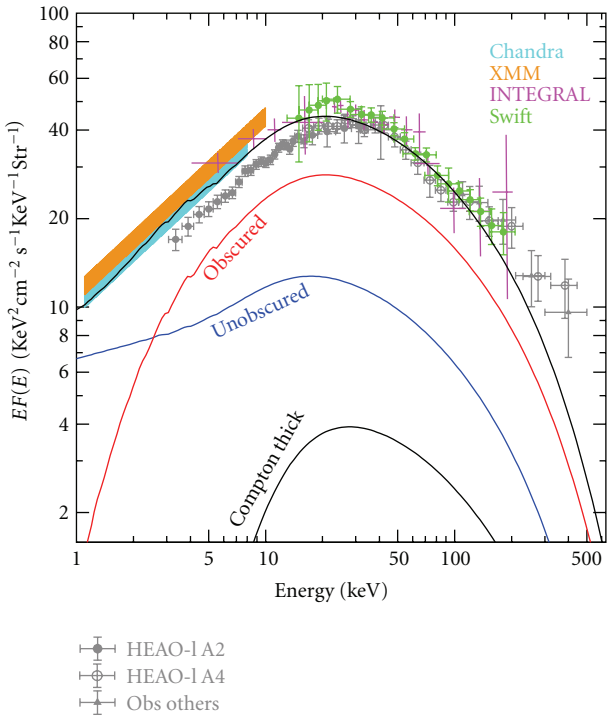


FIGURE 7: Observed spectrum of the extragalactic XRB from HEAO-1 [61], Chandra [32], XMM [62], INTEGRAL [63], and Swift [64] data. The *thick black solid line* shows the population synthesis model for the XRB spectrum of Treister et al. [35]. *Red*, *blue*, and *thin black* solid lines show the contribution to this model from unobscured, obscured Compton thin, and CT AGN, respectively.

the XRB which uses a local ratio of obscured to unosbcured AGN of ~3 : 1, plus a luminosity and redshift dependence, as described below [35]. The largest uncertainty in such models stems from the normalization mismatch of the data in the 1–10 keV energy range.

A possible dependence of the fraction of obscured AGN on luminosity was first suggested nearly 20 years ago [30] and confirmed since then by hard X-ray surveys [104, 105, 112]. A possible physical explanation is the so-called "receding torus," in which the size of the inner opening angle depends on luminosity [30, 113]. More recent observations found a luminosity dependence of the ratio of mid-IR to bolometric flux for unobscured AGN, consistent with this idea [98]. Alternatively, it has been proposed that the observed dependence of the obscured fraction on luminosity could be explained either by the effects of photoionization on the X-ray obscuring matter [114] or by the Eddington limit on a clumpy torus [65]. In Figure 8 we show the observed fraction of obscured AGN as a function of luminosity obtained by combining data from deep *Chandra* X-ray surveys [115]. A consistent result is observed for AGN selected in a hard X-ray (E = 14–195 keV) survey, as shown in Figure 8(b), indicating that this trend is not due to selection biases.

In the Figure 8(a) we compare the observed dependence of the fraction of obscured sources on luminosity with the expectations for different geometrical parameters of the obscuring material. If the height of the torus is roughly independent of luminosity, the change in covering fraction is due to a change in inner radius (the original "receding torus" model), hence a rough $L^{-1/2}$ dependence for the contrast should be expected [30, 116]. If the effects of radiation pressure are incorporated, in the case of a clumpy torus, a $L^{-1/4}$ dependence is expected. As can be seen in Figure 8, a $L^{-1/2}$ dependence is too steep compared to observed data. This implies that the height of the obscuring material cannot be independent of the source luminosity and provides evidence for a radiation-limited structure.

The dependence of the fraction of obscured AGN on redshift is more controversial. While some studies [31, 80, 111, 117, 118] found a small increase in the fraction of obscured AGN at higher redshifts, other results suggest that this fraction is constant [104, 119]. These discrepancies can be understood due to a combination of small samples and the use of N_H to classify AGN, which produces a well-known redshift bias [119]. The fraction of obscured AGN as a function of redshift for a large, ~2,000 sources, X-ray selected sample [115], using optical emission lines to separate obscured and unobscured AGN, is shown in Figure 9. It increases significantly with redshift, roughly as $(1+z)^{\alpha}$, with α = 0.3–0.5 (thin dashed lines, bottom panel, Figure 9; best fit, $\alpha \backsimeq 0.4$, thick dashed line). This value of α does not change significantly if a different host galaxy evolution is assumed, and it is consistent with the value of 0.3 reported by other studies [111, 117, 118].

Since star-forming galaxies may be expected to have more dust, the increase in the relative fraction of obscured AGN at high redshift may be due to an increase in the contribution to obscuration by galactic dust. By combining hard X-ray and mid-infrared observations, a similar ratio of hard X-ray to mid-infrared flux for obscured and unobscured AGN has been found [120], contrary to the predictions of the simplest AGN unification paradigm, in which the obscuration comes from the dusty torus and therefore the mid-infrared emission is reduced due to self-absorption. This result can be explained if the obscuration comes from a much more extended region, that is, kiloparsec, galactic scales rather than a compact parsec-scale torus. Furthermore, signatures for extended absorbing regions have been detected in nearby galaxies like NGC 1068 [121] and NGC 4151 [122]. Heavy absorption at kiloparsec scales has routinely been found in ultraluminous infrared galaxies (ULIRGs), which suffer a very strong evolution [123]. Hence, it seems likely that the change in the relative fraction of obscured AGN could be related to galactic-scale absorption, in particular since some ULIRGs also contain an obscured AGN (e.g., Arp 220; [124]).

Below, we review in more detail our current knowledge of the obscured AGN population at three different cosmic epochs: $z \simeq 0$, z = 1–3 and $z > 6$.

4.1. Obscured AGN in the Local Universe. Nearby AGN are found in the so-called Seyfert galaxies [125], which are known to host low luminosity and/or obscured nuclear activity [110, 126]. These growing supermassive black holes have been identified because of their high-ionization optical emission lines and in some cases their blue UV/optical

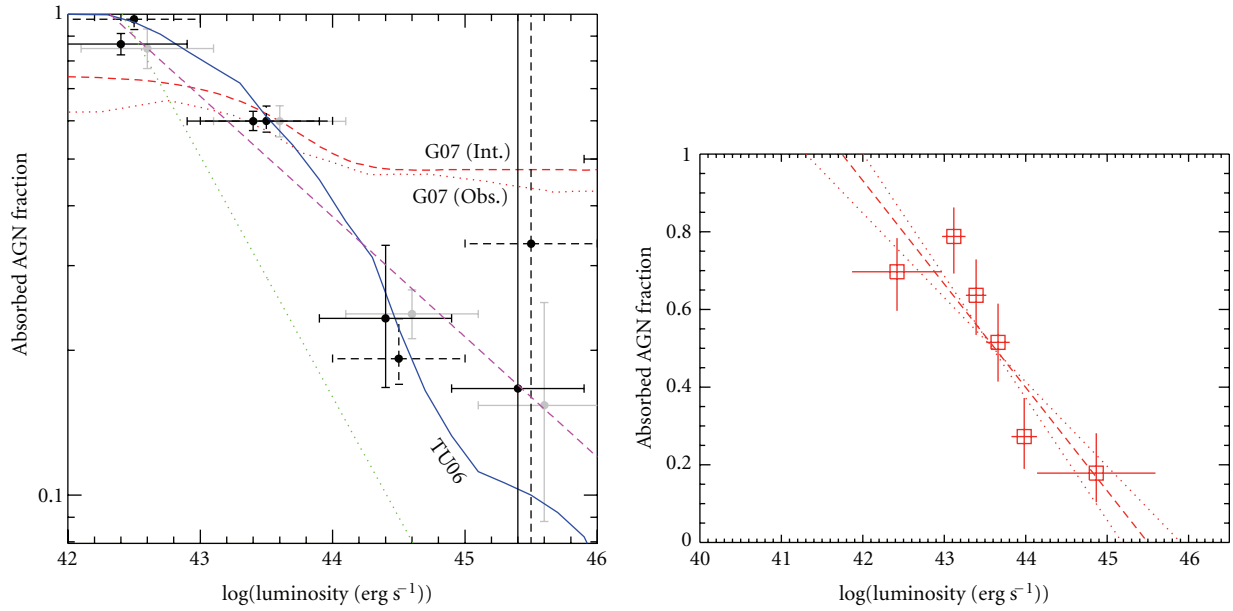


FIGURE 8: Ratio of obscured to total AGN as a function of hard X-ray luminosity. The black circles with *dashed error bars* show the obscured AGN fraction from the extended *Chandra* deep field south (ECDF-S) alone, while the *gray circles* show the results obtained using the meta-analysis by Treister and Urry [31]. Black circles with *solid error bars* show the fraction obtained combining these two samples. The dependence in the AGN population synthesis model of Gilli et al. [39] for the intrinsic and observed fractions of obscured AGN is shown by the *dashed* and *dotted* red lines; the dependence used by Treister and Urry [31] is shown by the *solid blue line*. The *dashed magenta line* shows the expected dependence for a radiation-limited torus [65], while the *dotted green line* shows the expectation for the original "receding torus" [30], both normalized to the observed value in the $10^{42\text{-}43}$ erg s^{-1} bin. (b) Same as (a) but using an AGN sample at $z \sim 0$ selected in hard X-rays from *Swift*/BAT observations [37]. The fact that the same luminosity dependence is observed in both samples indicates that it is not due to selection effects.

continuum. The first nearby AGN catalogues, produced ~40 years ago [127], contained ~200 quasars. Roughly speaking, ~5–15% of the galaxies near the Milky Way contain an active nucleus [110], and ~75% of these active galaxies are obscured. In fact, two of the three nearest AGN are Compton-thick (NGC 4945 and the Circinus galaxy [128]). Hence, optical surveys are not particularly efficient in unveiling this accretion, while observations at other wavelengths, in particular in the infrared [129], and hard X-rays, are more complete.

Surveys at hard X-ray energies, $E > 15$ keV, have been very successful in providing the most complete AGN samples in the local Universe. As long as the neutral hydrogen column density is lower than $\sim 10^{24}$ cm^{-2}, the direct AGN emission is mostly unaffected at these energies. Current observations at $E > 10$ keV with the International Gamma-Ray Astrophysics Laboratory (INTEGRAL; [130]) and Swift [131] satellites are available only at relatively high fluxes and hence low redshifts, $z < 0.05$.

Using the IBIS coded-mask telescope [132], INTEGRAL surveyed ~80% of the sky down to a flux of 5 mCrab in the 17–60 keV. Krivonos et al. [67] report the properties of 130 AGN detected in these all-sky observations. A large number of unidentified sources remain in this full INTEGRAL catalog (48) but only seven are found at high galactic latitude ($|b| > 5°$) and thus are likely of extragalactic origin. Five of the 130 known AGN are Compton thick. Using similar observations from the all-sky Swift/BAT survey, a catalog of 103 AGN [66] contains five AGN with estimated N_H greater than 10^{24} cm^{-2}. However, we caution that some of these N_H measurements were obtained by fitting a single absorbed power law to the X-ray spectrum, while heavily absorbed AGN have more complex spectra [133, 134], so the N_H estimates are likely to be lower limits.

Figure 10 shows the cumulative number counts of AGN, with CT sources shown separately, as a function of hard X-ray flux. In order to avoid the necessity of specifying a standard spectrum to convert fluxes to different energy bands, we show the INTEGRAL and Swift sources separately, but note that there is good agreement (within ~40%) in the normalization between the two distributions if a standard band conversion is assumed. At these high fluxes the slope of the $\log N$-$\log S$ is Euclidean, implying a uniform spatial distribution, as expected given the low redshifts of these sources. The number of CT AGN found by these surveys is surprisingly low, compared to the sample of known CT AGN in the local universe, most likely due to the effects of obscuration even at these high energies [35, 37, 68]. A study of optically selected local Seyfert 2 galaxies with hard X-ray information [126] found 12 CT AGN in a total of 45 Seyfert galaxies. Three were detected by Chandra and/or XMM, while the rest are mostly reflection-dominated sources, too faint to be detected by either INTEGRAL or Swift even though they are nearby, moderate-luminosity AGN. This suggests

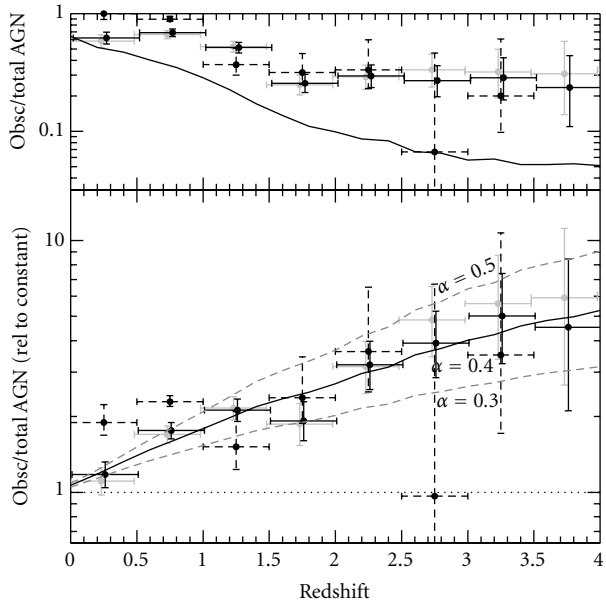


FIGURE 9: Fraction of obscured AGN as a function of redshift. *Upper panel*: direct measurements using the sources on the ECDF-S field only (*black circles with dashed error bars*), from the sample of Treister and Urry ([31]; *gray circles*) and combining both samples (*black circles with solid error bars*). The expected observed fraction for an intrinsic fraction of 3 : 1 obscured to unobscured AGN, accounting for optical and X-ray selection effects, is shown by the *black solid line*. As can be seen, while the observed fraction of obscured AGN declines toward higher redshifts, if the X-ray and optical selection effects and the luminosity dependence of the obscured AGN fraction are taken into account, this decline should be even stronger. *Bottom panel*: Inferred fraction of obscured AGN relative to an intrinsically constant fraction after correcting for selection effect and including the luminosity dependence of the obscured AGN fraction. Symbols are the same as for the upper panel. The corrected fraction of obscured AGN increases with redshift following a power law of the form $(1+z)^{\alpha}$ with $\alpha = 0.4 \pm 0.1$.

that even hard X-ray surveys miss quite a few Compton, thick AGN.

The observed fraction of CT AGN in the INTEGRAL and Swift/BAT hard X-ray selected samples is low, ∼5%. A very similar and consistent value, 4.6%, was recently obtained from a sample of 307 objects detected in the three-year all-sky *Swift*/BAT survey [37]. This was initially surprising, since previous AGN population synthesis models that can explain the XRB used much higher CT fractions of ∼15–20% [38, 39], that is, factors of 3-4 higher. We now know that even observations at these high energies can be affected by obscuration, if the column density is high enough. For example, ∼50% of the source flux in the 15–55 keV range can be lost if $\log N_H > 24.5$ [135]. As pointed out by Malizia et al. [68], and as can be clearly seen in Figure 11, the INTEGRAL all-sky observations, which have a similar flux limit as the *Swift*/BAT images, show a steep decline in the number of obscured sources, from ∼80% at $z < 0.015$ to ∼20–30% at higher redshifts. Also, all the CT AGN in this sample were found at $z < 0.015$. Hence, these authors concluded that the INTEGRAL observations are affected by obscuration at larger distances and that the intrinsic fraction of CT sources is ∼25%, as observed at $z < 0.015$. However, it is worth mentioning that these additional sources, because of their very high column densities, do not contribute significantly to the XRB, although they certainly contribute to black hole growth.

The cumulative contribution of CT AGN to the XRB, as a function of redshift, determined from population synthesis models, is shown in Figure 12. As can be seen, the total contribution of CT AGN to the XRB is ∼9%, and about 50% of it comes from sources at $z < 0.7$. Similarly, only ∼2% of the XRB is provided by CT AGN at $z > 1.4$, while CT AGN at $z > 2$ contribute only $\lesssim$1% to the XRB. Conversely, the 5% uncertainty in the absolute measurement of the XRB intensity translates into an uncertainty of a factor of ∼5 in the number of CT AGN at $z > 2$. Hence, the number of CT AGN at high redshift is largely unconstrained by the XRB.

4.2. Obscured AGN at Intermediate Redshifts ($1 \lesssim z \lesssim 3$). As shown in the previous section, the number of heavily obscured AGN at intermediate redshifts, $z \gtrsim 1$, is largely unconstrained by the XRB due to model degeneracies, or by current X-ray surveys at $E > 10$ keV, which do not have the required sensitivity. NuSTAR will change this situation dramatically. However, for now we are forced to use alternative methods to determine the amount of black hole growth occurring in these sources. We explore here two of these techniques, which have been particularly successful: X-ray stacking and mid-IR AGN selection.

Using the deepest available X-ray observations obtained with *Chandra*, as in the example shown in Figure 13, a number of moderately obscured AGN have been identified at $z > 1$. The observed-frame hard X-ray band (2–8 keV band) at these redshifts covers higher rest-frame energies, thus making them less affected by obscuration. For example, the vast majority of the sources the GOODS fields with high X-ray to optical flux ratios are obscured at $z \sim 2$ [136–138]. Furthermore, from X-ray spectral analysis, ∼30 CT AGN candidates have been identified in the CDF-S [70] and CDF-N [139]. However, it is clear that X-ray selection remains highly incomplete for obscured sources at these redshifts [91].

Because much of the energy absorbed at optical to X-ray wavelengths is later reemitted in the mid-to far-IR, it is expected that AGN, in particular the most obscured ones, should be very bright mid-IR sources [95]. Sources having mid-IR excesses, relative to their rest-frame optical and UV emission, have been identified as potential CT AGN candidates at $z \sim 2$ [75, 77, 140–142]. However, because of the strong connection between vigorous star formation and AGN activity in the most luminous infrared sources [26], the relative contribution of these two processes remains uncertain and controversial [143–145]. Significant progress has been made by virtue of deep Spitzer observations, in particular using the 24 μm band. At $z \sim$ 1-2, this emission corresponds to rest-frame wavelengths of ∼10 μm, where the contrast between AGN and star formation is largest. In order to look for high-luminosity obscured AGN missed by X-ray observations, Fiore et al. [78, 141] defined the "mid-IR excess" region as $f_{24}/f_R > 1000$ and R-$K > 4.5$ (Vega), as shown

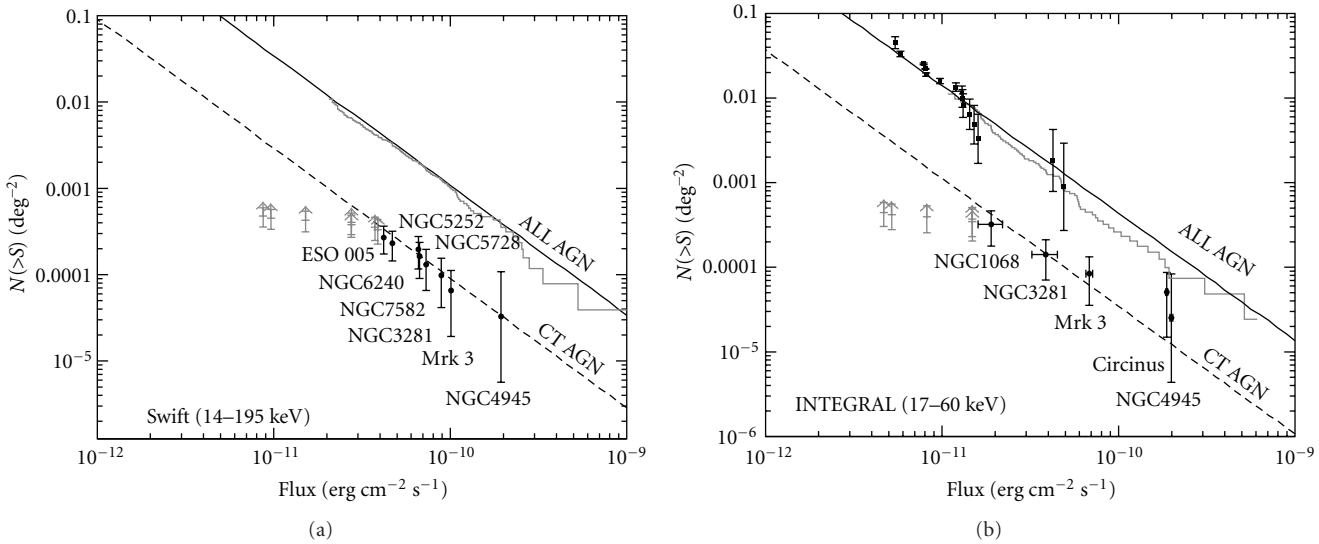


Figure 10: Log$N$-log$S$ distribution for AGN detected at high energies ($E > 10$ keV). The *gray line* in the (a) shows the AGN in the well-defined Swift/BAT samples in the 14–195 keV band [66], while the (b) shows the INTEGRAL sources [67] in the 17-60 keV band. *Solid squares* show the 15 sources detected in the ultradeep 3 Msec INTEGRAL observations of the XMM-LSS field. *Solid circles* mark the CT AGN detected with Swift (a) and INTEGRAL (b); the fraction is ∼5%. The *black solid lines* show the expected AGN Log$N$-log$S$ from the most complete AGN population synthesis model [38], which at these high fluxes has a Euclidean slope. The *dashed lines* mark the Euclidean slope normalized to the number of Swift and INTEGRAL CT AGN (5% of the total). The *gray* lower limits show the previously known transmission-dominated AGN with hard X-ray observations, not detected in the INTEGRAL or Swift surveys. These are lower limits since they were selected from pointed observations and are thus highly incomplete.

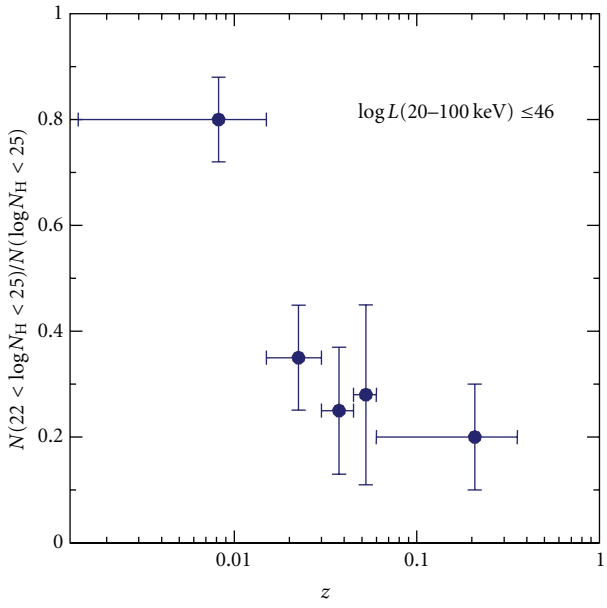


Figure 11: Fraction of obscured AGN ($N_H > 10^{22}$ cm^{-2}) as a function of redshift for a sample of AGN selected at high energies from INTEGRAL observations, as shown by Malizia et al. [68]. All of these sources have X-ray luminosities lower than 10^{46} erg s^{-1}. The remarkably strong decline in this fraction at $z \sim 0.015$ is likely due to selection effects and not intrinsic to the AGN population. That the fraction of CT AGN in the first bin is ∼20%, in contrast to the ∼5% found overall suggests that current hard X-ray surveys are not sensitive enough to observe CT AGN beyond $z \sim 0.01$. This will improve dramatically with NuSTAR [69].

in Figure 14. As argued by several authors, the fraction of CT AGN in these infrared-excess samples is very high, >70% [75, 78, 141].

Because sources in the mid-IR excess region are, by definition, very faint at optical wavelengths, it has been very difficult to use optical spectroscopy to measure accurate redshifts. Instead, most surveys have had to rely on (hopefully) accurate photometric redshifts [146, 147]. The distribution of photometric redshifts for the sources in the mid-IR excess region in the ECDF-S is shown in Figure 15; most mid-IR excess sources have $1 < z < 3$. While the majority of these sources are not detected in X-rays, a significant signal is found in X-ray stacks [75, 78, 141]. As shown in Figure 16, the strong stacked detection at $E > 5$ keV clearly indicates the presence of a large number of heavily obscured AGN in this infrared-excess subsample. Specifically, Treister et al. [75] reported that heavily obscured AGN were ∼80–90% of the mid-IR-excess sources in the ECDF-S. A similarly high fraction of ∼80% was found in the CDF-S [141] and other fields [78]. Optical spectral fitting of these sources indicates evidence for substantial young stellar populations, younger than 100 Myrs [75]. This suggests that these sources are simultaneously experiencing significant star formation and heavily obscured AGN activity. The best-fit stellar masses for ECDF-S infrared-excess sources range between 10^9 and $10^{12} M_\odot$ with a median stellar mass of $\sim 10^{11} M_\odot$ [75]. Hence, in general these are very massive galaxies.

In order to study in more detail the evolution of the CT AGN space density, in Figure 17 we present the existing measurements of the CT AGN space density as a function

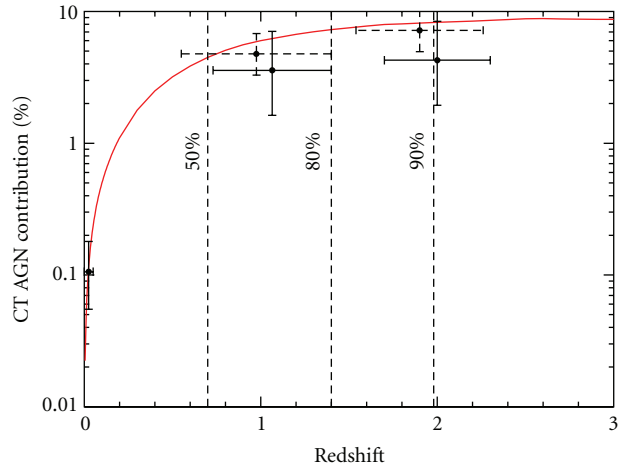


Figure 12: Cumulative fractional contribution of CT AGN to the XRB in the 14–195 keV Swift/BAT band as a function of redshift from population synthesis models [35]. As shown by the vertical dashed lines, 50%, 80%, and 90% of the total CT AGN contribution come from sources at $z < 0.7$, 1.4, and 2.0, respectively. Only ~1% of the total XRB intensity comes from CT AGN at $z > 2$. The data point at $z \sim 0$ corresponds to the contribution to the XRB by the CT AGN detected by Swift/BAT, while the data points at high redshift were obtained from the CT AGN in the CDF-S [70]. Solid error bars correspond to transmission-dominated sources only, while the data points with dashed error bars include all the sources in the sample.

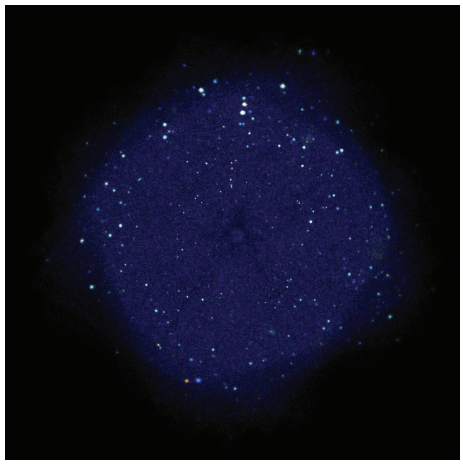


Figure 13: 4 Msec *Chandra* observations of the CDF-S. This is currently the deepest view of the X-ray sky. There are ~760 individual sources in ~450 arcmin^2. This is a smooth enhanced image corresponding to the full (0.5–8 keV) X-ray band. Image and data obtained from http://cxc.harvard.edu/cdo/cdfs.html.

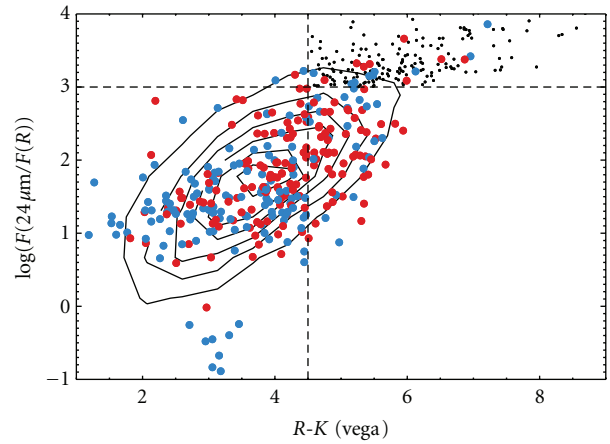


Figure 14: The ratio of 24-micron to R-band optical flux as a function of R-K color for the sources in the E-CDFS field [75]. Contours show the location of all the Spitzer 24 μm sources. *Red points* show the location of X-ray sources with hard X-ray spectra (hardness ratio HR > -0.3, defined as HR $= (H - S)/(H + S)$ where S and H are the background-subtracted counts in the soft and hard bands, resp.), while *blue points* have HR < -0.3. *Small black points* show the 193 sources in the IR-red excess region that were not detected individually in X-rays.

of redshift. Reasonable agreement, in particular at $z < 1$, is found between both observed values and existing hard X-ray luminosity functions and evolution [79]. However, at $z > 1.7$ and high luminosity, a clear discrepancy is found. Treister et al. [35] concluded that this difference of a factor of 2-3 could be due either to incompleteness in the Swift/BAT and INTEGRAL CT AGN samples at $z = 0$ used to fix the luminosity function normalization (because reflection-dominated AGN are missed) or to contamination by other types of sources in the observed values at high redshifts. However, after adding the measurements obtained using the infrared-excess sources in the ECDF-S, it appears not only that the systematic difference is still present but perhaps more importantly that there is a strong increase in the number of CT AGN from $z \simeq 1.7$ to 2.4. This is not described by any existing luminosity function. It is unlikely that this evolution is due to selection effects, as results from different fully independent surveys and selection techniques are combined in Figure 17, namely, X-ray-selected sources [70], 24 μm-selected sources [75, 78], and a sample of CT AGN found using mid-IR spectroscopy [77]. This result can be interpreted in the context of galaxy evolution models [148], where quasar activity is driven by galaxy mergers and the supermassive black hole is initially completely surrounded by dust, before radiation pressure removes it and a "classical" unobscured quasar is visible [25].

Several groups [149] have found that the fraction of galaxies containing an AGN is a strong function of their IR luminosity. In Figure 18 we present the stacked spectra for the sources in the CDF-S, grouped in bins of IR luminosity [150]. We can see by comparing these spectra that the relative emission at $E \gtrsim 5$ keV, where we expect the AGN emission to dominate even for heavily obscured sources, changes with IR luminosity. In other words, there is a clear trend, with stronger high energy X-ray emission at increasing IR luminosity. The spectra shown in Figure 18 cannot be directly interpreted, as the detector-plus-telescope response information is lost after the conversion to rest-frame energy and stacking. Hence, simulations assuming different intrinsic X-ray spectra have to be used in order to constrain the nature of the sources dominating the coadded signal.

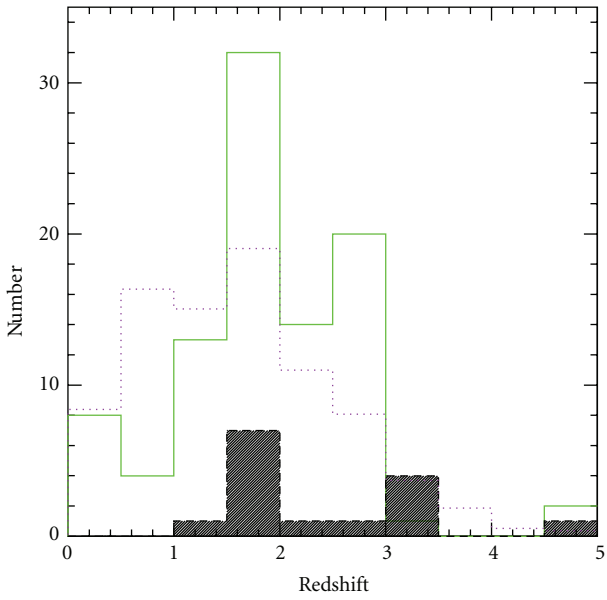


Figure 15: Photometric redshift distribution for 90 sources in the ECDF-S in the mid-IR excess region. The *solid histogram* shows the distribution for the sources not detected in X-rays, while the *dashed hatched histogram* considers only the X-ray-detected sources. A KS test shows that it is perfectly likely (~16%) that these two distributions were drawn from the same parent distribution. The *dotted histogram* shows the slightly lower redshift distribution (divided by 1,000) for all the sources with a 24 μm detection and a measured photometric redshift in the ECDF-S.

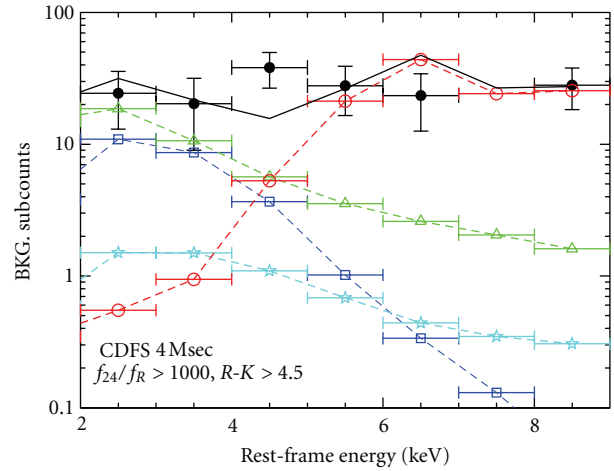


Figure 16: Stacked background-subtracted Chandra counts as a function of rest-frame energy, for sources with $f_{24}/f_R > 1000$ and $R\text{-}K > 4.5$ in the 4 Msec CDF-S field (*filled circles*; [75]). The *cyan dashed lines (stars)* show the simulated spectra for the high-mass X-ray binary (HMXB) population normalized using the relation between star-formation rate and X-ray luminosity [76]. The *blue dashed lines (open squares)* show simulated thermal spectra corresponding to a black body with $kT = 0.7$ keV. An absorbed AGN spectrum, given by a power law with $\Gamma = 1.9$ and a fixed $N_H = 10^{24}$ cm^{-2}, is shown by the *red dashed lines (open circles)*. In addition, a scattered AGN component, characterized by a 1% reflection of the underlying unobscured power law, is shown by the *green dashed lines (open triangles)*. The resulting summed spectrum (*black solid lines*) is in very good agreement with the observed counts. The strong detection in the stacked spectrum at $E > 5$ keV confirms the presence of a significant number of heavily obscured AGN in these IR-excess objects [75].

The observed stacked spectral shape cannot be explained by any plausible starburst spectrum. An AGN component dominating at $E > 5$ keV, is required. The average intrinsic rest-frame 2–10 keV AGN luminosity needed to explain the observed spectrum, assuming that every source in the sample contains an AGN of the same luminosity, is 6×10^{42} erg s^{-1} for sources with $L_{\rm IR} > 10^{11} L_\odot$, 3×10^{42} erg s^{-1} for sources with $L_{IR} > 5 \times 10^{10} L_\odot$, 5×10^{41} erg s^{-1} for $5 \times 10^{10} L_\odot > L_{\rm IR} > 10^{10} L_\odot$, and 7×10^{41} erg s^{-1} for $L_{IR} > 10^{10} L_\odot$. All of these are (intrinsically) very low-luminosity AGN; even if there is a range, it is extremely unlikely to include high-luminosity quasars like those discussed in previous stacking papers. This is not too surprising, actually, because the surveyed volume (even to high redshift) is small, so rare objects like high-luminosity quasars do not appear. If the heavily obscured AGN in these stacked samples have the same median intrinsic luminosity as the X-ray-detected sources with similar IR luminosities, this would indicate that 15% of the galaxies with $L_{\rm IR} > 10^{11} L_\odot$ contain a heavily obscured AGN. This fraction is ~10% in the $L_{\rm IR} > 5 \times 10^{10} L_\odot$ and $5 \times 10^{10} L_\odot > L_{IR} > 10^{10} L_\odot$ samples. For sources with $L_{\rm IR} > 10^{10} L_\odot$ this fraction is <5%. This extra AGN activity (in addition to the X-ray detected sources) can account for ~22% of the total black hole accretion. Adding this to the obscured black hole growth in X-ray-detected AGN [151], we confirm that most of this growth, ~70%, is significantly obscured and missed by even the deepest X-ray surveys [25, 91].

4.3. Obscured AGN at High Redshifts, $z > 3$. As mentioned in Section 3, most measurements of black hole accretion at high redshift, $z > 3$, come from optical observations of unobscured sources. This is not only because obscured sources are obviously fainter at most wavelengths, but also because large areas have to be covered in order to survey a significant volume at high redshifts. As can be seen in Figure 19, there is a clear decline in the number of luminous quasars at $z > 2$, although the decline is shallower for X-rays compared to optical surveys, since X-ray selection is less biased [79, 85]. However, it is important to point out that: (i) these results are limited to the highest luminosity sources, $\log l_x > 44.5$ erg s^{-1}, which do not represent the average AGN and do not contribute much to the extragalactic XRB [35] and (ii) only relatively unobscured sources are included. In particular, heavily obscured, Compton-thick, AGN are systematically underrepresented in these surveys. As we will describe, these missing populations can have a significant impact in our understanding of cosmic supermassive black hole growth.

In order to search for the presence of growing supermassive black holes in young galaxies, Treister et al. [152] stacked X-ray images of $z > 6$ galaxy candidates selected based on the optical and near-IR dropout techniques, selected from a

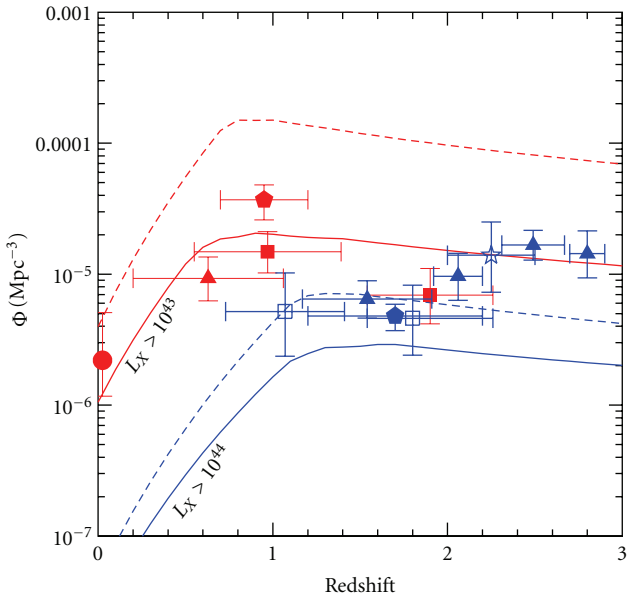


FIGURE 17: Space density of Compton thick AGN as a function of redshift, as published by Treister et al. [75]. *Filled triangles* show the ECDF-S results [75]. *Squares*: X-ray-selected sources in the CDF-S [70]. *Star*: Measurement obtained using mid-IR spectroscopy [77]. *Pentagons*: Values obtained using mid-IR excess sources in COSMOS [78]. *Solid lines* show the expected space density of Compton thick AGN from the luminosity function of Yencho et al. [79], with the overall normalization fixed to the results of the INTEGRAL and Swift/BAT surveys [35], while the *dashed lines* show the expectations based on the luminosity function of Della Ceca et al. [80]. *Red symbols* show measurements and expectations for $L_X > 10^{43}$ erg s^{-1} sources, while the *blue symbols* are for $L_X > 10^{44}$ erg s^{-1}. While for the lower luminosity sources a good agreement is found between observations and expectations, higher luminosity sources at $z > 1.8$ lie well above the luminosity function.

sample of 197 galaxies, 151 in the CDF-S and 46 in the CDF-N from Bouwens et al. [153]. Using the 4 Msec Chandra observations of the CDF-S, and the 2 Msec data available on the CDF-N, this corresponds to a total exposure time of $\sim 7 \times 10^8$ seconds ($\sim$23 years). Significant detections in both the soft and hard X-ray bands were obtained. However, these detections have recently been questioned (after the submission of this paper) by several authors [154–156] due to a possible bias in the background subtraction technique used by Treister et al. While a full discussion is clearly beyond the scope of this paper, we note that a full analysis, using an optimal weighting scheme (as in [152]), as well as considering the effects of faint, undetected, sources in the background, remains to be done. At a minimum, the results presented below can be considered as upper limits.

The corresponding average rest-frame 2–10 keV luminosity, derived from the observed-frame hard band, is 6.8×10^{42} erg s^{-1}. Since none of these sources were individually detected in X-rays, at least 30% of the galaxies in this sample likely contain an AGN [152]. Furthermore, there is a factor $\sim$9 difference between the fluxes measured in the observed-frame soft and hard bands. The only explanation for this relatively large flux ratio in the hard-to-soft bands is very high levels of obscuration. As can be seen in Figure 20, at $z \sim 6$, a minimum column density of $N_H \simeq 10^{24}$ cm^{-2}, that is, Compton thick obscuration, is required. Given that this ratio is observed in a stacked X-ray spectrum, this implies that there are very few sources with significantly lower levels of obscuration, which in turn means that these sources must be nearly Compton-thick along most directions ($\sim 4\pi$ obscuration). Similar sources have also been observed in the local universe [157] but appear to be rare. Furthermore, for $z \lesssim 3$ we know that the fraction of obscured AGN increases with decreasing luminosity [92, 104, 117] and increasing redshift [31, 111]. Hence, it is not entirely surprising that the sources studied here, given their low luminosities and high redshifts, are heavily obscured. In fact, the discovery of a Compton-thick AGN at $z \sim 5$ selected using the dropout technique has been recently reported [158].

This relatively large number of growing supermassive black holes at $z \gtrsim 6$ is contrary to the picture obtained from optical observations of high-luminosity quasars at similar distances, which are much rarer [21, 88]. This is particularly important in our understanding of the early hydrogen reionization, which can be either due to young stars and/or growing supermassive black holes [159]. The results presented by Treister et al. [152] show that while growing supermassive black holes do not contribute much to hydrogen reionization, this is not because their numbers drop steeply at $z > 4$ as previously suggested [160], but because large amounts of obscuration found in these sources imply that UV and soft X-rays do not escape.

5. The Cosmic History of Black Hole Accretion

Direct black hole mass measurements, either through stellar or gas dynamics, are available for only a few nearby galaxies. However, thanks to the tight correlation between mass of the supermassive black hole and other properties such as velocity dispersion and others, it has been possible to estimate the black hole mass function at $z \simeq 0$ [161–164]. This is commonly done starting from the observed galaxy luminosity or velocity function and assuming either a constant black hole to stellar mass ratio [161] or the M-σ relation [163]. Both the overall shape of the black hole mass function and the integrated black hole mass density, which can only be computed at $z \simeq 0$, can be used to infer properties of the AGN population. This was first used in the so-called "Soltan's argument" [24], which says that the intrinsic bolometric AGN luminosity, L, is directly linked to the amount of mass accreted by the black hole, $\dot{M}_{\rm acc}$:

$$L = \varepsilon \dot{M}_{\rm acc} c^2, \tag{1}$$

where ε is the accretion efficiency and c is the speed of light. A typical value assumed for the efficiency is $\sim$10% [24, 163].

Recent comparisons of the black hole mass function to the distribution inferred from the observed AGN luminosity indicate that the average efficiency is 8%, the Eddington ratio is $\sim$50%, and the average lifetime of the visible AGN phase is $\sim 10^8$ years [163, 164]. By studying the black hole mass distribution at the high mass end, $M > 10^9 M_\odot$, Natarajan

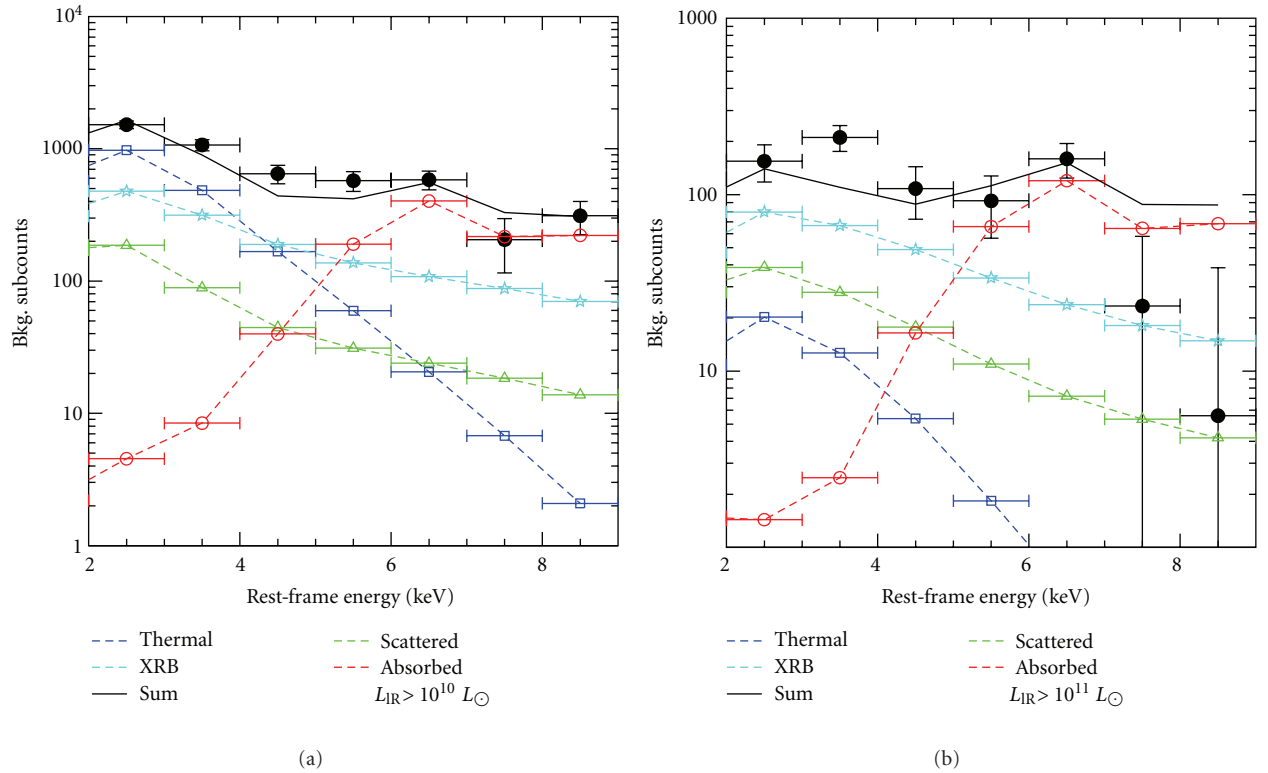


FIGURE 18: Stacked background-subtracted Chandra counts as a function of rest-frame energy for sources with low (*a*) and high (*b*) IR luminosity (*black filled circles*). The *cyan dashed lines* (*stars*), *blue dashed lines* (*open squares*), *red dashed lines* (*open circles*) and *green dashed lines* (*open triangles*) show the same model components as in Figure 16. The resulting summed spectra (*black solid lines*) are in very good agreement with the observed counts. The strong detection in the stacked spectrum at $E > 5$ keV, in particular at the higher IR luminosities, confirms the presence of a significant number of heavily obscured AGN in these samples.

and Treister [165] found that the observed number of ultra-massive black holes is significantly lower than the number density inferred from the AGN hard X-ray luminosity function. They concluded that this is evidence for an upper limit to the black hole mass, which can be explained by the presence of a self-regulation mechanism.

The observed black hole mass density at $z \simeq 0$, obtained by integrating the black hole mass function, ranges from 2.9×10^5 [162] to $4.6\ ^{+1.9}_{-1.4} \times 10^5 M_\odot$ Mpc^{-3} [163]; most recently, Shankar et al. [166] found 3.2–$5.4 \times 10^5 M_\odot$ Mpc^{-3}. For comparison, integrating the AGN hard X-ray LF, including the number of Compton-thick AGN constrained by INTEGRAL and Swift/BAT observations, Treister et al. [35] obtained a value of $4.5 \times 10^5 M_\odot$ Mpc^{-3}, perfectly consistent with the observed value, indicating that at least locally X-ray-detected AGN can account for most or all of the black hole growth.

The black hole mass function can be measured observationally for unobscured, high-luminosity AGN at higher redshift, taking advantage of the known correlation between black hole mass and observational quantities such as luminosity and emission line width [167]. These correlations are calibrated using more direct black hole measurements, available for a few, mostly local, sources [8, 9]. The large number of unobscured quasars with optical spectroscopy provided by the SDSS and other optical surveys has been very useful in determining the black hole mass function up to high redshifts, [168–170], and for lower luminosity sources using deeper surveys such as the AGN and Galaxy Evolution Survey (AGES; [171]) and COSMOS [172]. Using these black hole mass functions as a function of redshift as constraints, recently Natarajan [173] concluded that the observational data are inconsistent with the hypothesis that these black holes are created as the remnants of population III stars. Instead, they argue that massive seeding models are required [174].

While a clear picture of the history of black hole growth is emerging, significant uncertainties still remain. In particular, while the spectral shape and intensity of the extragalactic X-ray background have been used to constrain the AGN population, the number of heavily obscured accreting supermassive black holes beyond $z \sim 1$ is not properly bounded. Infrared and deep X-ray selection methods have been useful in that sense, but have not provided a final answer, due to confusion with star-forming galaxies in the infrared and the effects of obscuration in X-rays. At higher redshifts, the situation is even more unclear, and only a few, very rare, high luminosity quasars are known. Unless high-redshift AGN luminosity functions are pathological, these extreme sources do not represent the typical growing black holes in the early universe. As a consequence, and in spite of recent advances

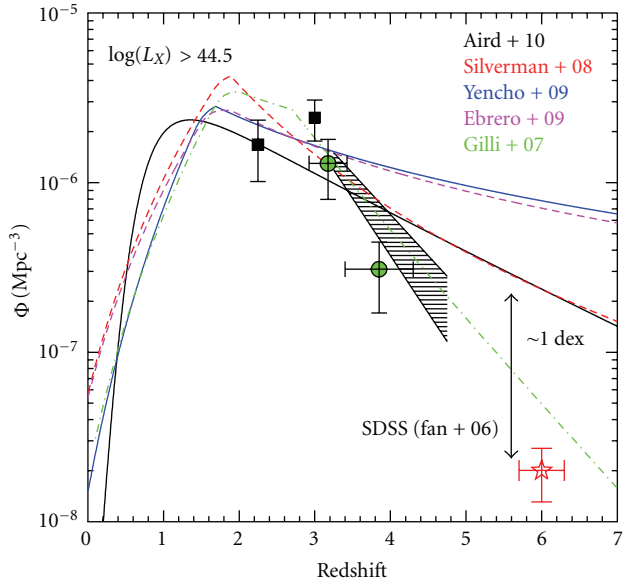


FIGURE 19: Number density of X-ray-selected AGN as a function of redshift for sources with $\log l_{2-10} > 44.5\,\mathrm{erg\,s^{-1}}$, as published by Brusa et al. [83]. *Lines* show AGN luminosity functions [39, 79, 84–86] while individual measurements are shown by green circles [87] and black squares [86]. The *red star* at $z = 6$ was obtained from the optical luminosity function assuming no evolution in α_{ox}. The *black-shaded area* shows the observations of optically bright SDSS QSOs [88].

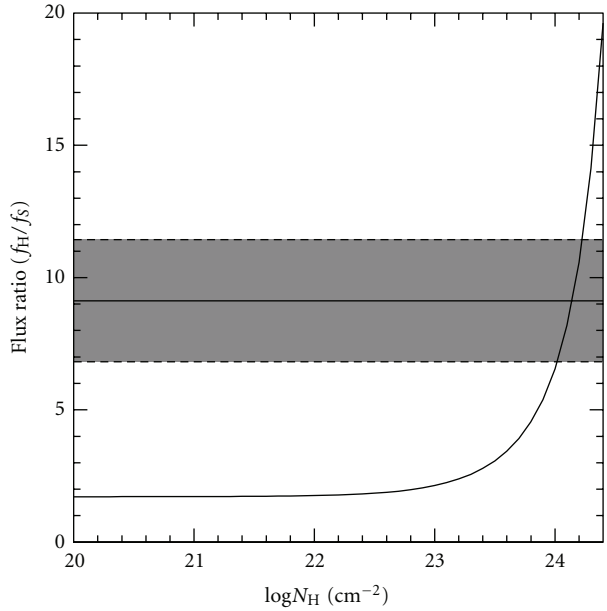


FIGURE 20: Expected ratio of the observed-frame hard-to-soft flux as a function of obscuring neutral Hydrogen column density (N_H). The *black solid line* was derived assuming an intrinsic power law spectrum with slope $\Gamma = 1.9$ and photoelectric absorption. The gray zone shows the measured ratio for the stack of galaxies at $z \simeq 6$ and the ± 1 standard deviation limits. A column density of $N_H \simeq 10^{24}\,\mathrm{cm^{-2}}$, that is, Compton-thick obscuration, is required to explain the observed hard-to-soft X-ray flux ratio [152].

[152, 173], the formation mechanism for the first black holes in the universe is still unknown.

6. Prospects

Scheduled for launch in February 2012, NuSTAR will be the first focusing high-energy ($E = 5$–$80\,\mathrm{keV}$) X-ray mission, reaching flux limits ~100 times fainter than *INTEGRAL* or *Swift*/BAT observations and comparable to *Chandra* and *XMM-Newton* at lower energies. During the first two years of operations, NuSTAR will likely observe, as part of the guaranteed time program, two extragalactic fields: the ECDF-S and the central $1\,\mathrm{deg^2}$ part of COSMOS, for a total of 3.1 Msec each. These deep high-energy observations will enable to obtain a nearly complete AGN survey, including heavily obscured Compton-thick sources, up to $z \sim 1.5$ [175]. A similar mission, ASTRO-H [176], will be launched by Japan in 2014. Both missions will provide angular resolutions. $\lesssim 1'$, in combination with observations at longer wavelengths will allow for the detection and identification of most growing supermassive black holes at $z \sim 1$.

There is little doubt that the Atacama Large Millimeter Array (ALMA) will revolutionize our understanding of galaxy evolution. Sources of mm and submm emission traced by ALMA include thermal emission of the warm/cold dust, which traces star formation, synchrotron radiation associated with relativistic particles, and free-free radiation from HII regions. In particular, CO rotational transition lines have been used to trace the spatial distribution, kinematics, temperature, and mass of the molecular gas [177]. The sensitivity of ALMA will allow for the detection of luminous IR galaxies ($L_{IR} > 10^{11} L_\odot$), which have been found to often host a heavily obscured AGN [150], up to $z \sim 10$. Furthermore, with ALMA it will be possible to study separately the molecular dust surrounding the central black hole and those in star-forming regions in the host galaxy. Due to their limited sensitivity and relatively bad angular resolution, currently available mm/sub-mm telescopes are not ideal to study star-forming regions even in nearby galaxies. This will dramatically change thanks to ALMA, which will have orders of magnitude better sensitivity and HST-like angular resolution. The first call for ALMA observations was released on March 31, 2011 for observations starting on September 30, 2011. It is expected that the complete array will be in full operation in 2013. The superb spatial resolution and sensitivity of ALMA will allow uniquely identify of the optical/near-IR counterpart of the mm-submm sources. Furthermore, ALMA will directly provide the redshift of the mm-submm sources through the detection of CO rotational transition lines, up to very high redshifts. Combining these new data with existing multiwavelength information will finally allow us to complete the census of supermassive black hole growth since the epoch of cosmic reionization.

Acknowledgments

Support for the work of E. Treiser was provided by the National Aeronautics and Space Administration through

Chandra/Einstein Postdoctoral Fellowship Award number PF8-90055, issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory (SAO) for and on behalf of the National Aeronautics Space Administration under Contract no. NAS8-03060. Additional support for this work was provided by NASA through Chandra Award SP1-12005X. E. Treister received partial support from Center of Excellence in Astrophysics and Associated Technologies (PFB 06). C. M. Urry acknowledges support from NSF Grants AST-0407295, AST-0449678, AST-0807570, and Yale University.

References

[1] J. A. Orosz, "Inventory of black hole binaries," in *A Massive Star Odyssey: From Main Sequence to Supernova*, K. van der Hucht, A. Herrero, and C. Esteban, Eds., vol. 212 of *IAU Symposium*, p. 365, 2003.

[2] D. Lynden-Bell, "Galactic nuclei as collapsed old quasars," *Nature*, vol. 223, no. 5207, pp. 690–694, 1969.

[3] T. R. Lauer, S. M. Faber, D. Richstone et al., "The masses of nuclear black holes in luminous elliptical galaxies and implications for the space density of the most massive black holes," *Astrophysical Journal*, vol. 662, no. 2, pp. 808–834, 2007.

[4] R. Genzel, A. Eckart, T. Ott, and F. Eisenhauer, "On the nature of the dark mass in the centre of the Milky Way," *Monthly Notices of the Royal Astronomical Society*, vol. 291, no. 1, pp. 219–234, 1997.

[5] A. M. Ghez, B. L. Klein, M. Morris, and E. E. Becklin, "High proper-motion stars in the vicinity of sagittarius A*: evidence for a supermassive black hole at the center of our galaxy," *Astrophysical Journal*, vol. 509, no. 2, pp. 678–686, 1998.

[6] J. Kormendy and D. Richstone, "Inward bound—the search for supermassive black holes in galactic nuclei," *Annual Review of Astronomy and Astrophysics*, vol. 33, no. 1, pp. 581–624, 1995.

[7] J. N. Bahcall, B.-Z. Kozlovsky, and E. E. Salpeter, "On the time dependence of emission-line strengths from a photoionized nebula," *Astrophysical Journal*, vol. 171, p. 467, 1972.

[8] R. D. Blandford and C. F. McKee, "Reverberation mapping of the emission line regions of Seyfert galaxies and quasars," *Astrophysical Journal*, vol. 255, pp. 419–439, 1982.

[9] B. M. Peterson, "Reverberation mapping of active galactic nuclei," *Astronomical Society of the Pacific*, vol. 105, pp. 247–268, 1993.

[10] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *Astrophysical Journal*, vol. 589, no. 1, pp. L21–L24, 2003.

[11] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[12] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.

[13] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.

[14] L. Ferrarese, "Beyond the bulge: a fundamental relation between supermassive black holes and dark matter halos," *Astrophysical Journal*, vol. 578, no. 1, pp. 90–97, 2002.

[15] V. Springel, T. Di Matteo, and L. Hernquist, "Modelling feedback from stars and black holes in galaxy mergers," *Monthly Notices of the Royal Astronomical Society*, vol. 361, no. 3, pp. 776–794, 2005.

[16] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes, and galaxy spheroids," *Astrophysical Journal Supplement Series*, vol. 163, no. 1, pp. 1–49, 2006.

[17] N. Menci, A. Fontana, E. Giallongo, A. Grazian, and S. Salimbeni, "The abundance of distant and extremely red galaxies: the role of AGN feedback in hierarchical models," *Astrophysical Journal*, vol. 647, no. 2, pp. 753–762, 2006.

[18] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[19] M. J. Rees, "Structure and properties of nearby galaxies," in *Proceedings of the Symposium*, vol. 1, pp. 237–244, Bad Muenstereifel, Germany, 1978.

[20] M. Volonteri, "Formation of supermassive black holes," *Astronomy and Astrophysics Review*, vol. 18, no. 3, pp. 279–315, 2010.

[21] C. J. Willott, L. Albert, D. Arzoumanian et al., "Eddington-limited accretion and the black hole mass function at redshift 6," *Astronomical Journal*, vol. 140, no. 2, pp. 546–560, 2010.

[22] V. Bromm and A. Loeb, "Formation of the first supermassive black holes," *Astrophysical Journal*, vol. 596, no. 1, pp. 34–46, 2003.

[23] B. Devecchi and M. Volonteri, "Formation of the first nuclear clusters and massive black holes at high redshift," *Astrophysical Journal*, vol. 694, no. 1, pp. 302–313, 2009.

[24] A. Soltan, "Masses of quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 200, pp. 115–122, 1982.

[25] E. Treister, P. Natarajan, D. B. Sanders, C. Megan Urry, K. Schawinski, and J. Kartaltepe, "Major galaxy mergers and the growth of supermassive black holes in quasars," *Science*, vol. 328, no. 5978, pp. 600–602, 2010.

[26] D. B. Sanders, B. T. Soifer, J. H. Elias et al., "Ultraluminous infrared galaxies and the origin of quasars," *Astrophysical Journal*, vol. 325, pp. 74–91, 1988.

[27] P. F. Hopkins and L. Hernquist, "Fueling low-level AGN activity through stochastic accretion of cold gas," *Astrophysical Journal*, vol. 166, no. 1, pp. 1–36, 2006.

[28] R. Antonucci, "Unified models for active galactic nuclei and quasars," *Annual Review of Astronomy and Astrophysics*, vol. 31, no. 1, pp. 473–521, 1993.

[29] C. M. Urry and P. Padovani, "Unified schemes for radio-loud active galactic nuclei," *Astronomical Society of the Pacific*, vol. 107, p. 803, 1995.

[30] A. Lawrence, "The relative frequency of broad-lined and narrow-lined active galactic nuclei—implications for unified schemes," *Monthly Notices of the Royal Astronomical Society*, vol. 252, pp. 586–592, 1991.

[31] E. Treister and C. M. Urry, "The evolution of obscuration in active galactic nuclei," *Astrophysical Journal*, vol. 652, no. 2, pp. L79–L82, 2006.

[32] R. C. Hickox and M. Markevitch, "Absolute measurement of the unresolved cosmic X-ray background in the 0.5-8 keV band with Chandra," *Astrophysical Journal*, vol. 645, no. 1, pp. 95–114, 2006.

[33] R. F. Mushotzky, L. L. Cowie, A. J. Barger, and K. A. Arnaud, "Resolving the extragalactic hard X-ray background," *Nature*, vol. 404, no. 6777, pp. 459–464, 2000.

[34] G. Setti and L. Woltjer, "Active galactic nuclei and the spectrum of the X-ray background," *Astronomy and Astrophysics*, vol. 224, pp. L21–L23, 1989.

[35] E. Treister, C. M. Urry, and S. Virani, "The space density of compton thick AGN and the X-ray background," *Astrophysical Journal*, vol. 696, pp. 110–120, 2009.

[36] A. R. Draper and D. R. Ballantyne, "Balancing the cosmic energy budget: the cosmic X-ray background, blazars, and the compton thick active galactic nucleus fraction," *Astrophysical Journal*, vol. 707, no. 1, pp. 778–786, 2009.

[37] D. Burlon, M. Ajello, J. Greiner, A. Comastri, A. Merloni, and N. Gehrels, "Three-year swift-BAT survey of active galactic nuclei: reconciling theory and observations?" *Astrophysical Journal*, vol. 728, p. 58, 2011.

[38] E. Treister and C. Megan Urry, "Active galactic nuclei unification and the X-ray background," *Astrophysical Journal*, vol. 630, no. 1, pp. 115–121, 2005.

[39] R. Gilli, A. Comastri, and G. Hasinger, "The synthesis of the cosmic X-ray background in the Chandra and XMM-Newton era," *Astronomy and Astrophysics*, vol. 463, no. 1, pp. 79–96, 2007.

[40] G. Hinshaw, J. L. Weiland, R. S. Hill et al., "Five-year wilkinson microwave anisotropy probe observations: data processing, sky maps and basic results," *Astrophysical Journal Supplement Series*, vol. 180, no. 2, pp. 225–245, 2009.

[41] J. C. Manners, *Obscuration and X-ray variability of active galactic nuclei*, Ph.D. thesis, University of Edinburgh, 2002.

[42] J. J. Condon, Q. F. Yin, T. X. Thuan, and T. Boller, "The ROSAT/IRAS galaxy sample revisited," *Astronomical Journal*, vol. 116, no. 6, pp. 2682–2716, 1998.

[43] G. A. Shields, "Thermal continuum from accretion disks in quasars," *Nature*, vol. 272, no. 5655, pp. 706–708, 1978.

[44] M. A. Malkan and W. L. W. Sargent, "The ultraviolet excess of Seyfert 1 galaxies and quasars," *Astrophysical Journal*, vol. 254, pp. 22–37, 1982.

[45] M. Schmidt and R. F. Green, "Quasar evolution derived from the Palomar bright quasar survey and other complete quasar surveys," *Astrophysical Journal*, vol. 269, pp. 352–374, 1983.

[46] B. J. Boyle, T. Shanks, S. M. Croom et al., "The 2dF QSO Redshift Survey-I. The optical luminosity function of quasi-stellar objects," *Monthly Notices of the Royal Astronomical Society*, vol. 317, no. 4, pp. 1014–1022, 2000.

[47] G. T. Richards, X. Fan, D. P. Schneider et al., "Colors of 2625 quasars at $0 < z < 5$ measured in the sloan digital sky survey photometric system," *Astronomical Journal*, vol. 121, no. 5, pp. 2308–2330, 2001.

[48] G. T. Richards, A. D. Myers, A. G. Gray et al., "Efficient photometric selection of quasars from the sloan digital sky survey. II. ~1,000,000 quasars from data release 6," *Astrophysical Journal Supplement Series*, vol. 180, no. 1, pp. 67–83, 2009.

[49] D. E. Vanden Berk, G. T. Richards, A. Bauer et al., "Composite quasar spectra from the sloan digital sky survey," *Astronomical Journal*, vol. 122, no. 2, pp. 549–564, 2001.

[50] L. J. Kewley, B. Groves, G. Kauffmann, and T. Heckman, "The host galaxies and classification of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 3, pp. 961–976, 2006.

[51] G. Kauffmann, T. M. Heckman, C. Tremonti et al., "The host galaxies of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 346, no. 4, pp. 1055–1077, 2003.

[52] L. J. Kewley, M. A. Dopita, R. S. Sutherland, C. A. Heisler, and J. Trevena, "Theoretical modeling of starburst galaxies," *Astrophysical Journal*, vol. 556, no. 1, pp. 121–140, 2001.

[53] S. M. Croom, G. T. Richards, T. Shanks et al., "The 2dF-SDSS LRG and QSO survey: the QSO luminosity function at $0.4 < z < 2.6$," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 4, pp. 1755–1772, 2009.

[54] S. M. Croom, R. J. Smith, B. J. Boyle et al., "The 2dF QSO Redshift Survey-XII. The spectroscopic catalogue and luminosity function," *Monthly Notices of the Royal Astronomical Society*, vol. 349, no. 4, pp. 1397–1418, 2004.

[55] G. T. Richards, S. M. Croom, S. F. Anderson et al., "The 2dF-SDSS LRG and QSO (2SLAQ) Survey: the $z < 2.1$ quasar luminosity function from 5645 quasars to $g = 21.85$," *Monthly Notices of the Royal Astronomical Society*, vol. 360, no. 3, pp. 839–852, 2005.

[56] G. T. Richards, M. A. Strauss, X. Fan et al., "The sloan digital sky survey quasar survey: Quasar luminosity function from data release 3," *Astronomical Journal*, vol. 131, no. 6, pp. 2766–2787, 2006.

[57] E. Glikman, S. G. Djorgovski, D. Stern, A. Dey, B. T. Jannuzi, and K.-S. Lee, "The faint end of the Quasar luminosity function at $z \sim 4$: implications for ionization of the intergalactic medium and cosmic downsizing," *Astrophysical Journal*, vol. 728, p. L26, 2011.

[58] E. Glikman, M. Bogosavljević, S. G. Djorgovski et al., "The faint end of the quasar luminosity function at $z \sim 4$," *Astrophysical Journal*, vol. 710, no. 2, pp. 1498–1514, 2010.

[59] H. Ikeda, T. Nagao, K. Matsuoka et al., "Probing the faint end of the Quasar luminosity function at $z \sim 4$ in the COSMOS field," *Astrophysical Journal*, vol. 728, p. L25, 2011.

[60] B. Siana, M. Del Carmen Polletta, H. E. Smith et al., "High-redshift QSOs in the swire surveyand the $z \sim 3$ QSO luminosity function," *Astrophysical Journal*, vol. 675, no. 1, pp. 49–70, 2008.

[61] D. E. Gruber, J. L. Matteson, L. E. Peterson, and G. V. Jung, "The spectrum of diffuse cosmic hard X-rays measured with HEAO 1," *Astrophysical Journal*, vol. 520, no. 1, pp. 124–129, 1999.

[62] A. De Luca and S. Molendi, "The 2-8 keV cosmic X-ray background spectrum as observed with XMM-Newton," *Astronomy and Astrophysics*, vol. 419, no. 3, pp. 837–848, 2004.

[63] E. Churazov, R. Sunyaev, M. Revnivtsev et al., "INTEGRAL observations of the cosmic X-ray background in the 5-100 keV range via occultation by the Earth," *Astronomy and Astrophysics*, vol. 467, no. 2, pp. 529–540, 2007.

[64] M. Ajello, J. Greiner, G. Sato et al., "Cosmic X-Ray background and Earth albedo spectra with swift bat," *Astrophysical Journal*, vol. 689, no. 2, pp. 666–677, 2008.

[65] S. F. Hönig and T. Beckert, "Active galactic nuclei dust tori at low and high luminosities," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 3, pp. 1172–1176, 2007.

[66] J. Tueller, R. F. Mushotzky, S. Barthelmy et al., "Swift bat survey of AGNs," *Astrophysical Journal*, vol. 681, no. 1, pp. 113–127, 2008.

[67] R. Krivonos, M. Revnivtsev, A. Lutovinov, S. Sazonov, E. Churazov, and R. Sunyaev, "INTEGRAL/IBIS all-sky survey in hard X-rays," *Astronomy and Astrophysics*, vol. 475, no. 2, pp. 775–784, 2007.

[68] A. Malizia, J. B. Stephen, L. Bassani, A. J. Bird, F. Panessa, and P. Ubertini, "The fraction of Comptonthick sources in an INTEGRAL complete AGN sample," *Monthly Notices of the Royal Astronomical Society*, vol. 399, pp. 944–951, 2009.

[69] F. A. Harrison et al., "The nuclear spectroscopic telescope array (NuSTAR)," in *Proceedings of the Space Telescopes and Instrumentation: Ultraviolet to Gamma Ray*, vol. 7732 of *Proceedings of SPIE*, July 2010.

[70] P. Tozzi, R. Gilli, V. Mainieri et al., "X-ray spectral properties of active galactic nuclei in the Chandra Deep Field South," *Astronomy and Astrophysics*, vol. 451, no. 2, pp. 457–474, 2006.

[71] M. Schmidt, "3C 273: a star-like object with large red-shift," *Nature*, vol. 197, no. 4872, p. 1040, 1963.

[72] K. I. Kellermann, R. Sramek, M. Schmidt, D. B. Shaffer, and R. Green, "VLA observations of objects in the Palomar Bright Quasar Survey," *Astronomical Journal*, vol. 98, pp. 1195–1207, 1989.

[73] M. C. Begelman, R. D. Blandford, and M. J. Rees, "Theory of extragalactic radio sources," *Reviews of Modern Physics*, vol. 56, no. 2, pp. 255–351, 1984.

[74] D. B. Sanders, E. S. Phinney, G. Neugebauer, B. T. Soifer, and K. Matthews, "Continuum energy distributionof quasars—shapes and origins," *Astrophysical Journal*, vol. 347, pp. 29–51, 1989.

[75] E. Treister, C. N. Cardamone, K. Schawinski et al., "Heavily obscured agn in star-forming galaxies at $z \simeq 2$," *Astrophysical Journal*, vol. 706, no. 1, pp. 535–552, 2009.

[76] P. Ranalli, A. Comastri, and G. Setti, "The 2-10 keV luminosity as a star formation rate indicator," *Astronomy and Astrophysics*, vol. 399, no. 1, pp. 39–50, 2003.

[77] D. M. Alexander, R. R. Chary, A. Pope et al., "Reliable identification of compton-thick quasars at $z \approx 2$: spitzer mid-infrared spectroscopy of HDF-oMD49," *Astrophysical Journal*, vol. 687, no. 2, pp. 835–847, 2008.

[78] F. Fiore et al., "Chasing highly obscured QSOs in the COSMOS field," *The Astrophysical Journal*, vol. 693, pp. 447–462, 2009.

[79] B. Yencho, A. J. Barger, L. Trouille, and L. M. Winter, "The optx project. II. hard X-ray luminosity functions of active galactic nuclei for $z \lesssim 5$," *Astrophysical Journal*, vol. 698, no. 1, pp. 380–396, 2009.

[80] R. Della Ceca, A. Caccianiga, P. Severgnini et al., "The cosmological properties of AGN in the XMM-Newton Hard Bright Survey," *Astronomy and Astrophysics*, vol. 487, no. 1, pp. 119–130, 2008.

[81] J. A. Baldwin, M. M. Phillips, and R. Terlevich, "Classification parameters for the emission-line spectra of extragalactic objects," *Astronomical Society of the Pacific*, vol. 93, pp. 5–19, 1981.

[82] A. D. Montero-Dorta, D. J. Croton, R. Yan et al., "The DEEP2 Galaxy Redshift Survey: the red sequence AGN fraction and its environment and redshift dependence," *Monthly Notices of the Royal Astronomical Society*, vol. 392, no. 1, pp. 125–134, 2009.

[83] M. Brusa, R. Gilli, F. Civano, A. Comastri, R. Fiore, and C. Vignali, "Identification of (high-redshift) AGN with WFXT: lessons from COSMOS and CDFS," *Memorie della Societa Astronomica Italiana Supplementi*, vol. 17, p. 106, 2011.

[84] J. D. Silverman, P. J. Green, W. A. Barkhouse et al., "The luminosity function of X-ray-selected active galactic nuclei: evolution of supermassive black holes at high redshift," *Astrophysical Journal*, vol. 679, no. 1, pp. 118–139, 2008.

[85] J. Ebrero, F. J. Carrera, M. J. Page et al., "The XMM-Newton serendipitous survey VI. The X-ray luminosity function," *Astronomy and Astrophysics*, vol. 493, no. 1, pp. 55–69, 2009.

[86] J. Aird, K. Nandra, E. S. Laird et al., "The evolution of the hard X-ray luminosity function of AGN," *Monthly Notices of the Royal Astronomical Society*, vol. 401, pp. 2531–2551, 2010.

[87] M. Brusa, A. Comastri, R. Gilli et al., "High-redshift quasars in the COSMOS survey: the space density of $z > 3$ X-ray selected QSOs," *Astrophysical Journal*, vol. 693, pp. 8–22, 2009.

[88] X. Fan, C. L. Carilli, and B. Keating, "Observational constraints on cosmic reionization," *Annual Review of Astronomy and Astrophysics*, vol. 44, pp. 415–462, 2006.

[89] M. Elvis, T. Maccacaro, A. S. Wilson et al., "Seyfert galaxies as X-ray sources," *Monthly Notices of the Royal Astronomical Society*, vol. 183, pp. 129–157, 1978.

[90] F. E. Bauer, D. M. Alexander, W. N. Brandt et al., "The fall of active galactic nuclei and the rise of star-forming galaxies: a close look at the Chandra Deep Field X-ray number counts," *Astronomical Journal*, vol. 128, no. 5, pp. 2048–2065, 2004.

[91] E. Treister, C. Megan Urry, E. Chatzichristou et al., "Obscured active galactic nuclei and the X-RAY, optical, and far-infrared number counts of active galactic nuclei in the goods fields," *Astrophysical Journal*, vol. 616, no. 1, pp. 123–135, 2004.

[92] S. Sazonov, M. Revnivtsev, R. Krivonos, E. Churazov, and R. Sunyaev, "Hard X-ray luminosity function and absorption distribution of nearby AGN: INTEGRAL all-sky survey," *Astronomy and Astrophysics*, vol. 462, no. 1, pp. 57–66, 2007.

[93] G. T. Richards, M. Lacy, L. J. Storrie-Lombardi et al., "Spectral energy distributions and multiwavelength selection of type 1 quasars," *Astrophysical Journal Supplement Series*, vol. 166, no. 2, pp. 470–497, 2006.

[94] M. Haas, U. Klaas, S. A. H. Müller et al., "The ISO view of Palomar-Green quasars," *Astronomy and Astrophysics*, vol. 402, no. 1, pp. 87–111, 2003.

[95] E. Treister, C. M. Urry, J. Van Duyne et al., "Spitzer number counts of active galactic nuclei in the goods fields," *Astrophysical Journal*, vol. 640, no. 2, pp. 603–611, 2006.

[96] E. Glikman, D. J. Helfand, and R. L. White, "A near-infrared spectral template for quasars," *Astrophysical Journal*, vol. 640, no. 2, pp. 579–591, 2006.

[97] M. Nenkova, Z. Ivezić, and M. Elitzur, "Dust emission from active galactic nuclei," *Astrophysical Journal*, vol. 570, no. 1, pp. L9–L12, 2002.

[98] E. Treister, J. H. Krolik, and C. Dullemond, "Measuring the fraction of obscured quasars by the infrared luminosity of unobscured quasars," *Astrophysical Journal*, vol. 679, no. 1, pp. 140–148, 2008.

[99] C. N. Cardamone, C. M. Urry, M. Damen et al., "Mid-infrared properties and color selection for X-ray-detected active galactic nuclei in the MUSYC Extended Chandra Deep Field-South," *Astrophysical Journal*, vol. 680, no. 1, pp. 130–142, 2008.

[100] M. Schmidt, "Space distribution and luminosity functions of Quasi-Stellar radio sources," *Astrophysical Journal*, vol. 151, p. 393, 1968.

[101] H. L. Marshall, "The evolution of optically selected quasars with Z less than 2.2 and B less than 20," *Astrophysical Journal*, vol. 299, pp. 109–121, 1985.

[102] G. Mathez, "Evolution of the luminosity function of quasars—a model with constant density and luminosity evolution," *Astronomy & Astrophysics*, vol. 53, pp. 15–21, 1976.

[103] G. Hasinger, T. Miyaji, and M. Schmidt, "Luminosity-dependent evolution of soft X-ray selected AGN: new Chandra and XMM-Newton surveys," *Astronomy and Astrophysics*, vol. 441, no. 2, pp. 417–434, 2005.

[104] Y. Ueda, M. Akiyama, K. Ohta, and T. Miyaji, "Cosmological evolution of the hard X-ray active galactic nucleus luminosity function and the origin of the hard X-ray background," *Astrophysical Journal*, vol. 598, no. 2, pp. 886–908, 2003.

[105] A. J. Barger, L. L. Cowie, R. F. Mushotzky et al., "The cosmic evolution of hard X-ray-selected active galactic nuclei," *Astronomical Journal*, vol. 129, no. 2, pp. 578–609, 2005.

[106] S. Cristiani, D. M. Alexander, F. Bauer et al., "The space density of high-redshift QSOs in the Great Observatories Origins Deep Survey," *Astrophysical Journal*, vol. 600, no. 2, pp. L119–L122, 2004.

[107] X. Fan, V. K. Narayanan, R. H. Lupton et al., "A survey of $z >$ 5.8 quasars in the sloan digital sky survey. I. Discovery of three new quasars and the spatial density of luminous quasars at $z \sim 6$," *Astronomical Journal*, vol. 122, no. 6, pp. 2833–2849, 2001.

[108] C. J. Willott, P. Delorme, C. Reylé et al., "The Canada-France high-z quasar survey: nine new quasars and the luminosity function at redshift 6," *Astronomical Journal*, vol. 139, no. 3, pp. 906–918, 2010.

[109] F. Civano et al., "The population of high-redshift active galactic nuclei in the Chandra-COSMOS survey," *Astrophysical Journal*, vol. 741, p. 91, 2011.

[110] R. Maiolino and G. H. Rieke, "Low-luminosity and obscured seyfert nuclei in nearby galaxies," *Astrophysical Journal*, vol. 454, no. 1, pp. 95–105, 1995.

[111] D. R. Ballantyne, J. E. Everett, and N. Murray, "Connecting galaxy evolution, star formation, and the cosmic X-ray background," *Astrophysical Journal*, vol. 639, no. 2, pp. 740–752, 2006.

[112] A. T. Steffen, A. J. Barger, L. L. Cowie, R. F. Mushotzky, and Y. Yang, "The changing active galactic nucleus population," *Astrophysical Journal*, vol. 596, no. 1, pp. L23–L26, 2003.

[113] C. Simpson, "The luminosity dependence of the type 1 active galactic nucleus fraction," *Monthly Notices of the Royal Astronomical Society*, vol. 360, no. 2, pp. 565–572, 2005.

[114] A. Akylas and I. Georgantopoulos, "Can photo-ionisation explain the decreasing fraction of X-ray obscured AGN with luminosity?" *Astronomy and Astrophysics*, vol. 479, no. 3, pp. 735–740, 2008.

[115] E. Treister, S. Virani, E. Gawiser et al., "Optical spectroscopy of X-ray sources in the extended Chandra deep field south," *Astrophysical Journal*, vol. 693, pp. 1713–1727, 2009.

[116] R. Barvainis, "Hot dust and the near-infrared bump in the continuum spectra of quasars and active galactic nuclei," *Astrophysical Journal*, vol. 320, pp. 537–544, 1987.

[117] F. L. A. Franca, F. Fiore, A. Comastri et al., "The HELLAS2XMM survey. VII. The hard X-ray luminosity function of AGNs up to $z = 4$: More absorbed AGNs at low luminosities and high redshifts," *Astrophysical Journal*, vol. 635, no. 2, pp. 864–879, 2005.

[118] G. Hasinger, "Absorption properties and evolution of active galactic nuclei," *Astronomy and Astrophysics*, vol. 490, no. 3, pp. 905–922, 2008.

[119] A. Akylas, I. Georgantopoulos, A. Georgakakis, S. Kitsionas, and E. Hatziminaoglou, "XMM-Newton and Chandra measurements of the AGN intrinsic absorption: dependence on luminosity and redshift," *Astronomy and Astrophysics*, vol. 459, no. 3, pp. 693–701, 2006.

[120] D. Lutz, R. Maiolino, H. W. W. Spoon, and A. F. M. Moorwood, "The relation between AGN hard X-ray emission and mid-infrared continuum from ISO spectra: scatter and unification aspects," *Astronomy and Astrophysics*, vol. 418, no. 2, pp. 465–473, 2004.

[121] J. J. Bock, K. A. Marsh, M. E. Ressler, and M. W. Werner, "High-resolution mid-infrared imaging of the nucleus of NGC 1068," *Astrophysical Journal*, vol. 504, no. 1, pp. L5–L10, 1998.

[122] J. T. Radomski, R. K. Piña, C. Packham et al., "Resolved mid-infrared emission in the narrow-line region of NGC 4151," *Astrophysical Journal*, vol. 587, no. 1, pp. 117–122, 2003.

[123] W. Saunders, M. Rowan-Robinson, A. Lawrence et al., "The 60-micron and far-infrared luminosity functions of IRAS galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 242, pp. 318–337, 1990.

[124] K. Iwasawa, D. B. Sanders, A. S. Evans, N. Trentham, G. Miniutti, and H. W. W. Spoon, "Fe K emission in the ultra-luminous infrared galaxy Arp 220," *Monthly Notices of the Royal Astronomical Society*, vol. 357, pp. 565–571, 2005.

[125] C. K. Seyfert, "Nuclear emission in spiral nebulae," *Astrophysical Journal*, vol. 97, p. 28, 1943.

[126] G. Risaliti, R. Maiolino, and M. Salvati, "The distribution of absorbing column densities among Seyfert 2 galaxies," *Astrophysical Journal*, vol. 522, no. 1, pp. 157–164, 1999.

[127] J. B. de Veny, W. H. Osborn, and K. Janes, "A catalogue of quasars," *Astronomical Society of the Pacific*, vol. 83, p. 611, 1971.

[128] G. Matt, A. C. Fabian, M. Guainazzi, K. Iwasawa, L. Bassani, and G. Malaguti, "The X-ray spectra of compton-thick Seyfert 2 galaxies as seen by BeppoSAX," *Monthly Notices of the Royal Astronomical Society*, vol. 318, pp. 173–179, 2000.

[129] M. H. K. de Grijp, G. K. Miley, J. Lub, and T. De Jong, "Infrared Seyferts: a new population of active galaxies?" *Nature*, vol. 314, no. 6008, pp. 240–242, 1985.

[130] C. Winkler, T. J. L. Courvoisier, G. Di Cocco et al., "The INTEGRAL mission," *Astronomy and Astrophysics*, vol. 411, no. 1, pp. L1–L6, 2003.

[131] N. Gehrels, G. Chincarini, P. Giommi et al., "The Swift gamma-ray burst mission," *Astrophysical Journal*, vol. 611, no. 2, pp. 1005–1020, 2004.

[132] P. Ubertini, F. Lebrun, G. Di Cocco et al., "IBIS: the imager on-board INTEGRAL," *Astronomy and Astrophysics*, vol. 411, no. 1, pp. L131–L139, 2003.

[133] P. Vignati, S. Molendi, G. Matt et al., "BeppoSAX unveils the nuclear component in NGC 6240," *Astronomy and Astrophysics*, vol. 349, no. 2, pp. L57–L60, 1999.

[134] N. A. Levenson, T. M. Heckman, J. H. Krolik, K. A. Weaver, and P. T. Zycki, "Penetrating the deep cover of compton-thick active galactic nuclei," *Astrophysical Journal*, vol. 648, no. 1, pp. 111–127, 2006.

[135] G. Ghisellini, F. Haardt, and G. Matt, "The contribution of the obscuring torus to the X-ray spectrum of Seyfert galaxies—a test for the unification model," *Monthly Notices of the Royal Astronomical Society*, vol. 267, p. 743, 1994.

[136] D. M. Alexander, W. N. Brandt, A. E. Hornschemeier et al., "The Chandra deep field north survey. VI. The nature of the optically faint X-ray source population," *Astronomical Journal*, vol. 122, no. 5, pp. 2156–2176, 2001.

[137] A. J. Barger, L. L. Cowie, P. Capak et al., "Optical and infrared properties of the 2 Ms Chandra Deep Field North X-ray sources," *Astronomical Journal*, vol. 126, no. 2, pp. 632–665, 2003.

[138] V. Mainieri, D. Rigopoulou, I. Lehmann et al., "Submillimetre detection of a high-redshift type 2 QSO," *Monthly Notices of the Royal Astronomical Society*, vol. 356, no. 4, pp. 1571–1575, 2005.

[139] I. Georgantopoulos, A. Akylas, A. Georgakakis, and M. Rowan-Robinson, "The Compton-thick AGN in the Chandra Deep Field North," *Astronomy and Astrophysics*, vol. 507, no. 2, pp. 747–756, 2009.

[140] E. Daddi, D. M. Alexander, M. Dickinson et al., "Multiwavelength study of massive galaxies at $z \sim 2$. II. Widespread

compton-thick active galactic nuclei and the concurrent growth of black holes and bulges," *Astrophysical Journal*, vol. 670, no. 1, pp. 173–189, 2007.

[141] F. Fiore, A. Grazian, P. Santini et al., "Unveiling obscured accretion in the Chandra Deep Field-South," *Astrophysical Journal*, vol. 672, no. 1, pp. 94–101, 2008.

[142] I. Georgantopoulos, A. Georgakakis, M. Rowan-Robinson, and E. Rovilos, "Searching for mid-IR obscured AGN in the Chandra deep field North," *Astronomy and Astrophysics*, vol. 484, no. 3, pp. 671–678, 2008.

[143] J. L. Donley, G. H. Rieke, P. G. Pérez-González, and G. Barro, "Spitzer's contribution to the AGN population," *Astrophysical Journal*, vol. 687, no. 1, pp. 111–132, 2008.

[144] A. Pope, R. S. Bussmann, A. Dey et al., "The nature of faint Spitzer-selected dust-obscured galaxies," *Astrophysical Journal*, vol. 689, no. 1, pp. 127–133, 2008.

[145] A. Georgakakis, M. Rowan-Robinson, K. Nandra, J. Digby-North, P. G. Pérez-González, and G. Barro, "Infrared excess sources: compton thick QSOs, low-luminosity seyferts or starbursts?" *Monthly Notices of the Royal Astronomical Society*, vol. 406, no. 1, pp. 420–433, 2010.

[146] M. Salvato, G. Hasinger, O. Ilbert et al., "Photometric redshift and classification for the XMM-cosmos sources," *Astrophysical Journal*, vol. 690, no. 2, pp. 1250–1263, 2009.

[147] C. N. Cardamone, P. G. Van Dokkum, C. M. Urry et al., "The multiwavelength Survey by Yale-Chile (MUSYC): deep medium-band optical imaging and high-quality 32-band photometric redshifts in the ECDF-S," *Astrophysical Journal Supplement Series*, vol. 189, no. 2, pp. 270–285, 2010.

[148] P. F. Hopkins, L. Hernquist, T. J. Cox, and D. Kerbs, "A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. i. galaxy mergers and quasar activity," *Astrophysical Journal Supplement Series*, vol. 175, no. 2, pp. 356–389, 2008.

[149] J. S. Kartaltepe, D. B. Sanders, E. Le Floc'H et al., "A multiwavelength study of a sample of 70 μm selected galaxies in the cosmos field. I. Spectral energy distributions and luminosities," *Astrophysical Journal*, vol. 709, no. 2, pp. 572–596, 2010.

[150] E. Treister, C. M. Urry, K. Schawinski, C. N. Cardamone, and D. B. Sanders, "Heavily obscured active galactic nuclei in high-redshift luminous infrared galaxies," *Astrophysical Journal*, vol. 722, no. 2, pp. L238–L243, 2010.

[151] B. Luo, F. E. Bauer, W. N. Brandt et al., "The Chandra deep field-south survey: 2 Ms source catalogs," *Astrophysical Journal Supplement Series*, vol. 179, no. 1, pp. 19–36, 2008.

[152] E. Treister, K. Schawinski, M. Volonteri, P. Natarajan, and E. Gawiser, "Black hole growth in the early Universe is self-regulated and largely hidden from view," *Nature*, vol. 474, no. 7351, pp. 356–358, 2011.

[153] R. J. Bouwens, G. D. Illingworth, J. P. Blakeslee, and M. Franx, "Galaxies at $z \sim 6$: the UV luminosity function and luminosity density from 506 HUDF, HUDF parallel ACS field, and goods i-dropouts," *Astrophysical Journal*, vol. 653, no. 1, pp. 53–85, 2006.

[154] F. Fiore, S. Puccetti, and S. Mathur, "Demography of high redshift,". In press.

[155] C. J. Willott, "No evidence of obscured, accreting black holes in most $z = 6$ star-forming galaxies," *Astrophysical Journal Letters*, vol. 742, no. 1, 2011.

[156] L. L. Cowie, A. J. Barger, and G. Hasinger, "The faintest X-ray sources from $z = 0$–8," . In press.

[157] Y. Ueda, S. Eguchi, Y. Terashima et al., "Suzaku observations of active galactic nuclei detected in the SWIFT BAT survey: discovery of a "new type" of buried supermassive black holes," *Astrophysical Journal*, vol. 664, no. 2 II, pp. L79–L82, 2007.

[158] R. Gilli, J. Su, C. Norman et al., "A Compton-thick active galactic nucleus at $z \sim 5$ in the 4 Ms Chandra deep field south," *The Astrophysical Journal Letters*, vol. 730, no. 2, p. L28, 2011.

[159] M. Volonteri and N. Y. Gnedin, "Relative role of stars and quasars in cosmic reionization," *Astrophysical Journal*, vol. 703, no. 2, pp. 2113–2117, 2009.

[160] C. A. Faucher-GIGUÈRE, A. Lidz, L. Hernquist, and M. Zaldarriaga, "Evolution of the intergalactic opacity: implications for the ionizing background, cosmic star formation, and quasar activity," *Astrophysical Journal*, vol. 688, no. 1, pp. 85–107, 2008.

[161] P. Salucci, E. Szuszkiewicz, P. Monaco, and L. Danese, "Mass function of dormant black holes and the evolution of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 307, no. 3, pp. 637–644, 1999.

[162] Q. Yu and S. Tremaine, "Observational constraints on growth of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 4, pp. 965–976, 2002.

[163] A. Marconi, G. Risaliti, R. Gilli, L. K. Hunt, R. Maiolino, and M. Salvati, "Local supermassive black holes, relics of active galactic nuclei and the X-ray background," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 1, pp. 169–185, 2004.

[164] F. Shankar, P. Salucci, G. L. Granato, G. De Zotti, and L. Danese, "Supermassive black hole demography: the match between the local and accreted mass functions," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 4, pp. 1020–1030, 2004.

[165] P. Natarajan and E. Treister, "Is there an upper limit to black hole masses?" *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 3, pp. 838–845, 2009.

[166] F. Shankar, D. H. Weinberg, and J. Miralda-Escudé, "Self-consistent models of the AGN and black hole populations: duty cycles, accretion rates, and the mean radiative efficiency," *Astrophysical Journal*, vol. 690, no. 1, pp. 20–41, 2009.

[167] M. Vestergaard and B. M. Peterson, "Determining central black hole masses in distant active galaxies and quasars. II. Improved optical and UV scaling relationships," *Astrophysical Journal*, vol. 641, no. 2, pp. 689–709, 2006.

[168] R. J. McLure and J. S. Dunlop, "The cosmological evolution of quasar black hole masses," *Monthly Notices of the Royal Astronomical Society*, vol. 352, no. 4, pp. 1390–1404, 2004.

[169] M. Vestergaard, "Early growth and efficient accretion of massive black holes at high redshift," *Astrophysical Journal*, vol. 601, no. 2, pp. 676–691, 2004.

[170] B. C. Kelly, M. Vestergaard, X. Fan, P. Hopkins, L. Hernquist, and A. Siemiginowska, "Constraints on black hole growth, quasar lifetimes, and Eddington ratio distributions from the SDSS broad-line quasar black hole mass function," *Astrophysical Journal*, vol. 719, no. 2, pp. 1315–1334, 2010.

[171] J. A. Kollmeier, C. A. Unken, C. S. Kochanek et al., "Black hole masses and eddington ratios at $0.3 < z < 4$," *Astrophysical Journal*, vol. 648, no. 1, pp. 128–139, 2006.

[172] J. R. Trump et al., "Observational limits on type 1 active galactic nucleus accretion rate in COSMOS," *The Astrophysical Journal*, vol. 700, pp. 49–55, 2009.

[173] P. Natarajan, "The formation and evolution of massive black hole seeds in the early universe," *Bulletin of the Astronomical Society of India*, vol. 39, no. 1, pp. 145–161, 2011.

[174] G. Lodato and P. Natarajan, "The mass function of high-redshift seed black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L64–L68, 2007.

[175] D. R. Ballantyne, A. R. Draper, K. K. Madsen, J. R. Rigby, and E. Treister, "Lifting the veil on obscured accretion: active galactic nuclei number counts and survey strategies for imaging hard X-ray missions," *Astrophysical Journal*, vol. 736, no. 1, p. 56, 2011.

[176] T. Takahashi, K. Mitsuda, R. Kelley et al., "The ASTRO-H mission," in *Proceedings of the Space Telescopes and Instrumentation: Ultraviolet to Gamma Ray*, vol. 7732 of *Proceedings of SPIE*, 2010.

[177] L. Yao, E. R. Seaquist, N. Kuno, and L. Dunne, "CO molecular gas in infrared-luminous galaxies," *Astrophysical Journal*, vol. 588, no. 2, pp. 771–791, 2003.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 970858, 21 pages
doi:10.1155/2012/970858

Review Article

Mass Functions of Supermassive Black Holes across Cosmic Time

Brandon C. Kelly[1, 2] and Andrea Merloni[3]

[1] *High Energy Astrophysics Division, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[2] *Department of Physics, Broida Hall, University of California, Santa Barbara, Santa Barbara, CA 93106, USA*
[3] *Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstraße 1, 85748 Garching, Germany*

Correspondence should be addressed to Brandon C. Kelly, bckelly@cfa.harvard.edu

Received 10 September 2011; Accepted 1 December 2011

Academic Editor: Francesca Civano

The black hole mass function of supermassive black holes describes the evolution of the distribution of black hole mass. It is one of the primary empirical tools available for mapping the growth of supermassive black holes and for constraining theoretical models of their evolution. In this paper, we discuss methods for estimating the black hole mass function, including their advantages and disadvantages. We also review the results of using these methods for estimating the mass function of both active and inactive black holes. In addition, we review current theoretical models for the growth of supermassive black holes that predict the black hole mass function. We conclude with a discussion of directions for future research which will lead to improvement in both empirical and theoretical determinations of the mass function of supermassive black holes.

1. Introduction

Understanding how and when supermassive black holes (SMBHs) grow is currently of central importance in extragalactic astronomy. A significant amount of empirical work has established correlations between SMBH mass and host galaxy spheroidal properties, such as luminosity [1–3], stellar velocity dispersion (the $M_{\rm BH}$-σ_* relationship, e.g., [4–6]), concentration or Sersic index [7, 8], bulge mass [9–11], and binding energy [12, 13]. These scaling relationships imply that the evolution of spheroidal galaxies and the growth of SMBHs are intricately tied together. The currently favored mechanism for linking the growth of SMBHs and their hosts is black hole feedback, whereby black holes grow by accreting gas in the so-called "active" phases, possibly fueled by a major merger of two gas-rich galaxies, until feedback energy from the SMBH expels gas and shuts off the accretion process [14–17]. Alternatively, it has been suggested that the origin of the scaling relationships does not necessarily require SMBH feedback but emerges from the stochastic nature of the hierarchical assembly of black hole and stellar mass through galaxy mergers [18, 19].

Feedback-driven "self-regulated" growth of black holes has been able to reproduce the local $M_{\rm BH}$-σ_* relationship in smoothed particle hydrodynamics simulations [20–22]. Moreover, AGN feedback has also been invoked as a means of quenching the growth of the most massive galaxies [23, 24]. There have been numerous models linking SMBH growth, the quasar phase, and galaxy evolution [25–33]. While feedback is likely important for regulating the growth of SMBHs and galaxies, the fueling mechanisms that contribute to growing the SMBH are likely diverse. Major mergers of gas-rich galaxies may fuel quasars at high redshift and grow the most massive SMBHs. However, major mergers alone do not appear to be sufficient to reproduce the number of X-ray faint AGN [34], and accretion of ambient gas via internal galactic processes [35, 36] may fuel these fainter, lower $M_{\rm BH}$ AGN at lower z. This is supported by the fact that many AGN are observed to live in late-type galaxies out to $z \approx 1$ [37, 38], and the X-ray luminosity function of AGN hosted by late-type galaxies suggests that fueling by minor interactions or internal instabilities represents a nonnegligible contribution to the accretion history of the universe [39].

The black hole mass function (BHMF) provides a complete census of the mass of SMBHs and their evolution. Because of this, the BHMF is one of the primary empirical tools available for investigating the growth of SMBHs, and for constraining theoretical models for the growth of the SMBH

population. Because SMBHs and galaxies are thought to be linked in their evolution, the BHMF provides insight into the fueling mechanisms that dominate black hole growth and therefore into the role of feedback in the evolution of the host galaxy. The BHMF is also an important tool in planning future surveys, as it provides an estimate of the distribution of SMBH mass expected for the survey. This in turn is important because mass is a fundamental quantity of the black hole and therefore is an important observational quantity for empirical studies of black hole accretion physics [40–45]. Of course, further improvement to our understanding of black hole accretion physics will further improve our modeling and understanding of black hole accretion and feedback, which in turn will improve our understanding of black-hole-galaxy coevolution. Therefore, the BHMF is an important empirical quantity for SMBH studies.

In this paper, we discuss the current status of BHMF estimation and theoretical modeling. In Section 2, we discuss the nontrivial task of estimating the BHMF. In Section 3, we discuss current estimates of the local SMBH BHMF. In Section 4, we discuss BHMF estimates derived by combining the local BHMF with the AGN luminosity function via a continuity equation. In Section 5, we discuss BHMFs estimated for AGN only. In Section 6, we review theoretical models for SMBH growth that predict the SMBH BHMF. Finally, in Section 7, we discuss directions for future improvements to the empirical and theoretical studies of the BHMF. We note that unlike, say, the luminosity function, the division between "observational" and "theoretical" studies is not as clear for the BHMF, as some amount of modeling is necessary in order to estimate the BHMF from strictly observational quantities. We have attempted to divide the studies according to whether the BHMF is constrained empirically, as in, say, a formal statistical fitting procedure, or if it is predicted from a theoretical model for SMBH growth. In reality, the line between theoretical and empirical studies is blurry and some procedures which we have considered to be empirical may be thought of as theoretical.

2. Estimating the Black Hole Mass Function

The black hole mass function, denoted as $\phi(M_{\rm BH}, z)dM_{\rm BH}$, is the number of sources per comoving volume $V(z)$ with black hole masses in the range $M_{\rm BH}$, $M_{\rm BH} + dM_{\rm BH}$. The black hole mass function is related to the joint probability distribution of $M_{\rm BH}$ and z, $p(M_{\rm BH}, z)$, as

$$\phi(M_{\rm BH}, z) = N\left(\frac{dV}{dz}\right)^{-1} p(M_{\rm BH}, z). \quad (1)$$

The normalization of the BHMF is N, the total number of SMBHs in the observable universe, and is given by the integral of ϕ over $M_{\rm BH}$ and $V(z)$.

2.1. Complications with Estimating Black Hole Mass Functions. Similar to luminosity function estimation, the BHMF may be estimated from astronomical surveys. However, while there are many well-established methods for estimating luminosity functions, there are two complications that make BHMF estimation a more difficult problem [46]. The first is the issue of incompleteness. Surveys are typically constructed by finding the set of objects of interest containing a SMBH that satisfy a flux criteria, for example, all objects brighter than some flux limit. Surveys are not constructed by selecting on mass. Because there is a distribution of luminosities at a given SMBH mass, whether it is the luminosity of the host galaxy or of the AGN, some SMBHs will scatter above the flux limit and some below. This creates a selection function which is less sensitive to $M_{\rm BH}$, and it is possible that a survey may be incomplete in all mass bins.

The second complication is the large uncertainty in SMBH mass among mass estimators. Currently, it is not possible to obtain reliable mass estimators for large numbers of SMBHs through dynamical and modeling of the stellar or gaseous components, and thus scaling relationships are employed. Masses may be estimated using scaling relationships between $M_{\rm BH}$ and the properties of the host galaxy bulge or the luminosity and the width of the broad emission lines for AGN [47, 48]. It has also recently been suggested that the X-ray variability properties of AGN may also provide another scaling relationship for estimating $M_{\rm BH}$ [45, 49, 50], but further work is needed for developing this. While these scaling relationships enable one to estimate $M_{\rm BH}$ for large numbers of SMBHs, they also contain a significant intrinsic statistical scatter. Gültekin et al. [51] find that for early-type galaxies there is an intrinsic scatter in $M_{\rm BH}$ of 0.31 ± 0.06 dex and 0.38 ± 0.09 dex at fixed host galaxy bulge dispersion and luminosity, respectively; the amplitude of the scatter is larger for late-type galaxies. For AGN with broad emission lines, Vestergaard and Peterson [48] estimate the scatter in $M_{\rm BH}$ at fixed luminosity and line width to be $\sim$0.4 dex, depending on which emission line is used.

The statistical uncertainty in the mass estimates can have a significant effect on the inferred BHMF. The distribution of the mass estimates is the convolution of the intrinsic BHMF with the error distribution in the mass estimates. In general, it is typically assumed that the error in the mass estimates is independent of the actual value of $M_{\rm BH}$. This is not the case for $M_{\rm BH}$ estimated through dynamical modeling; however, independence between $M_{\rm BH}$ and its error is likely to be a good approximation for $M_{\rm BH}$ estimated using scaling relationships. Because scaling relationships are the only feasible manner to estimate $M_{\rm BH}$ for a large sample of SMBHs, which is necessary for any estimate of the BHMF, we will assume that $M_{\rm BH}$ and its error are independent. Under the assumption of independence between the estimated $M_{\rm BH}$ and its error, the BHMF that would be inferred directly from the distribution of the mass estimates is broader than the intrinsic BHMF and is thus biased. Figure 1 illustrates this effect, where an intrinsic mass function is compared with the distribution of an unbiased mass estimator having a statistical uncertainty of 0.3, 0.4, and 0.5 dex, respectively. As can be seen, the distribution of mass estimates is significantly different from the intrinsic mass function. In particular, the distribution of the mass estimates falls off more slowly with increasing $M_{\rm BH}$, and overpredicts the number of SMBHs at the high $M_{\rm BH}$ end of the mass function. The bias is worse when the dispersion in the scatter in the mass estimates becomes larger.

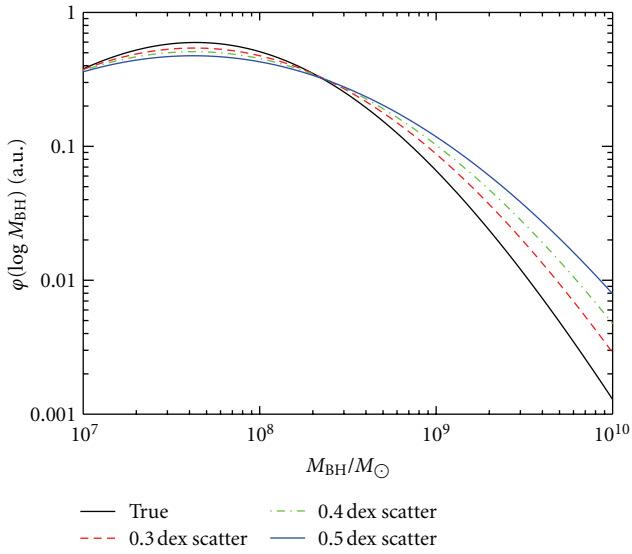


Figure 1: Illustration of the bias in the estimated BHMF derived from mass estimates. Shown is the true mass function (thick solid black line) for a simulated sample, and the mass function derived from the mass estimateswhen the statistical error in the mass estimates is 0.3 dex (red dashed line), 0.4 dex (green dot-dashed line), and 0.5 dex (solid blue thin line). The mass function estimated from the mass estimates is biased, especially at the high-MBH end and for large statistical error.

2.2. Methodology for Estimating the Black Hole Mass Function. In order to estimate the SMBH mass function in an unbiased manner, it is necessary to match the mass function with the observed distribution of the mass estimates and any additional observational quantities that the selection function (The selection function is the probability of including a source in one's sample as a function of its measured quantities) depends on. The basic idea is to start with an assumed mass function. Then, calculate the distribution of mass estimates implied by this mass function. In addition, calculate the distribution of observational quantities that one's sample is selected on, say, flux, that is implied by the assumed mass function. This step allows one to correct for incompleteness but requires an additional assumption about how to relate the mass function to the quantity that one's sample is selected on. Finally, impose the selection function for the sample and compare the predicted observed distributions of mass estimates and any other observables (e.g., flux) with the actual distributions. If they are not consistent, then the data rule out the assumed mass function and relationship between $M_{\rm BH}$ and the observable quantities.

We can make the above procedure more quantitative by deriving the likelihood function for the SMBH mass function. Kelly et al. [46] derived the likelihood function for the mass function when using masses estimated from AGN broad emission lines. They used this likelihood function for developing a Bayesian approach to estimating the SMBH mass function. Although their method was limited to broad-line mass estimates, it is straightforward to generalize their formalism for any generic mass estimator. Denote the black hole mass estimate as $\widehat{M}_{\rm BH}$. In addition, denote as X the set of observables that one uses to select one's sample. In the majority of cases, this will be flux at one or more wavelengths. Then, the likelihood function for the BHMF based on a sample of n SMBHs is

$$\begin{aligned} p\left(\widehat{\mathbf{M}}_{\rm BH}, \mathbf{X}, \mathbf{z} \mid \theta, \psi, N\right) \\ \propto C_n^N [s(\theta,\psi)]^{N-n} \prod_{i=1}^{n} \int_0^\infty p\left(\widehat{M}_{{\rm BH},i} \mid M_{{\rm BH},i}, z_i, X_i\right) \\ \times p(X_i \mid M_{{\rm BH},i}, z_i, \psi) \\ \times p(M_{{\rm BH},i}, z_i \mid \theta) dM_{{\rm BH},i}, \end{aligned} \tag{2}$$

where the BHMF is related to N and the probability distribution of $M_{\rm BH}$ and z via (1). Here, C_n^N is the binomial coefficient, θ denotes the parameters for the BHMF, ψ denotes the parameters for the distribution in X at fixed $M_{\rm BH}$ and z, and $s(\theta,\psi)$ is the probability of including a SMBH in one's sample as a function of θ and ψ. Here, we have assumed that the distribution in the mass estimates at fixed $M_{\rm BH}$, z, and X, $p(\widehat{M}_{\rm BH} \mid M_{\rm BH}, z, X)$ is known, although one could include additional free parameters for this as well. The binomial coefficient arises from the fact that the number of objects included in one's survey follows a binomial distribution (Often in the luminosity function literature the likelihood is assumed to be a Poisson distribution. A poisson distribution is an approximation to the binomial distribution when $N \to \infty$ and $s(\theta,\psi) \to 0$, so (2) converges to the Poisson distribution when $n \ll N$. See [52] for further details.) with N "trials" and probability of success $s(\theta,\psi)$. The probability of including a SMBH in one's survey as a function of the BHMF, $s(\theta,\psi)$, is calculated from the survey selection function $s(X,z)$ as

$$\begin{aligned} s(\theta,\psi) = \int_{X_{\rm min}}^{X_{\rm max}} s(X,z) \Bigg[\iint_0^\infty p(X \mid M_{\rm BH}, z, \psi) \\ \times p(M_{\rm BH}, z \mid \theta) dM_{\rm BH}\, dz \Bigg] dX. \end{aligned} \tag{3}$$

It is up to the researcher to choose the particular parametric form for the SMBH mass function, the distribution in the mass estimates at fixed $M_{\rm BH}$, z, and X, and the distribution of the observable that the sample is selected on (e.g., flux) at fixed $M_{\rm BH}$ and z. Typical choices are log-normal distributions, Schechter functions, and mixtures of log-normal distributions. Once one has done this, one can use (2) to compute a maximum-likelihood estimate for the BHMF or perform Bayesian inference.

An alternative form of estimating the BHMF can be used when the mass estimates are derived from an observational quantity, Y, and the intrinsic distribution of Y is known. This is commonly used to estimate the local mass function using host-galaxy scaling relationships [53]. In this case, the mass function is

$$\phi(M_{\rm BH}) = \int_{Y_{\rm min}}^{Y_{\rm max}} p(M_{\rm BH} \mid Y) \phi(Y) dY, \tag{4}$$

where $\phi(Y)$ is the comoving number density of SMBHs as a function of the quantity Y. When both $p(M_{\rm BH} \mid Y)$ and $\phi(Y)$ are known, then the BHMF follows directly from (4). As an example, if the mass function is derived from the scaling between $M_{\rm BH}$ and host galaxy spheroidal luminosity, $L_{\rm sph}$, then $Y = L_{\rm sph}$, $p(M_{\rm BH} \mid Y)$ is the $M_{\rm BH}$-$L_{\rm sph}$ relationship and $\phi(Y)$ is the luminosity function of stellar bulges hosting SMBHs. As with BHMFs determined from a mass estimator, improper treatment of the intrinsic scatter in $M_{\rm BH}$ at fixed Y will lead to a biased estimate of the BHMF. However, when calculating the BHMF from (4), ignoring the intrinsic scatter results in an estimated BHMF that is too narrow, underpredicting the number of SMBHs at the high-mass end of $\phi(M_{\rm BH})$. This is opposite to the case when one estimates the BHMF directly from mass estimates.

3. Black Hole Mass Functions Derived from Host Galaxy Scaling Relationships

The observed scaling between $M_{\rm BH}$ and the properties of the SMBH host galaxy bulge have motivated several groups to estimate the local BHMF [53–65], with decreasing statistical uncertainties. These estimates of the local BHMF have formed the basis for many studies which have attempted to map black hole growth by comparing with the AGN luminosity function, this is further discussed in Section 4. Typically, the local BHMF is estimated using the local $M_{\rm BH}$-σ_* relationship or the local $M_{\rm BH}$-$L_{\rm sph}$ relationship, combined with the local number density of galaxies as a function of stellar velocity dispersion or bulge luminosity.

The scaling relationships between $M_{\rm BH}$ and host galaxy properties are only determined for the local universe, and thus most authors have limited their determination of the BHMF based on them to the local BHMF. There are, however, a couple of exceptions. Tamura et al. [66], estimated the BHMF out to $z \approx 1$ assuming that evolution in the $M_{\rm BH}$-$L_{\rm sph}$ relation is driven only by passive evolution in $L_{\rm sph}$. Shankar et al. [67] (see also [68, 69]) used the local velocity dispersion function of spheroids in combination with their inferred age distributions to estimate the BHMF at $z < 6$. In order to do this, they assumed that most of the stars in nearby spheroids formed in a single event and that σ_* did not change once the spheroid was formed. In addition, Shankar et al. [67] allowed the normalization of the $M_{\rm BH}$-σ_* relationship to evolve, with the degree of evolution being a free parameter. They found evidence for mild evolution in the normalization of the $M_{\rm BH}$-σ_* relationship.

Evolution in the scaling relationships is currently an area of intense study, with most groups finding evidence that the normalization of the scaling relationships increases towards higher z [70–75], at least for active SMBHs. However, there are still concerns regarding potential biases due to selection effects [76], but see Treu et al. [70] and Bennert et al. [75] for procedures aimed at modeling and correcting for selection. There may also be biases due to extrapolating the AGN mass estimates derived from the broad emission lines to luminous quasars at high z [77]. As such, the uncertainties on the quantitative form of the evolution in the scaling relationships and their scatter are currently large, limiting their use for determining the BHMF outside of the local universe.

When the $M_{\rm BH}$-σ_* relationship is used to estimate the local BHMF, it is common to use the velocity dispersion distribution derived from the SDSS by Sheth et al. [78], with an additional component representing the brightest cluster galaxies [58]. Sheth et al. [78] estimate the velocity dispersion distribution for late-type galaxies by using the Tully-Fisher relation to convert the luminosity function of late-type galaxies to a circular velocity distribution and then set $\sigma = v_c/\sqrt{2}$. When the $M_{\rm BH}$-$L_{\rm sph}$ relation is used, it is typical to estimate the distribution of $L_{\rm sph}$ separately for early- and late-type galaxies by converting their respective luminosity functions to spheroidal luminosity functions using an assumed ratio of bulge luminosity to total luminosity. From this, it has been inferred that the local BHMF is dominated by early-type galaxies at $M_{\rm BH} \gtrsim 4 \times 10^7 M_\odot$ [62]. Shankar et al. [64] present a compilation of recently determined local BHMFs based on a variety of methods, scaling relations used, and data sets used. In Figure 2, we show the range of local BHMFs estimated from the $M_{\rm BH}$-σ_*, $M_{\rm BH}$-$L_{\rm sph}$, and $M_{\rm BH}$-$M_{\rm star}$ relationships, as presented in Shankar et al. [64]. In general, estimates of the local mass density of SMBHs span the range $\rho_{\rm BH} = (3.2\text{–}5.4) \times 10^5 M_\odot {\rm Mpc}^{-3}$ for $h = 0.7$ [64].

While the procedure for estimating the local BHMF is, in theory, straightforward, a number of significant systematics remain. First, there is the observational difficulty that most BHMFs derived from the $M_{\rm BH}$-σ_* relationship are based on SDSS spectra. Unfortunately, the SDSS velocity dispersions are based on a fixed aperture, and thus the size of the aperture relative to the bulge varies with the apparent size of the galaxy and its inclination. In addition, the spectral resolution of SDSS spectra is $\sim 100\,{\rm km s}^{-1}$, making it difficult to reliably measure σ_* for SMBHs with $M_{\rm BH} < 10^7 M_\odot$. Another concern is that the local BHMF is derived by assuming that the $M_{\rm BH}$-σ_* or $M_{\rm BH}$-$L_{\rm sph}$ relations are single power laws with a constant scatter in $M_{\rm BH}$ at fixed σ_* or $L_{\rm sph}$. However, recent work has shown these assumptions to be incorrect. For one, the $M_{\rm BH}$-σ_* and $M_{\rm BH}$-$L_{\rm sph}$ relations diverge at the high-$M_{\rm BH}$ end, which Lauer et al. [58] suggest implies that the $M_{\rm BH}$-σ_* relation is not a single power law. This divergence creates an inconsistency in the BHMFs derived from these two scaling relationships [58, 61]. Similarly, the σ-L relationships for the SDSS and dynamical $M_{\rm BH}$ SMBH samples are inconsistent, suggesting a possible selection bias in the estimated BHMFs [55, 79]. The scatter in the $M_{\rm BH}$-σ_* relation is larger for spirals [51, 80], and appears to increase at low $M_{\rm BH}$ such that most SMBHs lie below the $M_{\rm BH}$-σ_* relation, see (e.g., [81]). Several authors have found differences in the slope and scatter of the scaling relations for pseudobulges [80, 82–84]; however, it is unclear that this result is due to differences in the perceived bulge velocity dispersions for bulges as compared to pseudobulges or due to different scaling relationships. Recently, Kormendy et al. [85] argue that $M_{\rm BH}$ does not correlate with galaxy disks and only correlates weakly, if at all, with pseudobulges. On other hand, Graham et al. [86] analyzed a larger sample of barred galaxies and concluded that $M_{\rm BH}$ does correlate with σ_*, even though the $M_{\rm BH}$-σ_* relationship for barred galaxies

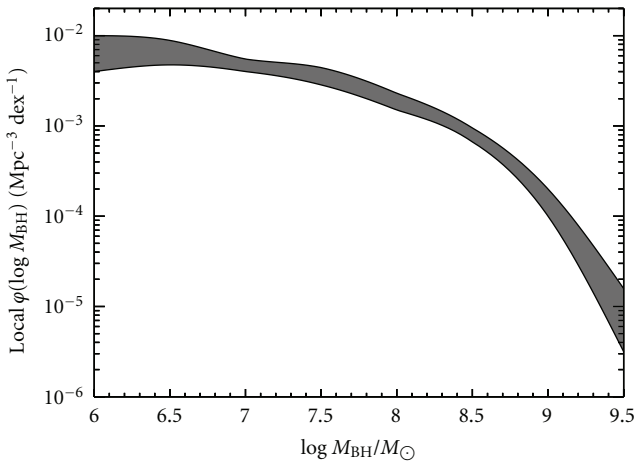


FIGURE 2: Local BHMF. The shaded region defines the spread in estimates obtained using the $M_{\rm BH}$-σ_*, $M_{\rm BH}$-$L_{\rm sph}$, and $M_{\rm BH}$-$M_{\rm star}$ relationships, as compiled by Shankar et al. [64]. Based on this estimate, the local universe is dominated by SMBHs with $M_{\rm BH} < 10^7 M_{\odot}$.

is offset from that of non barred galaxies. Although there is still much that we do not understand about the $M_{\rm BH}$ and host galaxy scaling relationship, these recent results suggest that the scaling relationships are not a single power law with constant intrinsic dispersion in $M_{\rm BH}$, representing a significant source of systematic uncertainty in the estimated local BHMF, especially at the low-mass end.

4. Black Hole Mass Functions Derived from the Local Mass Function and the AGN Luminosity Function

By employing the argument of Soltan [87], numerous studies have attempted to estimate the BHMF at a variety of redshifts by comparing the accreted mass distribution implied by the quasar luminosity function with the local BHMF [53, 55, 63, 88–92]. These methods employ a continuity equation describing the evolution of the number density of SMBHs [93, 94]:

$$\frac{\partial \phi_M(M_{\rm BH},t)}{\partial t} + \frac{\partial \phi_M(M_{\rm BH},t)\langle \dot{M}(M_{\rm BH},t)\rangle}{\partial M_{\rm BH}} = n_{\rm merge}(M_{\rm BH},t). \tag{5}$$

Here, $\langle \dot{M}(M_{\rm BH},t)\rangle$ is the average growth rate of SMBHs as a function of $M_{\rm BH}$ and cosmic age, $t(z)$, and $n_{\rm merge}(M_{\rm BH},t)$ is the rate at which the number density of SMBHs changes due to mergers of black holes, or ejections of black holes from their host galaxies due to gravitational recoil. Technically, $n_{\rm merge}(M_{\rm BH},t)$ can also include a contribution from SMBHs which are created, but this has not been thought to occur over the redshift range in which (5) is typically applied, that is, $z \lesssim 5$. Because the merger rate of black holes is currently unknown, many studies that have employed (5) set $n_{\rm merge}(M_{\rm BH},t) = 0$.

Under the assumption that SMBHs grow during phases of AGN activity, AGN demographics in combination with the local BHMF may be used to compute $\phi_M(M_{\rm BH},t)$. This is because the AGN luminosity function maps the accretion history onto SMBHs, and the local BHMF acts as a boundary condition on (5); it is also possible in principle to include the BHMF for AGN, which provides more information. Studies that have used (5) to estimate the BHMF generally fall into two categories: those that assume an AGN lightcurve and those that employ the BHMF of AGN. We discuss each of these separately.

4.1. Methods That Assume an AGN Lightcurve. Most authors employing (5) have assumed a parametric form for $\langle \dot{M}(M_{\rm BH},t)\rangle$. The accretion rate is related to the bolometric luminosity output of the accretion flow onto the SMBH as $L = \epsilon_r \dot{M}_{\rm acc} c^2$, where ϵ_r is the radiative efficiency of the accretion flow, $\dot{M}_{\rm acc}$ is the accretion rate of matter onto the SMBH, and c is the speed of light. The growth rate of the SMBH is $\dot{M} = (1-\epsilon_r)\dot{M}_{\rm acc}$, due to the fact that a fraction ϵ_r of accreted mass is radiated away as energy. Making this substitution, the continuity equation becomes

$$\frac{\partial \phi_M(M_{\rm BH},t)}{\partial t} + \frac{1-\epsilon_r}{\epsilon_r c^2}\frac{\partial \phi_M(M_{\rm BH},t)\langle L(M_{\rm BH},t)\rangle}{\partial M_{\rm BH}} = 0, \tag{6}$$

where we have ignored mergers of SMBHs. Equation (6) shows that it is possible to calculated the BHMF at a time given the local BHMF, an assumed average accretion flow lightcurve as a function of $M_{\rm BH}$, $\langle L(M_{\rm BH},t)\rangle$, and an assumed radiative efficiency. Because $\phi_M(M_{\rm BH},z)$ and $\langle L(M_{\rm BH},t(z))\rangle$ imply a luminosity function, the local BHMF and AGN luminosity function can be used to place constraints on ϵ_r and $\langle L(M_{\rm BH},t)\rangle$. This means that, in practice, one also has to assume a bolometric correction, which itself likely depends on both black hole mass [95] and $L/L_{\rm Edd}$ [96, 97]. In addition, an estimate of $\langle L(M_{\rm BH},t)\rangle$ also enables one to estimate the lifetime and duty cycle of AGN activity, modulo some luminosity-dependent definition of an AGN; note that the AGN duty cycle defines the fraction of SMBHs that are "active" at a given $M_{\rm BH}$ and z.

A variety of lightcurve models have been used when employing (6) to reconstruct the evolution of the BHMF. The simplest model is that where SMBHs spend a fraction of their time radiating at a constant Eddington ratio and spend the remainder of their time in quiescence. The free parameters in this model are the Eddington ratio, AGN lifetime or duty cycle, and radiative efficiency. This model has been used by [53–55, 57, 64] (technically, [54] assumed that the Eddington ratio was a weakly increasing function of luminosity) to study the build-up of the local black hole mass function, although [64] also considered models where the average accretion rate relative to Eddington falls off toward lower z and higher $M_{\rm BH}$. Raimundo and Fabian [98] employed a variation on the constant $L/L_{\rm Edd}$ models, assuming three different populations of AGN with their own Eddington ratio: a population of obscured low $L/L_{\rm Edd}$ AGN, a population of obscured AGN with higher $L/L_{\rm Edd}$, and a population of unobscured AGN. Yu and Lu [62] modeled the quasar lightcurve as radiating at the Eddington limit for a period of time, and then transitioning

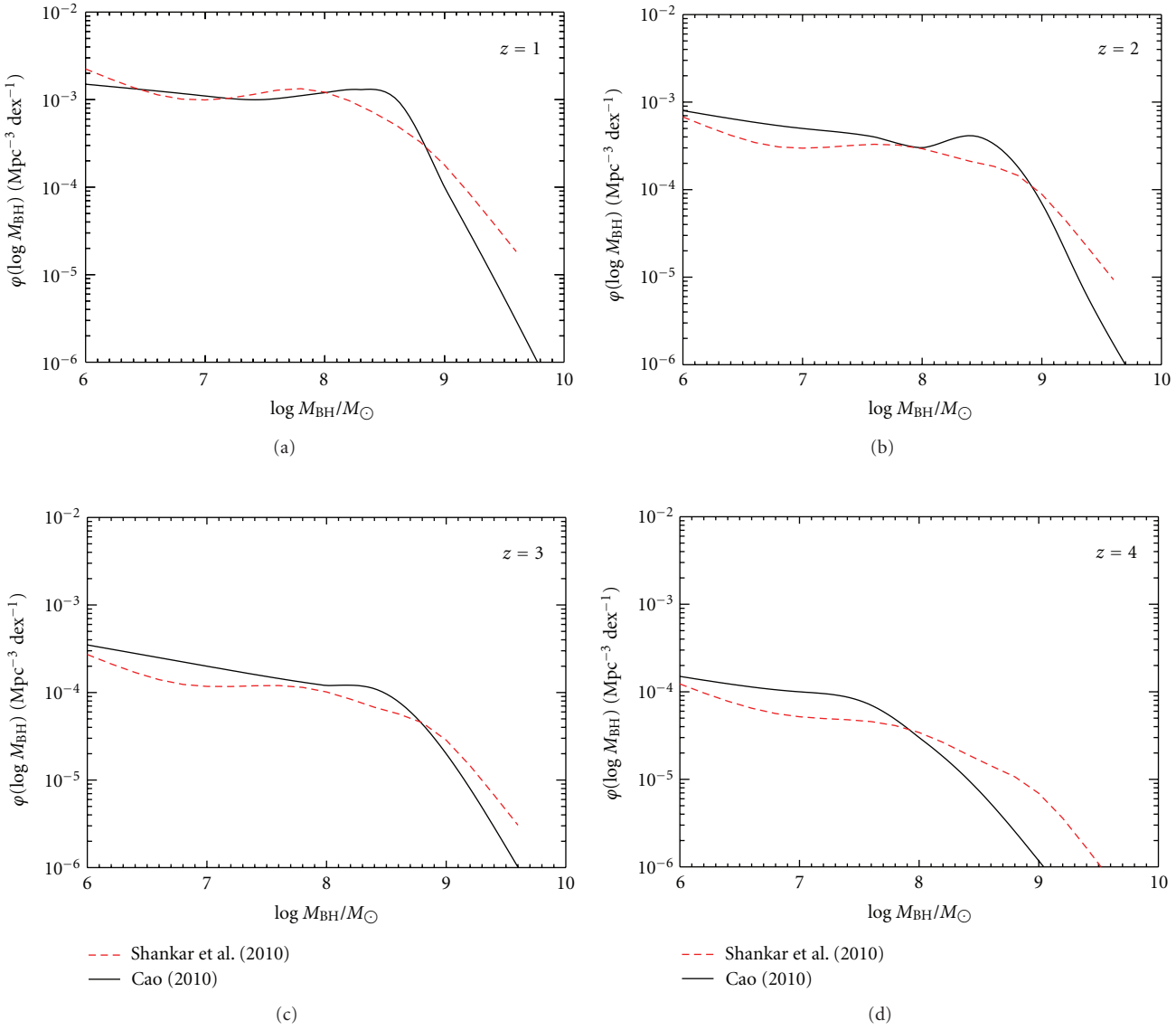


FIGURE 3: Comparison of two recently estimated BHMFs, calculated by Shankar et al. [64] (red dashed line) with that calculated by Cao [92] (solid black line). Both BHMFs were estimated by assuming a quasar lightcurve, where Shankar et al. [64] used a step function model while Cao [92] assumed a power-law decay. Despite the two different models, the BHMFs are similar at a variety of redshifts, except at possibly the high-mass end.

into a power-law decay. Cao [92] also modeled the quasar lightcurve as undergoing a power-law decay. Lightcurves undergoing a power-law decay arise from self-regulation models and describe the evolution of the lightcurve after black hole feedback unbinds the accreting gas, therefore quenching its fuel supply. The power-law decay occurs either as a result of evolution of a blast wave [36, 99] or from viscous evolution of the accretion disk [100, 101].

In Figure 3, we compare the BHMF calculated by Shankar et al.[64] with that calculated by Cao [92]. For Shankar et al. [64], we show their reference model, which assumes a radiative efficiency of $\epsilon_r = 0.065$, an accretion rate relative to Eddington of $\dot{M}/\dot{M}_{\rm Edd} = 0.6$, and that half of all SMBHs are active at $z = 6$. We show the model from Cao [92] which assumes a radiative efficiency of $\epsilon_r = 0.11$ and a quasar lifetime of 7.5×10^8 yr, as it better matches the Shankar et al. [64] estimates. The two estimates of the BHMF agree fairly well, despite the different quasar lightcurve models.

In general, most of the studies that have used (6) in combination with an assumed quasar lightcurve have concluded the following.

(i) Most SMBH growth occurs in periods when the quasar is radiating near the Eddington limit.

(ii) Most, if not all, of the local black hole mass function can be explained as the relic of previous AGN activity, implying that mergers of SMBHs are not important for building up the local mass function.

(iii) SMBH growth is antihierarchical, with the most massive black holes growing first. This has also been termed "downsizing" of active SMBHs.

(iv) The lifetime of AGN activity is $\sim$ a few $\times 10^8$ yr.

(v) Most SMBHs have nonzero spin, as implied by inferred radiative efficiencies of $\epsilon_r \gtrsim 0.06$.

However, while (6) has proven to be an important tool for studying SMBH growth and estimating the black hole mass function, it must be kept in mind that the use of (6) often entails some strong assumptions. These methods rely on the assumed form of the quasar lightcurve, distribution of radiative efficiencies, and bolometric corrections, all of which are subject to considerable uncertainty. Moreover, in general, these methods also rely on an estimate of the local black hole mass function, which, as discussed in Section 3, is itself subject to considerable uncertainty. Indeed, there is a strong degeneracy between the estimated radiative efficiency of accretion and the normalization of the local BHMF, and therefore the uncertainty in ϵ_r is linearly proportional to that in the normalization (or integral) of the local BHMF. All of these issues have the potential to introduce systematic error into methods based on (6), and further work is needed in reducing these systematics.

4.2. Methods That Include the Distribution of Active Supermassive Black Holes. An alternative to the methods described in Section 4.1 is to estimate the average value of the accretion rate onto SMBHs directly from the observational data. This avoids the issue of assuming a form for the quasar lightcurve, as instead $\langle L(M_{\rm BH}, t)\rangle$ is derived directly from the estimated distribution of $L/L_{\rm Edd}$. Techniques based on this approach require a means of linking the mass function of active SMBHs to observational quantities, which is done via scaling relationships. This was the approach of Merloni [89] and Merloni and Heinz [63], who employed the black hole "fundamental plane" (BHFP) [40, 41].

The BHFP is a scaling relationship between $M_{\rm BH}$, radio luminosity, and X-ray luminosity, that exists for low-accretion rate black holes (i.e., $\dot{M}/\dot{M}_{\rm Edd} \lesssim 0.01$), extending from galactic black holes to supermassive ones. The BHFP likely reflects the connection between $M_{\rm BH}$ and the conversion of the accretion flow into radiative energy and jet power. It, in principle, enables one to connect the radio and X-ray luminosity functions to the mass function of active SMBHs. Having obtained a distribution of $M_{\rm BH}$ and X-ray luminosity at a given redshift for the active SMBH population, Merloni [89] and Merloni and Heinz [63] then convert this to a joint distribution of $M_{\rm BH}$ and $\dot{M}_{\rm acc}$ assuming a conversion from X-ray luminosity to $\dot{M}_{\rm acc}$ which depends on the Eddington ratio. The joint distribution of $M_{\rm BH}$ and $\dot{M}_{\rm acc}$ at a given redshift for active SMBHs therefore enables calculation of the average growth rate $\langle \dot{M}(M_{\rm BH}, t(z))\rangle$, which can then be combined with the continuity equation to calculate the black hole mass function at the next redshift. Their estimated BHMF is shown in Figure 4, which is a recreation of their Figure 5. Similar to methods based on assuming a quasar lightcurve, Merloni and Heinz [63] concluded that SMBHs grow antihierarchically; however, in contrast to the lightcurve methods, Merloni and Heinz [63] concluded that most SMBHs have low spin as inferred from their derived radiative efficiency. In addition, Merloni and Heinz [63] concluded that the distribution of SMBH accretion rates is broad and that most SMBH growth occurs during a radiatively efficient accretion mode.

The method of estimating the BHMF from the BHFP developed by Merloni [89] and Merloni and Heinz [63] has the advantage that it derives the distribution of accretion rates empirically. However, there are also disadvantages to this approach. The uncertainties regarding the bolometeric correction, estimation of the local BHMF, and radiative efficiency also apply to the BHFP method as well. Moreover, as discussed in Merloni and Heinz [63], the BHFP is only defined for low-accretion rate objects, that is, objects with $L/L_{\rm Edd} \lesssim 10^{-2}$. Merloni and Heinz [63] extrapolate the BHFP to higher accretion rates, after rescaling the normalization to ensure that the radio luminosity is weak for AGN in the radiatively efficient mode (those objects with $L/L_{\rm Edd} \gtrsim 10^{-2}$ and lacking a jet). Unfortunately, the AGN in the radiatively efficient mode make a significant contribution to the X-ray luminosity function, from which $\langle \dot{M}(M_{\rm BH}, t)\rangle$ is derived. Moreover, most studies, including those based on the BHFP, have concluded that most SMBH growth occurs at $L/L_{\rm Edd} \gtrsim 10^{-2}$, which corresponds to the radiatively efficient mode. Because the radiatively efficient mode also corresponds to the regime of largest systematic uncertainty for the BHFP, there is the potential for significant systematic error in estimating the BHMF based on the BHFP, as well as in estimating the primary mode of SMBH growth. There is thus a need for further improvement to our understanding of the scaling relationships involving $M_{\rm BH}$ and the AGN SED.

5. Black Hole Mass Functions of AGN

Thus far, we have focused on methods for estimating the mass function of all SMBHs. In this section, we will describe methods for estimating the BHMF for those SMBHs in AGN and the results that have come from the application of these methods.

5.1. Methods Based on Scaling Relationships Involving the Broad Emission Lines. The steady improvement in reverberation mapping of AGN [102, 103] has revealed a correlation between the luminosity of AGN and the broad-line region radius [104, 105]. It is therefore possible, in principle, to obtain an estimate of $M_{\rm BH}$ for broad-line AGN (BLAGN) by combining a luminosity-based estimate of the broad-line region size with an estimate of the velocity dispersion of the broad-line region gas obtained from the width of the broad emission lines [47]. These virial mass estimates are then calibrated to the estimates of $M_{\rm BH}$ obtained from reverberation mapping, which themselves are calibrated to be consistent with the local $M_{\rm BH}$-σ_* relationship [106, 107]. Currently, calibrations exist for Hα [108], Hβ (e.g., [48]), Mg-II [109–111], and C IV [48]. The statistical scatter in the virial mass estimates is currently estimated to be $\sim$0.4 dex [48],

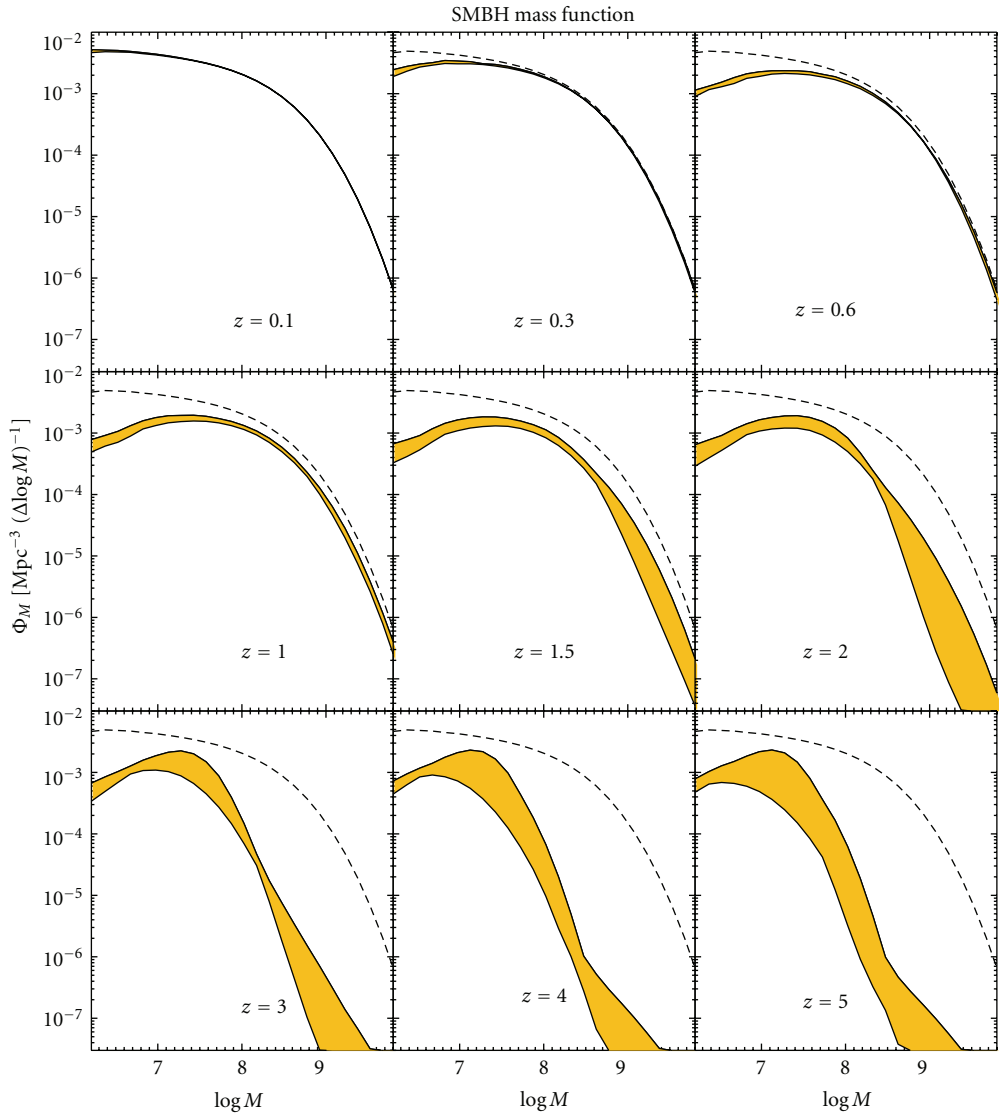


FIGURE 4: Figure 5 from Merloni and Heinz [63], showing the redshift evolution of their estimated BHMF. The dashed line is the local BHMF and the shaded regions reflect the uncertainty in the BHMF that is due to uncertainties in the AGN luminosity function. The high mass end of their estimated BHMF is built up faster than the low-mass end, a phenomenon that has been called "downsizing".

although there are indications that the scatter may be smaller, at least for the most luminous quasars [112–116]. Moreover, it should be noted that the calibration for Mg II is obtained by enforcing consistency in the mean values of the Mg II mass estimator and the Hβ and C IV ones, and therefore there is currently no direct estimate of the statistical scatter in Mg II-based virial mass estimates. In contrast, the amplitudes of the statistical scatter for Hβ and C IV are estimated by comparing mass estimates derived from these lines with the masses derived from reverberation mapping [48]. Although there is currently very little reverberation mapping data for C IV, the estimate of the dispersion in the C IV-based mass estimates should not be biased so long as the masses based on reverberation mapping are reliable estimates of the true $M_{\rm BH}$, regardless of which emission line was used in the reverberation mapping campaign.

Early estimates of the mass function of SMBHs in BLAGN were obtained by binning up the virial mass estimates and applying a $1/V_{\rm max}$ correction [110, 117–119], a technique borrowed from luminosity function estimation. Greene and Ho [118] estimated the local BHMF for BLAGN from the SDSS DR4, while Vestergaard et al. [119] estimated the BHMF for BLAGN over $0.3 < z < 5$ using the uniformly selected quasar sample from the SDSS DR3 [120]. Vestergaard and Osmer [110] estimated the BHMF for the brightest BLAGN using objects from a variety of surveys, as their sample was designed to complement the uniformly selected SDSS DR3 sample. Unfortunately, as discussed in Section 2.1, this method of binning up the mass estimates suffers from biases due to the large statistical scatter in the virial mass estimates, and due to the inability of a luminosity-based $1/V_{\rm max}$ correction to correct for incompleteness in

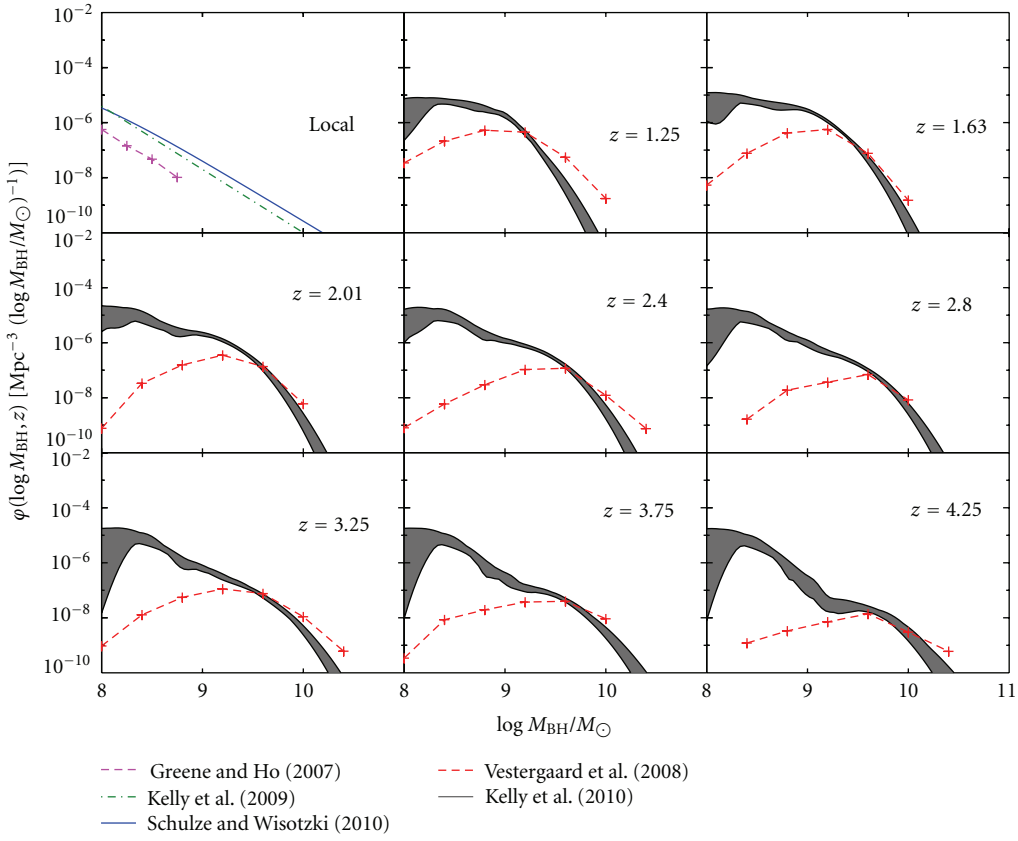


FIGURE 5: Broad-line AGN BHMFs at a variety of redshifts. Shown are the local BHMF estimated by Greene and Ho [118] (dashed magenta line), Kelly et al. [46] (dot-dashed green line), and Schulze and Wisotzki [121] (solid blue line). Also shown are the $z > 1$ BHMFs estimated by Vestergaard et al. [119] (dashed red line) and Kelly et al. [115] (shaded region), where the shaded region for the Kelly et al. [115] estimate defines an approximate 95% confidence region. The BHMFs estimated by Greene and Ho [118] and Vestergaard et al. [119] are flux-limited BHMFs, as they did not fully correct for incompleteness in $M_{\rm BH}$.

$M_{\rm BH}$. Subsequent attempts have further improved in their methodology, providing more accurate BHMFs.

Shen et al. [113] employed a forward-modeling approach where the mass function and Eddington ratio distribution were estimated by matching the observed distribution of mass estimates and luminosity to that implied by the model BHMF and Eddington ratio distribution. Their method accounts for incompleteness and the statistical scatter in the mass estimates but lacked statistical rigor in that the matching was done visually. Schulze and Wisotzki [121] employed a maximum-likelihood technique for estimating the local BHMF for BLAGN. Their method corrects for incompleteness in $M_{\rm BH}$ but does not correct the BHMF for the broadening caused by the statistical scatter in the virial mass estimates. Kelly et al. [46] developed a Bayesian method that corrects for both the statistical scatter in the mass estimates and incompleteness and used their method to estimate the local BHMF of BLAGN from the Bright Quasar Survey [122]. Kelly et al. [115] used the method of [46] to estimate the BHMF of BLAGN at $1 < z < 4.5$ from the mass estimates in the SDSS DR3 quasar sample [119]. The BLAGN BHMFs from a variety of studies are compiled in Figure 5, showing the evolution of the BHMF from the local universe out to $z = 4.5$. More recently, Shen and Kelly [116] extended the Bayesian method of [46] to include a possible luminosity-dependent bias in virial mass estimates derived from the emission line FWHM, the existence of which was suggested by Shen and Kelly [77]. Shen and Kelly [116] applied their method to the SDSS DR7 uniformly-selected quasar sample, independently estimating the BHMF and Eddington ratio distribution in different redshifts bins.

Similar to the methods based on the continuity equation, investigations of the BHMF for BLAGN have found evidence for the anti-hierarchical growth of SMBHs, that is, cosmic "down-sizing" of BLAGN activity. The inferred Eddington ratio distributions are wide, and the density of SMBHs continues to increase toward Eddington ratios which are below the survey completeness limit. In addition, Kelly et al. [115] used the BLAGN BHMF to estimate the lifetime of broad-line quasar activity to be $t_{\rm BL} \sim 150$ Myr among SMBHs with $M_{\rm BH} \sim 10^9 M_\odot$, which is similar to quasar lifetimes inferred from the continuity equation. Kelly et al. [115] also used their estimated BHMF to estimate the maximum mass of a SMBH to be $M_{\rm BH} \sim 3 \times 10^{10} M_\odot$, which is in agreement with theoretical expectations [123, 124].

Mass functions estimated from scaling relationships for BLAGN have the advantage that they are derived from estimates of $M_{\rm BH}$ that are obtained for individual sources, providing a more "direct" estimate of the mass function than those based on the continuity equation. However, they have the disadvantage that they are only available for a subset of the AGN population, which itself is only a subset of the SMBH population. This complicates comparison with other SMBH mass functions, as the fraction of AGN with broad emission lines is poorly constrained, especially as a function of mass. This being said, BHMFs of BLAGN represent a subset of SMBHs that are actively growing at the time that they are observed, and, as the aforementioned studies have demonstrated, their mass function still contains important information on SMBH growth.

As with all methods of BHMF estimation, the virial mass estimates and the mass functions derived from them still suffer from systematics. First, there is the usual problem of calculating a bolometeric correction, although this only affects the estimated Eddington ratio distribution and not the BHMF. Second, there are a few concerns with the virial mass estimates which could introduce systematic error; some of these have been discussed by Greene and Ho [125]. For one, most of the reverberation mapping data is only available for the Hβ line. Because of the limited Mg II data, the Mg II scaling relationship is in general not calibrated using objects with black hole mass estimates from reverberation mapping. There may be systematic effects with luminosity or Eddington ratio when using the FWHM-based scaling relationships [116, 126], possibly due to a dependence of the broad-line region structure on these quantities. Systematic effects on broad-line region geometry, which can effect the inferred velocity dispersion, are a particular concern for C IV, which is thought to arise in an accretion disk wind [127]. Along these lines, unaccounting for radiation pressure on broad-line clouds may also bias the virial masses, especially among those AGN radiating near the Eddington limit [128]; however, its importance is still debated [129–131]. In addition, the reliability of line width measurements can rapidly deteriorate for low S/N data [132]. And, finally, the BLAGN virial mass estimates are calibrated to the reverberation mapping derived masses, which themselves are calibrated to lie on the local $M_{\rm BH}$-σ_* relationship. Most of the AGN that are used to calibrate the reverberation mapping masses to the $M_{\rm BH}$-σ_* relationship have lower masses and are hosted by late-type galaxies, for which there is evidence that the $M_{\rm BH}$-σ_* relationship begins to break down [81]. Greene et al. [81] argue that the normalization of the scaling relationships inferred when limiting the calibration to low-mass SMBHs hosted in late-type galaxies may be about a factor of $\sim$1.5 lower than that used for the current broad-line mass estimates [81]. However, dynamical mass estimates exist for two reverberation mapped AGN: NGC 3227 [133, 134] and NGC 4151 [134, 135]. In both cases, the masses derived from dynamical modeling and reverberation mapping agree, so it is unclear if a smaller scaling factor is needed for late-type galaxies. These issues show that there are still many remaining questions regarding virial masses, highlighting the need for further study using high-quality reverberation mapping data.

5.2. Other Methods for Estimating the Black Hole Mass Function of AGN. Before broad-line mass estimates, there were two earlier attempts at estimating the BHMF for AGN, which we briefly mention here. Siemiginowska and Elvis [136] and Hatziminaoglou et al. [137] used a model for the AGN lightcurve arising due to thermal-viscous accretion disk instabilities [138] to calculate the expected distribution of luminosity at a given black hole mass. Based on this calculated distribution, they used the quasar luminosity function to constrain the quasar black hole mass function. Siemiginowska and Elvis [136] found evidence for SMBH downsizing in AGN, consistent with later work.

Franceschini et al. [139] found a tight correlation between $M_{\rm BH}$ and the total radio power observed in a sample of local galaxies. They then used their empirical relationship to estimate the local BHMF derived from the local radio luminosity function of galaxies. While many of the objects in their sample are not considered AGN in the traditional sense, Fraceschini et al. [139] argue that this correlation is a signature of an advection-dominated accretion flow, thought to dominate at low accretion rates relative to Eddington. Therefore, while these SMBHs may not be "active" in the quasar sense, the determination of their mass function relies on radio emission from the SMBH accretion flow, so this method may still be considered a method for estimating the BHMF for active SMBHs. Franceschini et al. [139] compared their BHMF to models of AGN activity and found that it was inconsistent with AGN activity being continuous and long lived, but consistent with AGN activity being transient and possibly recurrent.

6. Theoretical Models for Black Hole Mass Functions across Cosmic Time

There have been numerous theoretical models for the formation and growth of supermassive black holes, and coevolution with their host galaxies. Understanding this formation, growth, and coevolution is one of the current most important outstanding issues in extragalactic astrophysics. Because the black hole mass function provides a census of the SMBH population and its evolution, it is one of the most fundamental observational quantities available for constraining models of SMBH formation and growth. As such, many theoretical investigations have predicted a BMHF for comparison with the empirical BHMF. In this section, we review some of the models for SMBH formation and growth. There have been numerous theoretical models for SMBH growth and formation, and it is beyond the scope of this primarily empirically-focused review to review all of them; instead, we focus on those theoretical models that predict a BHMF.

6.1. Modeling the Coevolution of SMBHs and Galaxies: Predicted BHMFs. Early models for the coevolution of SMHBs and galaxies linked the growth of black holes to the properties of host dark matter halos, with periods of SMBH growth occurring in quasar phases initiated by mergers. In general, early studies that predicted a BHMF used various prescriptions to relate $M_{\rm BH}$ to the mass of the host halo [26, 140–143]. More recent models for the coevolution of SMBHs

and galaxies has incorporated AGN feedback from the SMBH. In addition, the availability of empirical BHMFs have enabled modelers to compare their more recent models with observational data. In general, the models are qualitatively in agreement with the empirical results, in that they are able to match the local BHMF fairly well and predict downsizing of SMBHs. However, considering the current systematic and statistical uncertainties in the empirical results, it is difficult to place rigorous empirical constraints on the models such that certain models may be ruled out. Because of this, we simply summarize some of the different recent models that have been developed which predict the BHMF.

Granato et al. [144] developed a model incorporating feedback from AGN and supernovae, where the feeding of the SMBH is driven by stellar radiation drag on gas. Their predicted local BHMF agrees with that estimated by Shankar et al. [57]. Cattaneo et al. [30] used halo merger trees constructed from N-body simulations to track the growth of SMBHs. In their model, the black hole fueling rate was proportional to the star formation rate of the host galaxy burst component and the density of the cold gas in the starburst component. Their model predicted SMBH downsizing, with the most massive part of the BHMF being built up first, in agreement with the subsequent empirical studies.

Hopkins et al. [145] describe a model for the coevolution of SMBHs and galaxies whereby all major mergers of gas-rich galaxies trigger a quasar. In this model, the final black hole mass is assumed to be on average proportional to the host spheroidal mass, in agreement with the local scaling relationships between SMBHs and their host galaxies. Hopkins et al. [145] estimated the merger rate of gas-rich galaxies by combining theoretical constraints of the halo and subhalo mass functions with empirical constraints on halo occupation models. Their model also predicts SMBH downsizing, and their predicted BHMF matches the local BHMF derived by Marconi et al. [53]. Similarly, Shen [146] also assumed that quasars are triggered by major mergers of gas-rich galaxies, with the SMBHs growing via accretion in these quasar phases. Shen [146] used a halo merger rate based on theoretical expectations from N-body simulations and assumed a universal quasar lightcurve shape having an exponential increase followed by a power law decay (see also [62]). The BHMF predicted by Shen [146] broadly agrees with the local one estimated by Shankar et al. [64] and predicts that most SMBHs with $M_{\rm BH} > 3 \times 10^8 M_\odot$ were in place by $z = 1$ but only 50% of them were assembled by $z = 2$.

Most recently, Fanidakis et al. [147, 148] extended the model of [23], which includes AGN feedback, to also follow the spin distribution of SMBHs. In their model, SMBHs are fueled through accretion of cold gas from mergers, disk instabilities, and cooling flows from hot halos. However, the inclusion of SMBH spin enabled them to include different radiative efficiencies, which dictates how much accreted material actually grows the black hole, and to provide an improved model for the amount of mechanical feedback imparted through an AGN jet, both of which depend on the spin of the black hole. Their model predicts that the present-day universe is dominated by SMBHs with $M_{\rm BH} \sim 10^7$–$10^8 M_\odot$, and that the BHMF at $M_{\rm BH} > 10^9 M_\odot$ was largely built up at $z < 2$ due to an increase in both lower-accretion rate "radio-mode" growth and mergers of SMBHs.

Almost all models for the cosmological coevolution of SMBHs and galaxies that predict a BHMF have been of an analytical or semi-analytical nature. An exception is the study done by Di Matteo et al. [31], who present the results from cosmological hydrodynamic simulations of the ΛCDM model that follow the growth of galaxies and SMBHs, including their feedback processes, at $z > 1$. Direct cosmological simulations such as these should, in principle, provide the most accurate results as to the predicted BHMF, and for identifying the relevant physical processes that are important in shaping the BHMF. However, current cosmological hydrodynamic simulations suffer from the fact that they cannot resolve processes on physical scales corresponding to the SMBH accretion flow. In fact, Di Matteo et al. [31] use a gravitational softening length of $\epsilon = 2.73\,{\rm h}^{-1}$ kpc. Instead, Di Matteo et al. [31] employ a subresolution model where the accretion onto the SMBH is estimated using a Bondi-Hoyle-Lyttleton parameterization [149–151] with a correction factor to account for the fact that the Bondi radius is not resolved. They assume a radiative feedback energy efficiency of 5% [20], which is the only free parameter in their model and required in order to match the normalization of the observed local $M_{\rm BH}$-σ_* relationship. Their calculated BHMF at $z = 1$ matches the local BHMF for $M_{\rm BH} > 2 \times 10^8 M_\odot$. In addition, Di Matteo et al. [31] also find downsizing in their model, in agreement with observations, with the high-mass end of the BHMF being largely in place by $z \sim 2$.

In Figure 6, we compile predicted BHMFs from several recent models for SMBH formation and growth [31, 145–147, 152]. In general, the models tend to agree to within a factor of a few with regards to the BHMF. However, they diverge at $M_{\rm BH} \gtrsim 10^9 M_\odot$, where their predicted SMBH number densities can differ by over an order of magnitude.

6.2. Modeling the BHMF of SMBH Seeds. Recent work has made improvement to models for the BHMF by focusing on theoretical modeling of the distribution of seed SMBHs. The discovery of quasars at $z \approx 6$–7 with $M_{\rm BH} \sim 10^9 M_\odot$ [153, 154] places strong constraints on the formation of SMBH seeds due to the very limited amount of time available at that redshift to grow SMBHs (e.g., [155]). Lodato and Natarajan [156] derive the BHMF of SMBH seeds at $z \sim 15$ that are the result of the collapse of pregalactic disks which have not yet been enriched by metals [157]. Black holes formed through such a mechanism have masses $M_{\rm BH} \sim 10^5 M_\odot$, while black holes which are the remnants of Pop III stars have $M_{\rm BH} \sim 10^3 M_\odot$. A similar seed black hole formation mechanism is through "quasistars" [158–160], which are also able to produce seed black holes with $M_{\rm BH} \sim 10^5 M_\odot$.

Volonteri et al. [161] describe a model for the growth of SMBHs seeded according to the direct collapse model of Lodato and Natarajan [157] with varying formation efficiencies. In addition, they also compared the results from this model using SMBHs seeded from Pop III remnants. Volonteri et al. [161] grow SMBHs through major mergers, and force the black hole mass after the galaxy merger to scale with the circular velocity of the host halo; additional growth is also

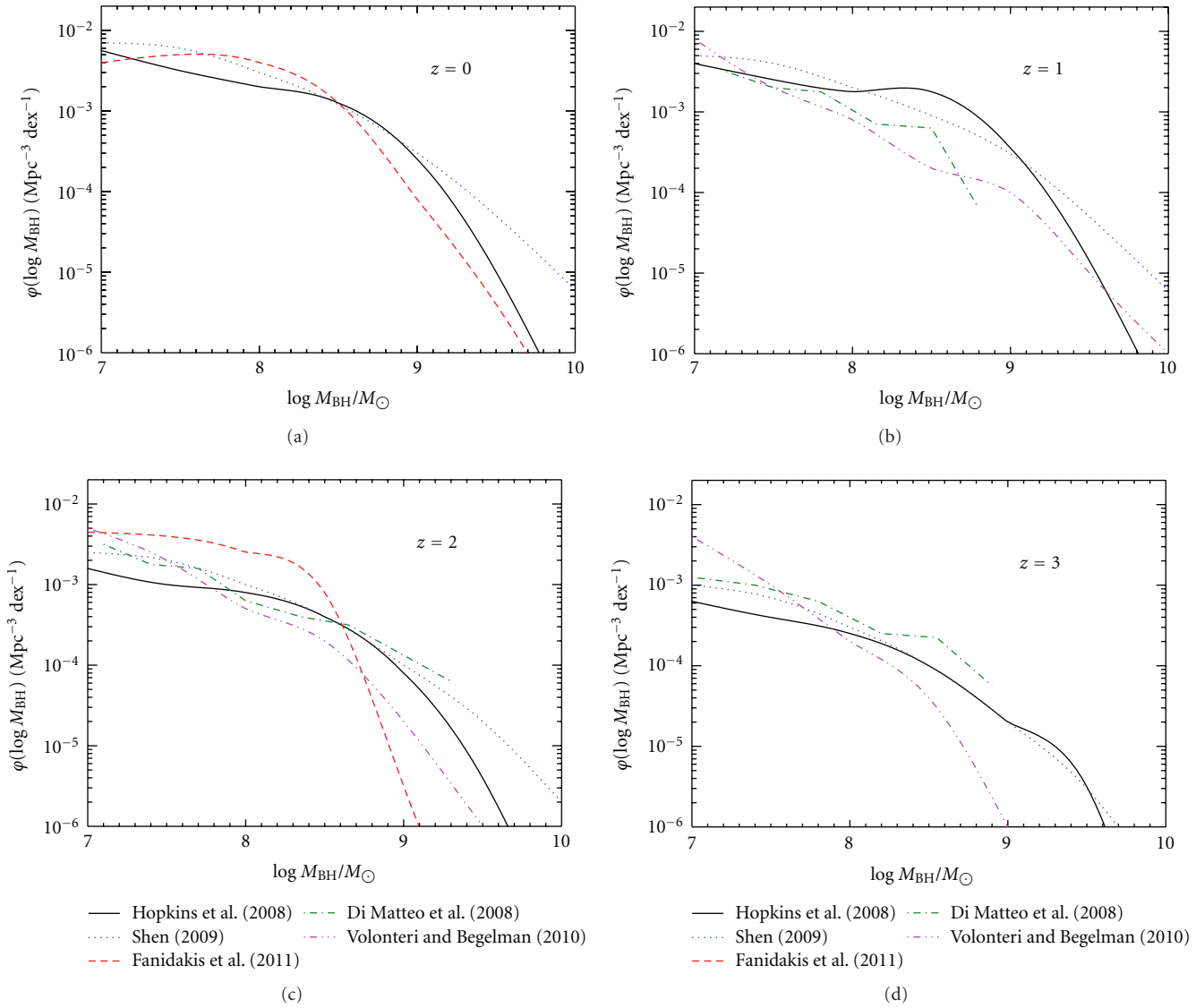


FIGURE 6: Compilation of BHMFs predicted by several recent models for SMBH formation and growth. Shown are the BHMFs predicted by Hopkins et al. [145] (solid black line), Shen [146] (dotted blue line), Fanidakis et al. [147] (dashed red line), Di Matteo et al. [31] (dot-dashed green line), and Volonteri and Begelman [152] (dashed-dotted-dotted-dotted magenta line). In general the number densities predicted by the models agree to within a factor of a few, although they diverge at $M_{\rm BH} \gtrsim 10^9 M_\odot$. Some of these authors did not report BHMFs at each redshift shown, so we only show those available at each redshift.

provided through black hole mergers. Their merger trees are based on a Monte Carlo algorithm based on the extended Press-Schechter formalism. They find that most significant differences in the local BHMF with respect to black hole formation efficiency occur at $M_{\rm BH} < 10^7 M_\odot$, with the number density of SMBHs with $M_{\rm BH} < 10^7 M_\odot$ increasing with increasing formation efficiency. Volonteri and Begelman [152] performed a similar analysis as that of Volonteri et al. [161] but instead used SMBH seeds formed via quasistars. The BHMFs calculated by Volonteri and Begelman [152] match those of Merloni and Heinz [63] at the high-mass end, at least at $z > 2$.

Natarajan and Volonteri [162] used a growth and seeding model which is very similar to that employed by Volonteri et al. [161]. However, they also predict the BHMF for broad line quasars, assuming that 20% of quasars are unobscured. They compare the BHMF derived from their model at $1 < z < 4.5$ to the BHMF for broad line quasars reported by Kelly et al. [115], and to the BHMF for all SMBHs reported by Merloni and Heinz [63]. Natarajan and Volonteri [162] concluded that seeds from Pop III stars have difficulty reproducing the BLAGN BHMF, especially at high redshift, while seeds resulting from the direct collapse of pregalactic disks do better at fitting the high mass end of the BLAGN BHMF at $z > 2$.

Lippai et al. [163] studied the impact that the fraction of halos that form a SMBH seed at $z > 6$ has on the local BHMF. They are able to reproduce the observed quasar luminosity

function and local BHMF for a suitable range of radiative efficiency and quasar lifetime, so long as at least ≈10% of high-z halos contained SMBH seeds. Tanaka and Haiman [164] present BHMFs at $z = 6$ based on a comprehensive exploration of the parameter space governing the buildup of the $z > 6$ BHMF, including models for SMBH seeds that form from Pop III remnants and through direct collapse, variations in the occupation fraction, and a detailed treatment of gravitational recoil. Comparing their predicted BHMFs to observations of $z \sim 6$ quasars from the SDSS, they concluded that $\sim 100 M_\odot$ seeds can grow into the SMBHs observed at $z \sim 6$, so long as they are nearly continuously embedded in dense gas, form at $z > 30$, have low occupation fractions, and stop forming by $z \sim 20$.

7. Directions for Future Work and Improvement

Before concluding this paper, we present a discussion of possible future empirical and theoretical work relevant to BHMF studies. These include the following.

7.1. Better Characterization of the SMBH-Host Galaxy Scaling Relationships. Currently, the local BHMF is estimated from the distribution of host galaxy properties assuming that $M_{\rm BH}$ has a constant log-normal scatter about a single-power law scaling relationship. As discussed in Section 3, recent observations have provided reason to doubt this assumption, suggesting that the correlations break down at the highest and lowest masses. This will create biases in the BHMF determined from the scaling relationships, which in turn will also affect the BHMF estimated from the continuity equation. Further direct $M_{\rm BH}$ estimates from dynamical and kinematic modeling should be obtained for a variety of galaxy types, especially at the high- and low-$M_{\rm BH}$ end. The next class of 25+ m telescopes should provide a significantly improved picture of the scaling relationships, thus providing us with more accurate estimates of the local BHMF.

7.2. Improvements to Techniques Based on the Continuity Equation. Most studies that have invoked the continuity equation to link the local BHMF to the AGN luminosity function have assumed a single radiative efficiency, which is equivalent to assuming a single black hole spin, and a universal AGN lightcurve. Neither of these assumptions are likely to be true, and improvements to this type of modeling should include a distribution of SMBH spin and AGN lightcurves. In addition, we need to better characterize the bolometeric corrections, which remain a significant source of systematic uncertainty. The continuity equation techniques should also be extended to map the evolution of the full joint 3-dimensional distribution of black hole mass, accretion rate, and spin. While this will not necessarily have a direct effect on estimating the BHMF, it will provide insight into the dominant accretion modes experienced by active SMBH and into the dominant fueling mechanism for AGN activity, as the spin distribution traces the SMBH fueling history [165].

7.3. Better Characterization of Scaling Relationships Involving $M_{\rm BH}$ for AGN. The dominant scaling relationship for estimating $M_{\rm BH}$ in AGN involves the broad emission lines. However, as discussed in Section 5.1, a number of systematic uncertainties remain. In order to reduce these systematics, we need to better understand the broad-line region geometry for the different emission lines and how it scales with luminosity, which will provide us with a more accurate conversion from line width to velocity dispersion. Moreover, accurate characterization of the broad-line region geometry should remove the need to calibrate the scaling relationships to the local $M_{\rm BH}$-σ_* relationship, which has its own set of systematics. Improvements to reverberation mapping campaigns and modeling [166–170], as well as increasing the number of AGN monitored for reverberation mapping, will be needed in order to really understand the systematics involved in the broad line mass estimates.

There is also the need to better characterize the black hole fundamental plane. Because the BHFP describes how the emission mechanisms responsible for the radio and X-ray flux scale with $M_{\rm BH}$, the BHFP coefficients depend on these emission mechanisms. However, these emission mechanisms depend on the geometry of the accretion state and the existence of a jet, which in turn depends on the accretion rate [171], and therefore the BHFP coefficients will be different for different classes of AGN [172–174]. In particular, the BHFP is currently poorly constrained for "soft-state" galactic black holes and radio-quiet AGN. Therefore, to reduce the systematics involved with the BHFP, it will be necessary to characterize the scaling relationships and their scatter for radio-quiet objects. A correlation between the radio and X-ray luminosity has been observed for radio-quiet objects [175], implying that a BHFP should also exist for these objects. In order to better characterize the BHFP for radio-quiet objects it will be necessary to obtain radio detections for a well-defined sample of radio-quiet AGN with reliable $M_{\rm BH}$ estimates and X-ray detections.

Finally, there has recently been the discovery of scaling relationships involving $M_{\rm BH}$ and the optical [176, 177] and X-ray [42, 45, 49] variability properties of AGN. Mass estimators based on these scaling relationships have not been rigorously developed yet, nor have they seen widespread use. However, the existence of these scaling relationships implies that the variability properties may offer another avenue for estimating $M_{\rm BH}$ and BHMFs, which may become increasingly valuable in the era of current and future large time-domain surveys, such as *Pan-Starrs* and LSST.

7.4. Understanding the Redshift Evolution of Scaling Relations. From the theoretical point of view, it is clear that high-redshift scaling relations (or the lack thereof) between SMBH and their hosts provide unique and powerful constraints to models for AGN feeding and feedback, which cannot be otherwise distinguished (see, e.g., Merloni et al. [74] and references therein).

In practical terms, a better understanding of the evolution of scaling relations may also be very advantageous for BHMF studies. As we discussed above, current technique for BH mass estimation at $z > 0$ involve unobscured, broad

emission line QSOs. One can argue that, as long as we are restricted to just this class of QSOs, we will have to make critical assumptions about the properties of a significant part of the population to draw conclusions about the full BHMF (if we wanted, e.g., to compare with "continuity-equation-based" methods). On the other hand, large multiwavelength surveys do and will provide a wealth of information on the host galaxies of obscured AGN at high redshift, that represent the numerically dominant part of the growing black holes population [178–181]. Therefore, if we had an independent way to put constraints on the nature of the BH-host relation for these objects, we could explore the uncharted territory of BHMF for obscured AGN (and for the entire population). Such an independent information could come, for example, from IR studies of broad emission lines which could act as probes of the BH potential less affected by obscuration. The first exploratory works pursuing this line of research have recently been published, for example, [182, 183].

From the technical point of view, a lot of work of course is needed to better understand how reliable these estimators are. Another big "technical" challenge of all studies of the evolution of scaling relations is the fact that they require a thorough assessment of the many observational biases one encounters in studying high redshift AGN and their hosts [184].

7.5. Accounting for Black Hole Kicks in Theoretical Models. Many theoretical models for the BHMF do not include recoiling effects caused by the merger of two black holes. However, recent theoretical work on black hole recoils suggests that black holes can spend a significantly large enough amount of time offset from the central region of the host galaxy to alter their growth, thereby increasing the scatter about the scaling relationships and decreasing the final black hole mass [185–188]. On the other hand, Volonteri et al. [189] and Volonteri et al. [190] find that ejected SMBHs are rare at $z < 5$, especially for massive SMBHs, suggesting that accounting for ejected black holes will not make a significant difference in the BHMF. Moreover, Tanaka and Haiman [164] find that recoiling black holes only modestly effect the BHMF. Further improvement in our understanding of the effects and frequency of black hole recoil will ensure the accurate implementation of black hole recoil into models for SMBH growth.

7.6. Improvements to Our Understanding of AGN Feedback. Most current theoretical models for SMBH growth involve AGN feedback and assume a single efficiency for coupling feedback energy to the gas; this feedback efficiency is usually treated as a free parameter. An improved physical understanding of AGN feedback will improve theoretical models for the BHMF, as the feedback efficiency affects the dynamics of the SMBH's fuel supply and therefore the amount that the SMBH accretes as a function of redshift. Recent high-resolution hydrodynamic simulations in one dimension [191–193] and two dimensions [194, 195] have concluded that AGN feedback efficiency increases with the Eddington ratio and that the values are below the value of ~5% assumed in many current theoretical models for SMBH growth. Further improvements to simulations developed for studying AGN feedback will lead to a better physical understanding of AGN feedback, which will improve theoretical models for SMBH growth and the BHMF.

On the observational side, future X-ray observations should provide considerable improvement in our understanding of AGN feedback. X-ray spectra are needed in order to determine the total column density of the gas, and thus its kinetic energy flux, which can be compared to the energy output of the SMBH. Current X-ray observations from *Chandra* have found evidence that AGN feedback exists in the local universe [196]. However, X-ray calorimeters on future X-ray satellites will be needed for further improvement as they provide the high throughput and spectral resolution needed to measure column densities and velocities of ionized gas, and consequently the kinetic energy flux, in a large sample of AGN across a broad redshift range.

7.7. Improvements in Resolution and Subresolution Modeling for Direct Hydrodynamic Simulations. Full hydrodynamic cosmological simulations offer the most promising avenue for providing a physically motivated BHMF without free parameters and for unambiguously identifying the relevant physical processes in building up the BHMF. However, they currently cannot resolve scales relevant to the accretion flow onto the SMBH. Numerical codes based on *adaptive mesh refinement* techniques will provide improvement in resolution, but it will likely be a while before hydrodynamic cosmological simulations are able to follow SMBH growth in large cosmological volumes while simultaneously resolving the scales relevant for individual black holes. In the meantime, further improvement can be made to the subresolution modeling employed by current hydrodynamic simulations.

One way of improving current sub-resolution models may be to implement the results on AGN feedback based on the type of work described in the previous bullet point. Another improvement is in modifying the sub-resolution model for the SMBH accretion flow. Current methods assume the Bondi rate combined with a correction factor to account for the fact that the temperature and density of the gas are not resolved at the Bondi radius. Not surprisingly, the growth of the SMBH is sensitive to how this correction factor is modeled [33]. Moreover, sub-resolution models based on the Bondi rate neglect the angular momentum of the gas, and thus the Bondi rate may not be representative of the actual accretion rate onto the SMBH. Hopkins and Quataert [197] used high-resolution hydrodynamic simulations to conclude that the Bondi rate was a poor estimate of the actual accretion rate onto the SMBH, and describe a sub-resolution model which accurately estimated the actual accretion rate in their simulations. In addition, Power et al. [198] suggest an alternative sub-resolution model based on an "accretion particle" to provide a more accurate estimate of the black hole accretion rate. The implementation of improved sub-resolution models for accretion rate and feedback into cosmological hydrodynamic simulations, as well as further improvements to the sub-resolution models, will result in more accurate predicted BHMFs, allowing a more insightful comparison with empirical BHMFs.

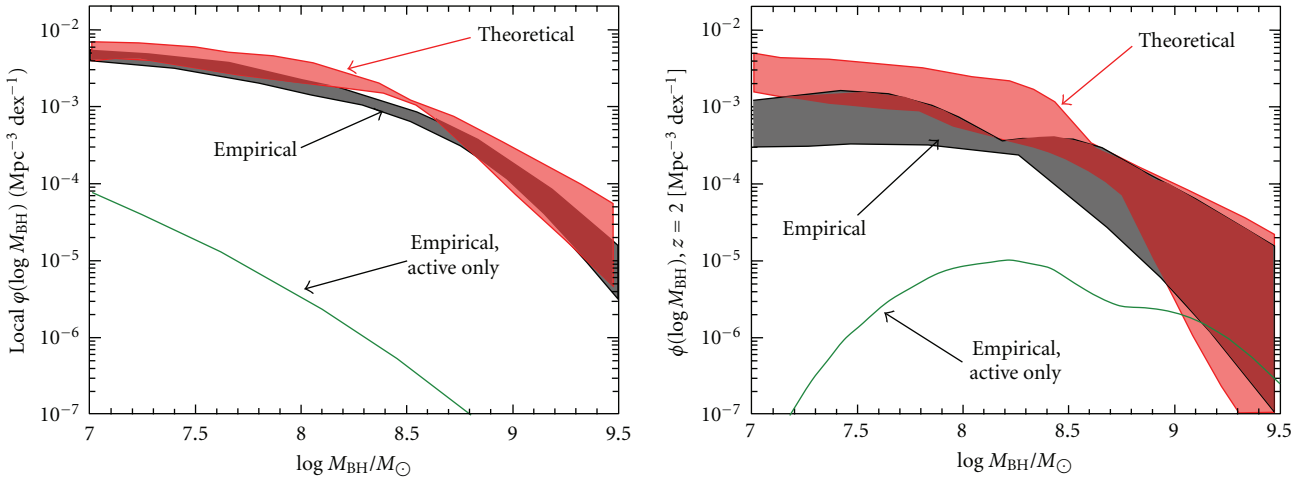


Figure 7: Comparison of empirical estimates of the BHMF (grey shaded region) with BHMFs predicted by theoretical models for SMBH formation and growth (red shaded region), in both the local universe (a) and at $z = 2$ (b). Also shown are the estimated BHMFs of broad-line AGN only (solid green line), from Schulze and Wisotzki [121] (a) and Kelly et al. [115] (b). The empirical estimates of the BHMF are those shown in Figures 3 and 4, while the theoretical estimates are those shown in Figure 6. The shaded regions define the spread in the estimates and models and may be considered to be a crude estimate of their uncertainty. In general, the theoretical number densities are consistent with the empirical ones to within a factor of a few.

In this paper we have reviewed current estimates of the SMBH mass function, as well as theoretical models for the BHMF. As discussed above, each of the methods for estimating the BHMF has their own set of systematics. In Figure 7, we compare the empirical estimates of the local BHMF (defined by the shaded region in Figure 2) with the BHMFs predicted by the theoretical models compiled in Figure 6. In addition, in Figure 7, we also compare the empirical BHMFs at $z = 2$, as estimated using the lightcurve method (shown in Figure 3) and the black hole fundamental plane (shown in Figure 4), with BHMFs predicted by the theoretical models. In both Figures we also include the BHMFs for broad-line AGN as estimated by Schulze and Wisotzki [121] and Kelly et al. [115]. In general, the theoretical models are consistent to within a factor of a few with the empirical estimates of the BHMF, although there is a large spread in the models and empirical estimates at $z = 2$. Moreover, the estimated number densities of broad-line AGN are significantly lower than those of all SMBHs, suggesting that only a small fraction of SMBHs are active across a broad range in $M_{\rm BH}$, except for possibly SMBHs at $z \sim 2$ with $M_{\rm BH} \gtrsim 10^9 M_\odot$.

Despite the differences in the methods for estimating the BHMF and the theoretical models, they have lead to a number of common conclusions. In particular, the empirical results have presented a picture whereby SMBHs grow primarily via accretion in active phases (Eddington ratios $L_{\rm bol}/L_{\rm Edd} > 0.01$), that quasar activity is a relatively short-lived phenomenon relative to the lifetime of the SMBH and host galaxy (i.e., small "duty cycles" for AGN activity) and that SMBH growth is anti-hierarchical with the most massive end of the BHMF being built up first. These empirical results are qualitatively in agreement with the steadily improving theoretical models.

Acknowledgments

The authors would like to thank Tommaso Treu, Priya Natarajan, Marta Volonteri, Aneta Siemiginowska, Xiaohui Fan, Marianne Vestergaard, Alister Graham, and Zoltán Haiman for helpful comments. In addition, they would like to thank two anonymous referees whose comments improved our paper. B. C. Kelly acknowledges support by NASA through Hubble Fellowship Grant no. HF-51243.01 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under Contract NAS 5-26555.

References

[1] J. Kormendy and D. Richstone, "Inward bound-the search for supermassive black holes in galactic nuclei," *Annual Review of Astronomy and Astrophysics*, vol. 33, no. 1, pp. 581–624, 1995.

[2] R. J. McLure and J. S. Dunlop, "The black hole masses of Seyfert galaxies and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 327, no. 1, pp. 199–207, 2001.

[3] R. J. McLure and J. S. Dunlop, "On the black hole-bulge mass relation in active and inactive galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 331, no. 3, pp. 795–804, 2002.

[4] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.

[5] D. Merritt and L. Ferrarese, "The M$_\bullet$-σ relation for supermassive black holes," *Astrophysical Journal*, vol. 547, no. 1, pp. 140–145, 2001.

[6] S. Tremaine, K. Gebhardt, R. Bender et al., "The slope of the black hole mass versus velocity dispersion correlation," *Astrophysical Journal*, vol. 574, no. 2, pp. 740–753, 2002.

[7] A. W. Graham, P. Erwin, N. Caon, and I. Trujillo, "A correlation between galaxy light concentration and supermassive black hole mass," *Astrophysical Journal*, vol. 563, no. 1, pp. L11–L14, 2001.

[8] A. W. Graham and S. P. Driver, "A log-quadratic relation for predicting supermassive black hole masses from the host bulge sérsic index," *Astrophysical Journal Letters*, vol. 655, no. 1, pp. 77–87, 2007.

[9] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[10] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *Astrophysical Journal Letters*, vol. 589, no. 1, pp. L21–L24, 2003.

[11] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *Astrophysical Journal Letters*, vol. 604, no. 2, pp. L89–L92, 2004.

[12] M. C. Aller and D. O. Richstone, "Host galaxy bulge predictors of supermassive black hole mass," *Astrophysical Journal*, vol. 665, no. 1, pp. 120–156, 2007.

[13] P. F. Hopkins, L. Hernquist, T. J. Cox, B. Robertson, and E. Krause, "A theoretical interpretation of the black hole fundamental plane," *Astrophysical Journal*, vol. 669, no. 1, pp. 45–66, 2007.

[14] J. Silk and M. J. Rees, "Quasars and galaxy formation," *Astronomy and Astrophysics*, vol. 331, no. 2, pp. L1–L4, 1998.

[15] A. C. Fabian, "The obscured growth of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 308, no. 4, pp. L39–L43, 1999.

[16] M. C. Begelman and B. B. Nath, "Self-regulated black hole accretion, the M-σ relation and the growth of bulges in galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 361, no. 4, pp. 1387–1392, 2005.

[17] N. Murray, E. Quataert, and T. A. Thompson, "On the maximum luminosity of galaxies and their central black holes: feedback from momentum-driven winds," *Astrophysical Journal*, vol. 618, no. 2, pp. 569–585, 2005.

[18] C. Y. Peng, "How mergers may affect the mass scaling relation between geuvitationally bound systems," *Astrophysical Journal Letters*, vol. 671, no. 2, pp. 1098–1107, 2007.

[19] K. Jahnke and A. V. Macciò, "The non-causal origin of the black-hole-galaxy scaling relations," *Astrophysical Journal Letters*, vol. 734, no. 2, article 92, 2011.

[20] T. Di Matteo, V. Springel, and L. Ilernquist, "Energy input from quasars regulates the growth and activity of black holes and their host galaxies," *Nature*, vol. 433, no. 7026, pp. 604–607, 2005.

[21] V. Springel, T. Di Matteo, and L. Hernquist, "Black holes in galaxy mergers: the formation of red elliptical galaxies," *Astrophysical Journal*, vol. 620, no. 2, pp. L79–L82, 2005.

[22] P. H. Johansson, T. Naab, and A. Burkert, "Equal- and unequal-mass mergers of disk and elliptical galaxies with black holes," *Astrophysical Journal*, vol. 690, no. 1, pp. 802–821, 2009.

[23] R. G. Bower, A. J. Benson, R. Malbon et al., "Breaking the hierarchy of galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 2, pp. 645–655, 2006.

[24] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[25] M. G. Haehnelt, P. Natarajan, and M. J. Rees, "High-redshift galaxies, their active nuclei and central black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 300, no. 3, pp. 817–827, 1998.

[26] G. Kauffmann and M. Haehnelt, "A unified model for the evolution of galaxies and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 311, no. 3, pp. 576–588, 2000.

[27] M. G. Haehnelt and G. Kauffmann, "The correlation between black hole mass and bulge velocity dispersion in hierarchical galaxy formation models," *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 3, pp. L35–L38, 2000.

[28] J. S. B. Wyithe and A. Loeb, "Self-regulated growth of supermassive black holes in galaxies as the origin of the optical and X-ray luminosity functions of quasars," *Astrophysical Journal*, vol. 595, no. 2, pp. 614–623, 2003.

[29] M. Volonteri, F. Haardt, and P. Madau, "The assembly and merging history of supermassive black holes in hierarchical models of galaxy formation," *Astrophysical Journal*, vol. 582, no. 2, pp. 559–573, 2003.

[30] A. Cattaneo, J. Blaizot, J. Devriendt, and B. Guiderdoni, "Active galactic nuclei in cosmological simulations -I. Formation of black holes and spheroids through mergers," *Monthly Notices of the Royal Astronomical Society*, vol. 364, no. 2, pp. 407–423, 2005.

[31] T. Di Matteo, J. Colberg, V. Springel, L. Hernquist, and D. Sijacki, "Direct cosmological simulations of the growth of black holes and galaxies," *Astrophysical Journal*, vol. 676, no. 1, pp. 33–53, 2008.

[32] R. S. Somerville, P. F. Hopkins, T. J. Cox, B. E. Robertson, and L. Hernquist, "A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 391, no. 2, pp. 481–506, 2008.

[33] C. M. Booth and J. Schaye, "Cosmological simulations of the growth of supermassive black holes and feedback from active galactic nuclei: method and tests," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 1, pp. 53–74, 2009.

[34] F. Marulli, E. Branchini, L. Moscardini, and M. Volonteri, "Modelling active galactic nuclei: ongoing problems for the faint-end of the luminosity function," *Monthly Notices of the Royal Astronomical Society*, vol. 375, no. 2, pp. 649–656, 2007.

[35] L. Ciotti and J. P. Ostriker, "Cooling flows and quasars. II. Detailed models of feedback-modulated accretion flows," *Astrophysical Journal*, vol. 551, no. 1, pp. 131–152, 2001.

[36] P. F. Hopkins and L. Hernquist, "Fueling low-level AGN activity through stochastic accretion of cold gas," *Astrophysical Journal Supplement Series*, vol. 166, no. 1, pp. 1–36, 2006.

[37] O. Guyon, D. B. Sanders, and A. Stockton, "Near-infrared adaptive optics imaging of QSO host galaxies," *Astrophysical Journal Supplement Series*, vol. 166, no. 1, pp. 89–127, 2006.

[38] J. M. Gabor et al., "Active galactic nucleus host galaxy morphologies in COSMOS," *The Astrophysical Journal*, vol. 691, no. 1, pp. 705–722, 2009.

[39] A. Georgakakis, A. L. Coil, E. S. Laird et al., "Host galaxy morphologies of X-ray selected AGN: assessing the significance of different black hole fuelling mechanisms to the accretion density of the Universe at $z \sim 1$," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 2, pp. 623–633, 2009.

[40] A. Merloni, S. Heinz, and T. Di Matteo, "A fundamental plane of black hole activity," *Monthly Notices of the Royal Astronomical Society*, vol. 345, no. 4, pp. 1057–1076, 2003.

[41] H. Falcke, E. Körding, and S. Markoff, "A scheme to unify low-power accreting black holes. Jet-dominated accretion

flows and the radio/X-ray correlation," *Astronomy and Astrophysics*, vol. 414, no. 3, pp. 895–903, 2004.

[42] I. M. McHardy, E. Koerding, C. Knigge, P. Uttley, and R. P. Fender, "Active galactic nuclei as scaled-up Galactic black holes," *Nature*, vol. 444, no. 7120, pp. 730–732, 2006.

[43] S. W. Davis and A. Laor, "The radiative efficiency of accretion flows in individual active galactic nuclei," *Astrophysical Journal Letters*, vol. 728, no. 2, article 98, 2011.

[44] M. A. Sobolewska, A. Siemiginowska, and M. Gierliński, "Simulated spectral states of active galactic nuclei and observational predictions," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 3, pp. 2259–2268, 2011.

[45] B. C. Kelly, M. Sobolewska, and A. Siemiginowska, "A stochastic model for the luminosity fluctuations of accreting black holes," *Astrophysical Journal Letters*, vol. 730, no. 1, article 52, 2011.

[46] B. C. Kelly, M. Vestergaard, and X. Fan, "Determining quasar black hole mass functions from their broad emission lines: application to the bright quasar survey," *The Astrophysical Journal*, vol. 692, article 1388, 2009.

[47] A. Wandel, B. M. Peterson, and M. A. Malkan, "Central masses and broad-line region sizes of active galactic nuclei. I. Comparing the photoionization and reverberation techniques," *Astrophysical Journal*, vol. 526, no. 2, pp. 579–591, 1999.

[48] M. Vestergaard and B. M. Peterson, "Determining central black hole masses in distant active galaxies and quasars. II. Improved optical and UV scaling relationships," *Astrophysical Journal*, vol. 641, no. 2, pp. 689–709, 2006.

[49] M. Nikolajuk, I. E. Papadakis, and B. Czerny, "Black hole mass estimation from X-ray variability measurements in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 350, no. 2, pp. L26–L30, 2004.

[50] X.-L. Zhou, S.-N. Zhang, D.-X. Wang, and L. Zhu, "Calibrating the correlation between black hole mass and x-ray variability amplitude: X-ray only black hole mass estimates for active galactic nuclei and ultra-luminous x-ray sources," *Astrophysical Journal Letters*, vol. 710, no. 1, pp. 16–23, 2010.

[51] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal*, vol. 698, no. 1, pp. 198–221, 2009.

[52] B. C. Kelly, X. Fan, and M. Vestergaard, "A flexible method of estimating luminosity functions," *Astrophysical Journal*, vol. 682, no. 2, pp. 874–895, 2008.

[53] A. Marconi, G. Risaliti, R. Gilli, L. K. Hunt, R. Maiolino, and M. Salvati, "Local supermassive black holes, relics of active galactic nuclei and the X-ray background," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 1, pp. 169–185, 2004.

[54] P. Salucci, E. Szuszkiewicz, P. Monaco, and L. Danese, "Mass function of dormant black holes and the evolution of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 307, no. 3, pp. 637–644, 1999.

[55] Q. Yu and S. Tremaine, "Observational constraints on growth of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 4, pp. 965–976, 2002.

[56] M. C. Aller and D. Richstone, "The cosmic density of massive black holes from galaxy velocity dispersions," *Astronomical Journal*, vol. 124, no. 6, pp. 3035–3041, 2002.

[57] F. Shankar, P. Salucci, G. L. Granato, G. De Zotti, and L. Danese, "Supermassive black hole demography: the match between the local and accreted mass functions," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 4, pp. 1020–1030, 2004.

[58] T. R. Lauer, S. M. Faber, D. Richstone et al., "The masses of nuclear black holes in luminous elliptical galaxies and implications for the space density of the most massive black holes," *Astrophysical Journal*, vol. 662, no. 2, pp. 808–834, 2007.

[59] A. W. Graham, S. P. Driver, P. D. Allen, and J. Liske, "The Millennium Galaxy Catalogue: the local supermassive black hole mass function in early- and late-type galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 378, no. 1, pp. 198–210, 2007.

[60] A. W. Graham and S. P. Driver, "The local supermassive black hole mass density: corrections for dependencies on the Hubble constant," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 1, pp. L15–L19, 2007.

[61] E. Tundo, M. Bernardi, J. B. Hyde, R. K. Sheth, and A. Pizzella, "On the inconsistency between the black hole mass function inferred from M.-σ and M.-L correlations," *Astrophysical Journal*, vol. 663, no. 1, pp. 53–60, 2007.

[62] Q. Yu and Y. Lu, "Toward precise constraints on the growth of massive black holes," *Astrophysical Journal Letters*, vol. 689, no. 2, pp. 732–754, 2008.

[63] A. Merloni and S. Heinz, "A synthesis model for AGN evolution: supermassive black holes growth and feedback modes," *Monthly Notices of the Royal Astronomical Society*, vol. 388, no. 3, pp. 1011–1030, 2008.

[64] F. Shankar, D. H. Weinberg, and J. Miralda-Escudé, "Self-consistent models of the AGN and black hole populations: duty cycles, accretion rates, and the mean radiative efficiency," *Astrophysical Journal Letters*, vol. 690, no. 1, pp. 20–41, 2009.

[65] M. Vika, S. P. Driver, A. W. Graham, and J. Liske, "The Millennium Galaxy Catalogue: the Mbh-Lspheroid derived supermassive black hole mass function," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 3, pp. 1451–1460, 2009.

[66] N. Tamura, K. Ohta, and Y. Ueda, "Supermassive black hole mass functions at intermediate redshifts from spheroid and AGN luminosity functions," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 134–146, 2006.

[67] F. Shankar, M. Bernardi, and Z. Haiman, "The evolution of the MBH-σ relation inferred from the age distribution of local early-type galaxies and active galactic nuclei evolution," *Astrophysical Journal Letters*, vol. 694, no. 2, pp. 867–878, 2009.

[68] Z. Haiman, L. Ciotti, and J. P. Ostriker, "Reasoning from fossils: learning from the local black hole population about the evolution of quasars," *Astrophysical Journal Letters*, vol. 606, no. 2, pp. 763–773, 2004.

[69] Z. Haiman, R. Jimenez, and M. Bernardi, "Reconstructing the cosmic evolution of quasars from the age distribution of local early-type galaxies," *Astrophysical Journal Letters*, vol. 658, no. 2, pp. 721–730, 2007.

[70] T. Treu, M. A. Malkan, and R. D. Blandford, "The relation between black hole mass and velocity dispersion at $z \sim 0.37$," *Astrophysical Journal*, vol. 615, no. 2, pp. L97–L100, 2004.

[71] C. Y. Peng, C. D. Impey, H. W. Rix et al., "Probing the coevolution of supermassive black holes and galaxies using gravitationally lensed quasar hosts," *Astrophysical Journal*, vol. 649, no. 2, pp. 616–634, 2006.

[72] T. Treu, J. H. Woo, M. A. Malkan, and R. D. Blandford, "Cosmic evolution of black holes and spheroids. II. Scaling

relations at z = 0.36," *Astrophysical Journal*, vol. 667, no. 1, pp. 117–130, 2007.

[73] J. H. Woo, T. Treu, M. A. Malkan, and R. D. Blandford, "Cosmic evolution of black holes and spheroids. III. The M BH-$\sigma*$ relation in the last six billion years," *Astrophysical Journal*, vol. 681, no. 2, pp. 925–930, 2008.

[74] A. Merloni, A. Bongiorno, M. Bolzonella et al., "On the cosmic evolution of the scaling relations between black holes and their host galaxies: broad-line active galactic nuclei in the zCosmos survey," *Astrophysical Journal*, vol. 708, no. 1, pp. 137–157, 2010.

[75] V. N. Bennert, T. Treu, J. H. Woo et al., "Cosmic evolution of black holes and spheroids. IV. the M BH-L sph relation," *Astrophysical Journal*, vol. 708, no. 2, pp. 1507–1527, 2010.

[76] T. R. Lauer, S. Tremaine, D. Richstone, and S. M. Faber, "Selection bias in observing the cosmological evolution of the M-σ- and M-L relationships," *Astrophysical Journal*, vol. 670, no. 1, pp. 249–260, 2007.

[77] Y. Shen and B. C. Kelly, "The impact of the uncertainty in single-epoch virial black hole mass estimates on the observed evolution of the black hole-bulge scaling relations," *Astrophysical Journal Letters*, vol. 713, no. 1, pp. 41–45, 2010.

[78] R. K. Sheth, M. Bernardi, P. L. Schechter et al., "The velocity dispersion function of early-type galaxies," *Astrophysical Journal*, vol. 594, no. 1, pp. 225–231, 2003.

[79] M. Bernardi, R. K. Sheth, E. Tundo, and J. B. Hyde, "Selection bias in the M$_\bullet$-σ and M$_\bullet$-L correlations and its consequences," *Astrophysical Journal*, vol. 660, no. 1, pp. 267–275, 2007.

[80] A. W. Graham, "Fundamental planes and the barless MBH-σ relation for supermassive black holes," *Astrophysical Journal Letters*, vol. 680, no. 1, pp. 143–153, 2008.

[81] J. E. Greene, C. Y. Peng, M. Kim et al., "Precise black hole masses from megamaser disks: black hole-bulge relations at low mass," *Astrophysical Journal Letters*, vol. 721, no. 1, pp. 26–45, 2010.

[82] J. Hu, "The black hole mass-stellar velocity dispersion correlation: bulges versus pseudo-bulges," *Monthly Notices of the Royal Astronomical Society*, vol. 386, no. 4, pp. 2242–2252, 2008.

[83] J. E. Greene, L. C. Ho, and A. J. Barth, "Black holes in pseudobulges and spheroidals: a change in the black hole-bulge scaling relations at low mass," *Astrophysical Journal*, vol. 688, no. 1, pp. 159–179, 2008.

[84] A. W. Graham and I. H. Li, "The M bh-σ diagram and the offset nature of barred active galaxies," *Astrophysical Journal*, vol. 698, no. 1, pp. 812–818, 2009.

[85] J. Kormendy, R. Bender, and M. E. Cornell, "Supermassive black holes do not correlate with galaxy disks or pseudobulges," *Nature*, vol. 469, no. 7330, pp. 374–376, 2011.

[86] A. W. Graham, C. A. Onken, E. Athanassoula, and F. Combes, "An expanded Mbh-σ diagram, and a new calibration of active galactic nuclei masses," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2211–2228, 2011.

[87] A. Soltan, "Masses of quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 200, pp. 115–122, 1982.

[88] Q. Yu and Y. Lu, "Constraints on QSO models from a relation between the QSO luminosity function and the local black hole mass function," *Astrophysical Journal Letters*, vol. 602, no. 2 I, pp. 603–624, 2004.

[89] A. Merloni, "The anti-hierarchical growth of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 353, no. 4, pp. 1035–1047, 2004.

[90] P. F. Hopkins, G. T. Richards, and L. Hernquist', "An observational determination of the bolometric quasar luminosity function," *Astrophysical Journal*, vol. 654, no. 2, pp. 731–753, 2007.

[91] X. Cao and F. Li, "Rapidly spinning massive black holes in active galactic nuclei: evidence from the black hole mass function," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 2, pp. 561–566, 2008.

[92] X. Cao, "Cosmological evolution of massive black holes: Effects of eddington ratio distribution and quasar lifetime," *Astrophysical Journal*, vol. 725, no. 1, pp. 388–393, 2010.

[93] A. Cavaliere, P. Morrison, and K. Wood, "On Quasar evolution," *Astrophysical Journal*, vol. 170, pp. 223–231, 1971.

[94] T. A. Small and R. D. Blandford, "Quasar evolution and the growth of black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 259, pp. 725–737, 1992.

[95] B. C. Kelly, J. Bechtold, J. R. Trump, M. Vestergaard, and M. Siemiginowska, "Observational constraints on the dependence of radio-quiet quasar X-ray emission on black hole mass and accretion rate," *Astrophysical Journal Supplement Series*, vol. 176, no. 2, pp. 355–373, 2008.

[96] R. V. Vasudevan and A. C. Fabian, "Piecing together the X-ray background: bolometric corrections for active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 381, no. 3, pp. 1235–1251, 2007.

[97] R. V. Vasudevan, R. F. Mushotzky, L. M. Winter, and A. C. Fabian, "Optical-to-X-ray emission in low-absorption AGN: results from the Swift-BAT 9-month catalogue," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 3, pp. 1553–1575, 2009.

[98] S. I. Raimundo and A. C. Fabian, "Eddington ratio and accretion efficiency in active galactic nuclei evolution," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 3, pp. 1217–1221, 2009.

[99] P. F. Hopkins, L. Hernquist, T. J. Cox, B. Robertson, T. Di Matteo, and V. Springel, "The evolution in the faint-end slope of the quasar luminosity function," *Astrophysical Journal*, vol. 639, no. 2, pp. 700–709, 2006.

[100] Q. Yu, Y. Lu, and G. Kauffmann, "Evolution of accretion disks around massive black holes: constraints from the demography of active galactic nuclei," *Astrophysical Journal*, vol. 634, no. 2, pp. 901–909, 2005.

[101] A. R. King and J. E. Pringle, "Fuelling active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L25–L28, 2007.

[102] B. M. Peterson, L. Ferrarese, K. M. Gilbert et al., "Central masses and broad-line region sizes of active galactic nuclei. II. A homogeneous analysis of a large reverberation-mapping database," *Astrophysical Journal Letters*, vol. 613, no. 2, pp. 682–699, 2004.

[103] M. C. Bentz et al., "The lick AGN monitoring project: broad-line region radii and black hole masses from reverberation mapping of Hβ," *Astrophysical Journal*, vol. 705, article 199, 2009.

[104] S. Kaspi, D. Maoz, H. Netzer, B. M. Peterson, M. Vestergaard, and B. T. Jannuzi, "The relationship between luminosity and broad-line region size in active galactic nuclei," *Astrophysical Journal*, vol. 629, no. 1, pp. 61–71, 2005.

[105] M. C. Bentz, B. M. Peterson, H. Netzer, R. W. Pogge, and M. Vestergaard, "The radius-luminosity relationship for active galactic nuclei: the effect of host-galaxy starlight on luminosity measurements. II. the full sample of reverberation-mapped agns," *Astrophysical Journal*, vol. 697, no. 1, pp. 160–181, 2009.

[106] C. A. Onken, L. Ferrarese, D. Merritt et al., "Supermassive black holes in active galactic nuclei. II. Calibration of the black hole mass-velocity dispersion relationship for active galactic nuclei," *Astrophysical Journal*, vol. 615, no. 2, pp. 645–651, 2004.

[107] J.-H. Woo, T. Treu, A. J. Barth et al., "The lick AGN monitoring project: the MBH- $\sigma*$ relation for reverberation-mapped active galaxies," *Astrophysical Journal Letters*, vol. 716, no. 1, pp. 269–280, 2010.

[108] J. E. Greene and L. C. Ho, "Estimating black hole masses in active galaxies using the Hα emission line," *Astrophysical Journal*, vol. 630, no. 1, pp. 122–129, 2005.

[109] R. J. McLure and M. J. Jarvis, "Measuring the black hole masses of high-redshift quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 337, no. 1, pp. 109–116, 2002.

[110] M. Vestergaard and P. S. Osmer, "Mass functions of the active black holes in distant quasars from the large bright quasar survey, the bright quasar survey, and the color-selected sample of the sdss fall equatorial stripe," *Astrophysical Journal*, vol. 699, no. 1, pp. 800–816, 2009.

[111] Y. Shen, G. T. Richards, M. A. Strauss et al., "A catalog of quasar properties from Sloan Digital Sky Survey Data Release 7," *Astrophysical Journal Supplement Series*, vol. 194, no. 2, article 45, 2011.

[112] J. A. Kollmeier, C. A. Unken, C. S. Kochanek et al., "Black hole masses and eddington ratios at $0.3 < z < 4$," *Astrophysical Journal*, vol. 648, no. 1, pp. 128–139, 2006.

[113] Y. Shen, J. E. Greene, M. A. Strauss, G. T. Richards, and D. P. Schneider, "Biases in virial black hole masses: an SDSS perspective," *Astrophysical Journal*, vol. 680, no. 1, pp. 169–190, 2008.

[114] C. L. Steinhardt and M. Elvis, "The quasar mass-luminosity plane-III. Smaller errors on virial mass estimates," *Monthly Notices of the Royal Astronomical Society*, vol. 406, no. 1, pp. L1–L5, 2010.

[115] B. C. Kelly, M. Vestergaard, X. Fan, P. Hopkins, L. Hernquist, and A. Siemiginowska, "Constraints on black hole growth, quasar lifetimes, and Eddington ratio distributions from the SDSS broad-line quasar black hole mass function," *Astrophysical Journal*, vol. 719, no. 2, pp. 1315–1334, 2010.

[116] Y. Shen and B. C. Kelly, "The demographics of broad-line quasars in the mass-luminosity plane. I. testing FWHM-based virial black hole masses," *Astrophysical Journal*. In press.

[117] J.-M. Wang, Y. M. Chen, and F. Zhang, "Cosmological evolution of the duty cycle of quasars," *Astrophysical Journal Letters*, vol. 647, no. 1, pp. L17–L20, 2006.

[118] J. E. Greene and L. C. Ho, "The mass function of active black holes in the local universe," *Astrophysical Journal*, vol. 667, no. 1, pp. 131–148, 2007.

[119] M. Vestergaard, X. Fan, C. A. Tremonti, P. S. Osmer, and G. T. Richards, "Mass functions of the active black holes in distant quasars from the sloan digital sky survey data release 3," *Astrophysical Journal*, vol. 674, no. 1, pp. L1–L4, 2008.

[120] G. T. Richards, M. A. Strauss, X. Fan et al., "The sloan digital sky survey quasar survey: quasar luminosity function from data release 3," *Astronomical Journal*, vol. 131, no. 6, pp. 2766–2787, 2006.

[121] A. Schulze and L. Wisotzki, "Low redshift AGN in the Hamburg/ESO Survey: II. The active black hole mass function and the distribution function of Eddington ratios," *Astronomy and Astrophysics*, vol. 516, no. 15, article A87, 2010.

[122] M. Schmidt and R. F. Green, "Quasar evolution derived from the Palomar bright quasar survey and other complete quasar surveys," *Astrophysical Journal*, vol. 269, pp. 352–374, 1983.

[123] P. Natarajan and E. Treister, "Is there an upper limit to black hole masses?" *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 3, pp. 838–845, 2009.

[124] D. Sijacki, V. Springel, and M. G. Haehnelt, "Growing the first bright quasars in cosmological simulations of structure formation," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 1, pp. 100–122, 2009.

[125] J. E. Greene and L. C. Ho, "Active galaxies and the study of black hole demographics," *Publications of the Astronomical Society of the Pacific*, vol. 121, no. 885, pp. 1167–1171, 2009.

[126] S. Collin, T. Kawaguchi, B. M. Peterson, and M. Vestergaard, "Systematic effects in measurement of black hole masses by emission-line reverberation of active galactic nuclei: eddington ratio and inclination," *Astronomy and Astrophysics*, vol. 456, no. 1, pp. 75–90, 2006.

[127] G. T. Richards, N. E. Kruczek, S. C. Gallagher et al., "Unification of luminous type 1 quasars through Civ emission," *Astronomical Journal*, vol. 141, no. 5, article 167, 2011.

[128] A. Marconi, D. J. Axon, R. Maiolino et al., "The effect of radiation pressure on virial black hole mass estimates and the case of narrow-line Seyfert 1 galaxies," *Astrophysical Journal*, vol. 678, no. 2, pp. 693–700, 2008.

[129] H. Netzer, "Radiation pressure force and black hole mass determination in low redshift type-I and type-II active galactic nuclei," *The Astrophysical Journal*, vol. 695, no. 1, article 793, 2009.

[130] A. Marconi, D. J. Axon, R. Maiolino et al., "On the observed distributions of black hole masses and eddington ratios from radiation pressure corrected virial indicators," *The Astrophysical Journal Letters*, vol. 698, no. 2, article L103, 2009.

[131] H. Netzer and P. Marziani, "The effect of radiation pressure on emission-line profiles and black hole mass determination in active galactic nuclei," *Astrophysical Journal*, vol. 724, no. 1, pp. 318–328, 2010.

[132] K. D. Denney, B. M. Peterson, M. Dietrich, M. Vestergaard, and M. C. Bentz, "Systematic uncertainties in measuring black hole masses from single epoch spectra," *The Astrophysical Journal*, vol. 692, article 246, 2009.

[133] R. I. Davies, J. Thomas, R. Genzel et al., "The star-forming torus and stellar dynamical black hole mass in the Seyfert 1 nucleus of NGC 3227," *Astrophysical Journal*, vol. 646, no. 2, pp. 754–773, 2006.

[134] E. K. S. Hicks and M. A. Malkan, "Circumnuclear gas in seyfert 1 galaxies: morphology, kinematics, and direct measurement of black hole masses," *Astrophysical Journal Supplement Series*, vol. 174, no. 1, pp. 31–73, 2008.

[135] C. A. Onken, M. Valluri, B. M. Peterson et al., "The black hole mass of NGC 4151: comparison of reverberation mapping and stellar dynamical measurements," *Astrophysical Journal*, vol. 670, no. 1, pp. 105–115, 2007.

[136] A. Siemiginowska and M. Elvis, "Deriving the quasar luminosity function from accretion-disk instabilities," *Astrophysical Journal*, vol. 482, no. 1, pp. L9–L12, 1997.

[137] E. Hatziminaoglou, A. Siemioinowska, and M. Elvis, "Accretion disk instabilities, cold dark matter models, and their role in quasar evolution," *Astrophysical Journal*, vol. 547, no. 1, pp. 90–98, 2001.

[138] A. Semicinowska, B. Czerny, and V. Kostyunin, "Evolution of an accretion disk in an active galactic nucleus," *Astrophysical Journal*, vol. 458, no. 2, pp. 491–507, 1996.

[139] A. Franceschini, S. Vercellone, and A. C. Fabian, "Supermassive black holes in early-type galaxies: relationship with radio emission and constraints on the black hole mass function," *Monthly Notices of the Royal Astronomical Society*, vol. 297, no. 3, pp. 817–824, 1998.

[140] M. G. Haehnelt and M. J. Rees, "The formation of nuclei in newly formed galaxies and the evolution of the quasar population," *Monthly Notices of the Royal Astronomical Society*, vol. 263, no. 1, pp. 168–178, 1993.

[141] T. Hosokawa, "Constraining the lifetime of quasars with the present-day mass function of supermassive black holes," *Astrophysical Journal*, vol. 576, no. 1, pp. 75–80, 2002.

[142] A. Cavaliere and V. Vittorini, "Supermassive black holes in galactic nuclei," *Astrophysical Journal*, vol. 570, no. 1, pp. 114–118, 2002.

[143] M. Enoki, M. Nagashima, and N. Gouda, "Relations between galaxy formation and the environments of quasars," *Publications of the Astronomical Society of Japan*, vol. 55, no. 1, pp. 133–142, 2003.

[144] G. L. Granato, G. De Zotti, L. Silva, A. Bressan, and L. Danese, "A physical model for the coevolution of QSOs and their spheroidal hosts," *Astrophysical Journal*, vol. 600, no. 2, pp. 580–594, 2004.

[145] P. F. Hopkins, L. Hernquist, T. J. Cox, and D. Kerbs, "A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. i. galaxy mergers and quasar activity," *Astrophysical Journal Supplement Series*, vol. 175, no. 2, pp. 356–389, 2008.

[146] Y. Shen, "Supermassive black holes in the hierarchical universe: a general framework and observational tests," *Astrophysical Journal Letters*, vol. 704, no. 1, pp. 89–108, 2009.

[147] N. Fanidakis, C. M. Baugh, A. J. Benson et al., "Grand unification of AGN activity in the ΛCDM cosmology," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 1, pp. 53–74, 2011.

[148] N. Fanidakis, Baugh C. M., A. J. Benson et al., "The evolution of active galactic nuclei across cosmic time: what is downsizing?" *Monthly Notices of the Royal Astronomical Society*. In press.

[149] F. Hoyle and R. A. Lyttleton, "The effect of interstellar matter on climatic variation," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 35, no. 3, pp. 405–415, 1939.

[150] H. Bondi and F. Hoyle, "On the mechanism of accretion by stars," *Monthly Notices of the Royal Astronomical Society*, vol. 104, p. 273, 1944.

[151] H. Bondi, "On spherically symmetrical accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 112, p. 195, 1952.

[152] M. Volonteri and M. C. Begelman, "Quasi-stars and the cosmic evolution of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 409, no. 3, pp. 1022–1032, 2010.

[153] L. Jiang, X. Fan, M. Vestergaard et al., "Gemini near-infrared spectroscopy of luminous z ∼ 6 quasars: chemical abundances, black hole masses, and Mg II absorption," *Astronomical Journal*, vol. 134, no. 3, pp. 1150–1161, 2007.

[154] D. J. Mortlock, S. J. Warren, B. P. Venemans et al., "A luminous quasar at a redshift of z = 7.085," *Nature*, vol. 474, no. 7353, pp. 616–619, 2011.

[155] Z. Haiman and A. Loeb, "What is the highest plausible redshift of luminous quasars?" *Astrophysical Journal Letters*, vol. 552, no. 2, pp. 459–463, 2001.

[156] G. Lodato and P. Natarajan, "The mass function of high-redshift seed black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L64–L68, 2007.

[157] G. Lodato and P. Natarajan, "Supermassive black hole formation during the assembly of pre-galactic discs," *Monthly Notices of the Royal Astronomical Society*, vol. 371, no. 4, pp. 1813–1823, 2006.

[158] M. C. Begelman, M. Volonteri, and M. J. Rees, "Formation of supermassive black holes by direct collapse in pre-galactic haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 1, pp. 289–298, 2006.

[159] M. C. Begelman, E. M. Rossi, and P. J. Armitage, "Quasi-stars: accreting black holes inside massive envelopes," *Monthly Notices of the Royal Astronomical Society*, vol. 387, no. 4, pp. 1649–1659, 2008.

[160] M. C. Begelman, "Evolution of supermassive stars as a pathway to black hole formation," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 673–681, 2010.

[161] M. Volonteri, G. Lodato, and P. Natarajan, "The evolution of massive black hole seeds," *Monthly Notices of the Royal Astronomical Society*, vol. 383, no. 3, pp. 1079–1088, 2008.

[162] P. Natarajan and M. Volonteri, "The mass function of black holes at 1<z<4.5: comparison of models with observations," *Monthly Notices of the Royal Astronomical Society*. In press.

[163] Z. Lippai, Z. Frei, and Z. Haiman, "On the occupation fraction of seed black holes in high-redshift dark matter halos," *Astrophysical Journal Letters*, vol. 701, no. 1, pp. 360–368, 2009.

[164] T. Tanaka and Z. Haiman, "The assembly of supermassive black holes at high redshifts," *The Astrophysical Journal*, vol. 696, p. 1798, 2009.

[165] E. Berti and M. Volonteri, "Cosmological black hole spin evolution by mergers and accretion," *Astrophysical Journal*, vol. 684, no. 2, pp. 822–828, 2008.

[166] K. Horne, B. M. Peterson, S. J. Collier, and H. Netzer, "Observational requirements for high-fidelity reverberation mapping," *Publications of the Astronomical Society of the Pacific*, vol. 116, no. 819, pp. 465–476, 2004.

[167] M. C. Bentz, K. Horne, A. J. Barth et al., "The lick agn monitoring project: velocity-delay maps from the maximum-entropy method for Arp 151," *Astrophysical Journal*, vol. 720, no. 1, pp. L46–L51, 2010.

[168] A. Pancoast, B. J. Brewer, and T. Treu, "Geometric and dynamical models of reverberation mapping data," *Astrophysical Journal Letters*, vol. 730, no. 2, article 139, 2011.

[169] Y. Zu, C. S. Kochanek, and B. M. Peterson, "An alternative approach to measuring reverberation lags in active galactic nuclei," *Astrophysical Journal Letters*, vol. 735, no. 2, article 80, 2011.

[170] B. J. Brewer, T. Treu, A. Pancoast et al., "The mass of the black hole in Arp 151 from bayesian modeling of reverberation mapping data," *Astrophysical Journal Letters*, vol. 733, no. 2, article L33, 2011.

[171] A. Merloni, S. Heinz, and T. Di Matteo, "A fundamental plane of black hole activity: pushing forward the unification scheme," *Astrophysics and Space Science*, vol. 300, no. 1-3, pp. 45–53, 2005.

[172] E. Körding, H. Falcke, and S. Corbel, "Refining the fundamental plane of accreting black holes," *Astronomy and Astrophysics*, vol. 456, no. 2, pp. 439–450, 2006.

[173] K. Gültekin, E. M. Cackett, J. M. Miller, T. Di Matteo, S. Markoff, and D. O. Richstone, "The fundamental plane of accretion onto black holes with dynamical masses," *Astrophysical Journal*, vol. 706, no. 1, pp. 404–416, 2009.

[174] R. M. Plotkin, S. Markoff, B. C. Kelly, E. Kö rding, and S. F. Anderson, "Using the Fundamental Plane of black hole activity to distinguish X-ray processes from weakly accreting black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 419, no. 1, pp. 267–286, 2012.

[175] A. Laor and E. Behar, "On the origin of radio emission in radio-quiet quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 2, pp. 847–862, 2008.

[176] B. C. Kelly, J. Bechtold, and A. Siemiginowska, "Are the variations in quasar optical flux driven by thermal fluctuations?" *Astrophysical Journal*, vol. 698, no. 1, pp. 895–910, 2009.

[177] C. L. MacLeod, Z. Ivezić, C. S. Kochanek et al., "Modeling the time variability of SDSS Stripe 82 quasars as a damped random walk," *Astrophysical Journal*, vol. 721, no. 2, pp. 1014–1033, 2010.

[178] M. Brusa, F. Fiore, P. Santini et al., "Black hole growth and starburst activity at z = 0.6-4 in the Chandra Deep Field South," *Astronomy and Astrophysics*, vol. 507, no. 3, pp. 1277–1289, 2009.

[179] C. N. Cardamone, C. Megan Urry, K. Schawinski, E. Treister, G. Brammer, and E. Gawiser, "Dust-corrected colors reveal bimodality in the host-galaxy colors of active galactic nuclei at z ∼ 1," *Astrophysical Journal*, vol. 721, no. 1, pp. L38–L42, 2010.

[180] Y. Q. Xue, W. N. Brandt, B. Luo et al., "Color-magnitude relations of active and non-active galaxies in the Chandra Deep fields: high-redshift constraints and stellar-mass selection effects," *Astrophysical Journal*, vol. 720, no. 1, pp. 368–391, 2010.

[181] J. Aird, A. L. Coil, J. Moustakas et al., "PRIMUS: the dependence of AGN accretion on host stellar mass and color," *The Astrophysical Journal*. In press.

[182] D. M. Alexander, W. N. Brandt, I. Smail et al., "Weighing the black holes in Z 2 submillimeter-emitting galaxies hosting active galactic nuclei," *Astronomical Journal*, vol. 135, no. 5, pp. 1968–1981, 2008.

[183] J. E. Sarria, R. Maiolino, F. La Franca et al., "The MBH-Mstar relation of obscured AGNs at high redshift," *Astronomy and Astrophysics*, vol. 522, no. 1, article L3, 2010.

[184] A. Schulze and L. Wisotzki, "Selection effects in the black hole-bulge relation and its evolution," *Astronomy and Astrophysics*, vol. 535, article A87, 2011.

[185] M. Volonteri, "Gravitational recoil: signatures on the massive black hole population," *Astrophysical Journal*, vol. 663, no. 1, pp. L5–L8, 2007.

[186] L. Blecha and A. Loeb, "Effects of gravitational-wave recoil on the dynamics and growth of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 4, pp. 1311–1325, 2008.

[187] D. Sijacki, V. Springel, and M. G. Haehnelt, "Gravitational recoils of supermassive black holes in hydrodynamical simulations of gas-rich galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 4, pp. 3656–3670, 2011.

[188] L. Blecha, T. J. Cox, A. Loeb, and L. Hernquist, "Recoiling black holes in merging galaxies: relationship to active galactic nucleus lifetimes, starbursts and the MBH-$\sigma*$ relation," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2154–2182, 2011.

[189] M. Volonteri, K. Gültekin, and M. Dotti, "Gravitational recoil: effects on massive black hole occupation fraction over cosmic time," *Monthly Notices of the Royal Astronomical Society*, vol. 404, no. 4, pp. 2143–2150, 2010.

[190] M. Volonteri, P. Natarajan, and K. Gültekin, "How important is the dark matter halo for black hole growth?" *Astrophysical Journal*, vol. 737, no. 2, article 50, 2011.

[191] L. Ciotti, J. P. Ostriker, and D. Proga, "Feedback from central black holes in elliptical galaxies. I. Models with either radiative or mechanical feedback but not both," *Astrophysical Journal*, vol. 699, no. 1, pp. 89–104, 2009.

[192] M. S. Shin, J. P. Ostriker, and L. Ciotti, "Feedback from central black holes in elliptical galaxies. II. Can purely mechanical energy feedback models work?" *Astrophysical Journal*, vol. 711, no. 1, pp. 268–283, 2010.

[193] L. Ciotti, J. P. Ostriker, and D. Proga, "Feedback from central black holes in elliptical galaxies. III. Models with both radiative and mechanical feedback," *Astrophysical Journal*, vol. 717, no. 2, pp. 708–723, 2010.

[194] R. Kurosawa, D. Proga, and K. Nagamine, "ON the feedback efficiency of active galactic nuclei," *Astrophysical Journal*, vol. 707, no. 1, pp. 823–832, 2009.

[195] J. P. Ostriker, E. Choi, L. Ciotti, G. S. Novak, and D. Proga, "Momentum driving: which physical processes dominate active galactic nucleus feedback?" *Astrophysical Journal*, vol. 722, no. 1, pp. 642–652, 2010.

[196] A. C. Fabian, J. S. Sanders, S. W. Allen et al., "A deep Chandra observation of the Perseus cluster: shocks and ripples," *Monthly Notices of the Royal Astronomical Society*, vol. 344, no. 3, pp. L43–L47, 2003.

[197] P. F. Hopkins and E. Quataert, "An analytic model of angular momentum transport by gravitational torques: from galaxies to massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 2, pp. 1027–1050, 2011.

[198] C. Power, S. Nayakshin, and A. King, "The accretion disc particle method for simulations of black hole feeding and feedback," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 1, pp. 269–276, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 846875, 15 pages
doi:10.1155/2012/846875

Review Article

The Role of Gravitational Instabilities in the Feeding of Supermassive Black Holes

Giuseppe Lodato

Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria 16, Milano, Italy

Correspondence should be addressed to Giuseppe Lodato, giuseppe.lodato@unimi.it

Received 28 August 2011; Accepted 7 October 2011

Academic Editor: Francesco Shankar

I review the recent progresses that have been obtained, especially through the use of high-resolution numerical simulations, on the dynamics of self-gravitating accretion discs. A coherent picture is emerging, where the disc dynamics is controlled by a small number of parameters that determine whether the disc is stable or unstable, whether the instability saturates in a self-regulated state or runs away into fragmentation, and whether the dynamics is local or global. I then apply these concepts to the case of AGN discs, discussing the implications of such evolution on the feeding of supermassive black holes. Nonfragmenting, self-gravitating discs appear to play a fundamental role in the process of formation of massive black hole seeds at high redshift ($z \sim 10$–15) through direct gas collapse. On the other hand, the different cooling properties of the interstellar gas at low redshifts determine a radically different behaviour for the outskirts of the accretion discs feeding typical AGNs. Here the situation is much less clear from a theoretical point of view, and while several observational clues point to the important role of massive discs at a distance of roughly a parsec from their central black hole, their dynamics is still under debate.

1. Introduction

The accretion discs surrounding the growing supermassive black holes (SMBH) in active galactic nuclei (AGN) are expected to become gravitationally unstable at a distance of ~ 0.01 pc from the black hole [1, 2]. Traditionally, this occurrence has been interpreted in relation to star formation: a self-gravitating disc, in this picture, would rapidly fragment and form stars [3, 4]. At the same time, it has been noted very early that the development of gravitational instability may also act as an efficient mechanism to produce torques through the effect of the resulting spiral structure and thus might be very effective in redistributing angular momentum within the disc and promote accretion [5–7]. As we shall see, the modern debate about these issues still concentrates on these two extreme cases. While we now have a much clearer understanding of the mechanism of growth and saturation of the instability in gaseous discs, and—especially though the use of high-resolution numerical simulations—we have clarified what are the main parameters regulating the disc structure and evolution, some questions are still unanswered. Are massive discs effectively truncated by star formation at the radius where they become self-gravitating, thus preventing accretion beyond these scales? Or does accretion proceed effectively through gravitational torques even in fragmenting discs, allowing the central black holes to be fed by gas on parsec scales?

From the observational point of view, on the one hand, it has now become quite clear that fragmentation in massive discs can be very important for the formation of compact, young stellar clusters in AGNs, and in particular in our own Galaxy [8, 9]. On the other hand, it is also clear that rotating gaseous discs exist on parsec scales in AGN [10–12], often displaying a clear Keplerian rotation [10].

All the issues discussed above bear important consequences not just for the dynamics of the disc itself, but, in a broader context, relate to the overall process of coevolution between the supermassive black hole and the host galaxy. The efficiency of star formation in the disc, the efficiency of the accretion process, and the related timescales and duty cycle of AGN activity are all often assumed as subgrid physics in simulations of galactic evolution on larger scales [17–19], which turn out to be quite sensitive to the chosen subgrid prescriptions.

In this contribution, I will not try to give an exhaustive answer to the questions above. I will rather give an account of the progresses we have made in recent years in our understanding of these phenomena and highlight their importance in several contexts related to the feeding of supermassive black holes. I will first summarize, in Section 2, the state of the art about the evolution of gravitational instabilities in gaseous discs, from a purely theoretical point of view. In Section 3, I will describe the possible importance of gravitational torques in the formation of the seeds of supermassive black holes by direct collapse in the early evolution of pregalactic discs. In Section 4, I will address the issue of feeding the SMBH in AGN through gravitational torques, and the related issue of fragmentation of AGN discs. Finally, in Section 5 I will draw some conclusions.

2. Gravitational Instabilities in Gaseous Discs

The issue of the nonlinear evolution of gravitational instabilities in gaseous discs has been studied in great detail over the last 10–15 years [13, 20–25]. As a result, despite the differences in the numerical methods adopted and in the setup used, a coherent picture of the overall dynamics is emerging. This issue has also been covered in several reviews, see for example Lodato [26] and Durisen et al. [27], and the reader is referred to these papers for further details and for an application of these concepts to different astrophysical systems (such as protostellar and protoplanetary discs), which share similar characteristics.

Consider an accretion disc with surface density $\Sigma(R)$, where R is the cylindrical distance to the central object of mass M, around which the disc is rotating in approximately centrifugal balance with angular velocity $\Omega(R)$. Let us also define the epicyclic frequency κ, which is equal to Ω in the case in which the rotation curve of the disc is Keplerian, $\Omega^2 = GM/R^3$. If the disc mass is high enough ($M_{\rm disc} \approx M$), deviations from Keplerian rotation might arise [28] and κ is not going to be exactly equal to Ω. The disc midplane temperature is $T(R)$, and the sound speed is $c_s \propto T^{1/2}$. The disc thickness is $H = c_s/\Omega$ for a non-self-gravitating disc and $H = c_s^2/\pi G\Sigma$ for a self-gravitating disc: we shall see that for gravitationally unstable discs the two definitions are equivalent. For most cases, we will consider thin discs, for which $H/R \ll 1$.

Fundamentally, the dynamics of self-gravitating accretion discs depends on three dimensionless parameters. Firstly, there is the well-known axisymmetric stability parameter $Q = c_s\kappa/\pi G\Sigma$ [29]. The second important parameter is the ratio between the cooling time $t_{\rm cool}$ and the dynamical time $t_{\rm dyn} = \Omega^{-1}$, a parameter often called $\beta = \Omega t_{\rm cool}$. Thirdly, we have the ratio between the disc mass and the central object mass $M_{\rm disc}/M$. As we shall see, each of these parameters controls some important features about the evolution of the gravitational instability.

2.1. The Role of Q: Linear Stability. As mentioned above, the basic, and most widely used, criterion to determine the stability of a massive disc against gravitational perturbations is related to the linear dispersion relation in the WKB approximation for an infinitesimally thin disc [30]:

$$(\omega - m\Omega)^2 = c_s^2 k^2 - 2\pi G\Sigma|k| + \kappa^2, \tag{1}$$

where ω is the frequency of the perturbation, k is the radial wave number, and m is the azimuthal wave number. The above dispersion relation is quadratic in k from which one easily sees that, for $m = 0$ (axisymmetric perturbations), ω^2 is positive (and the perturbation is stable) at all wavelengths if

$$Q = \frac{c_s\kappa}{\pi G\Sigma} > 1. \tag{2}$$

Marginal stability occurs at $Q = 1$.

Here we should note that the above (local) dispersion relation is strictly speaking only appropriate for infinitesimally thin discs and for tightly wound perturbations ($m/kR \ll 1$) for which the WKB approximation holds. Finite thickness effects generally act so as to dilute the effect of self-gravity, thus making the disc more stable and decreasing the marginal stability value of Q below unity (i.e., allowing a colder disc to remain stable). On the contrary, global perturbations are more unstable [31] thus effectively increasing the marginal stability value of Q.

For most cases considered here, the disc is close to being in Keplerian rotation, for which $\kappa = \Omega$. In this case, it is easy to show that the requirement of marginal stability ($Q \approx 1$) is equivalent to

$$\frac{M_{\rm disc}}{M} \approx \frac{H}{R}, \tag{3}$$

where $M_{\rm disc} = \pi\Sigma R^2$ is a measure of the enclosed disc mass within radius R. Thus, for marginally stable discs, "thin" and "light" on the one hand and "thick" and "massive" on the other hand are equivalent. Also note that, as mentioned above, when $Q \approx 1$, the two expressions for the disc thickness in the non-self-gravitating and in the self-gravitating regime are indeed equivalent.

AGN discs are generally quite thin, with $H/R \approx 10^{-3}$, and thus even a relatively light disc, much less massive than the central black hole, can be marginally stable. It is then easy, based on standard models of accretion discs around supermassive black holes [32], to calculate the distance from the black hole at which the disc first becomes gravitationally unstable [2, 33]. This turns out to be of the order of $10^3\, R_\bullet$ (where $R_\bullet$ is the Schwarzschild radius of the black hole), or 0.01 pc, for a $10^8\, M_\odot$ black hole. Thus, discs that extend beyond this radius are going to be gravitationally unstable: in order to determine their evolution, we need to understand the behaviour of the instability at the nonlinear stage: this is addressed in the next subsection.

2.2. The Role of β: Fragmentation versus Self-Regulation. The details of the nonlinear evolution of the gravitational instability are best understood through the use of hydrodynamical simulations, which include the disc self-gravity. However, before discussing such simulations, let us make

some preliminary remarks to guide us in the interpretation of the results of the simulation.

The very same fact that the linear stability of the disc depends on Q, which is directly proportional to the sound speed $c_s \propto T^{1/2}$ (where T is the disc temperature), offers a possible way to predict the nonlinear evolution of the system. In fact, the development of the instability will act as to feed back energy into the disc and to heat it up, thus making it more stable. In practice, the linear stability condition works as a "thermostat" for the disc, so that heating turns on only when Q drops below the marginal stability value, which we have seen is of order unity. If the thermostat works, we would expect the disc to be always close to marginal stability, at least under some conditions, in a so-called "self-regulated state" [34, 35].

From a numerical point of view, it is clear then that if we want to catch the dynamics associated with self-regulation, we need to make sure that the instability is able to feed back energy into the disc, and we should not then use isothermal simulations (such as the pioneering ones of [36]), which by constraint do not allow the disc to heat up. At the same time, we need to make sure that the disc is able to cool; otherwise, once the instability sets in, it will stabilize the disc forever (cf. the "perennial heating" problem for the spiral structure in galaxies), and we should thus also avoid pure N-body simulations, unless special arrangements are made to artificially cool the disc down [37].

One such approach has been taken by Gammie [20], who ran local, shearing-sheet simulations of razor-thin discs, which were allowed to heat up through shocks and $p\mathrm{d}V$ work and to cool down, according to a simple cooling prescription, such that

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{T}{t_{\mathrm{cool}}}, \tag{4}$$

where the cooling time t_{cool} is a free input parameter for the simulation. While more complex approaches, which consider the details of the radiative transfer within the disc [38], can certainly be adopted, such an approach should be considered as a useful "numerical experiment," in order to evaluate the disc response as a function of the main parameters, rather than as a "realistic" simulation of some particular system. Having clarified the main dependencies from the physical parameters, we may then establish the disc response in any particular system. Following this approach, a number of papers have considered the details of the process [13, 21, 24, 25], extending the simulations to full 3D and considering thus global and potentially thick configurations, as a function of the main parameters of the system, such as the disc mass and thickness. Here, I will present a summary of the main results concerning the issue of fragmentation and self-regulation of the instability. In the next subsection, I will address the important issue of the locality of the induced transport.

It turns out that the behaviour of the disc is actually determined by the ratio of t_{cool} to the dynamical time in the disc,

$$\beta = \Omega t_{\mathrm{cool}}. \tag{5}$$

It should be noted that, in most of the simulations described here [13, 20, 21, 24, 25], the parameter β is taken to be a single-free parameter for each simulation, with no dependance on either time or position in the disc. This is certainly not realistic, as in fact the cooling time should and will depend on the local microphysics associated with the disc opacity and radiative properties. These simulations should thus be regarded as simple "numerical experiments," where we test the disc response in a controlled configuration, as a function of the main parameters. For an actual, astrophysically relevant disc, we would thus calculate at any given radius the cooling properties and thus infer the disc behaviour from our controlled experiments. In doing this, care should be taken that the results are not affected by global effects (see below) or by nonlinear effects induced by a temperature dependence of the cooling rate, which has been studied by Johnson and Gammie [39] and Cossins et al. [40].

If the cooling timescale is larger than a few dynamical timescales, an initially stable (large Q) disc cools down until Q becomes of the order of unity. At this stage, the disc becomes gravitationally unstable and develops a spiral structure which provides a heating source, through compressional heating and shock dissipation, able to balance the externally imposed cooling. Once in thermal equilibrium, the disc is characterized by an approximately constant value of Q very close to marginal stability. In such a state, a spiral structure persists in the disc, to provide the required heating. Therefore, the self-regulation mechanism described above determines the disc structure and evolution. Figure 1(a) shows the result of one such simulations, where in this case $\beta = 10$ and the total disc mass $M_{\mathrm{disc}} = 0.1M$ [13]. The colour plot shows the disc surface density, in which a spiral structure is clearly seen. Figure 1(b) shows the azimuthally and vertically averaged value of Q as a function of radius, for several simulations with the same mass ratio but with different values of β, as indicated. The disc in this case extends from $R = 0.25$ to $R = 25$ in code units. It is then seen that far from the boundaries (where the density drops and Q correspondingly grows) the disc is self-regulated, with $Q \approx 1$ over a wide radial range. Cossins et al. [13] have also computed the amplitude of the perturbed surface density as a function of β. Analysis of the disc structure showed that while the cooling rate β does not influence the spectrum of wavenumbers that are excited, it does affect that amplitude of the density perturbations, such that

$$\frac{\delta\Sigma}{\Sigma} \approx \frac{1}{\sqrt{\beta}}, \tag{6}$$

which is shown in Figure 2. Thus as the cooling becomes more rapid (and thus as β decreases), the amplitude of the density perturbation increases. Similarly, it was found [13] that the spectrum of the radial wavenumber k peaks strongly where $kH = 1$, a result that can be predicted from the dispersion relation (1) but has now also been demonstrated numerically. This result is independent of both the cooling rate and the disc to star mass ratio.

The behaviour described above changes when the cooling time is decreased to smaller values [20]. In this case,

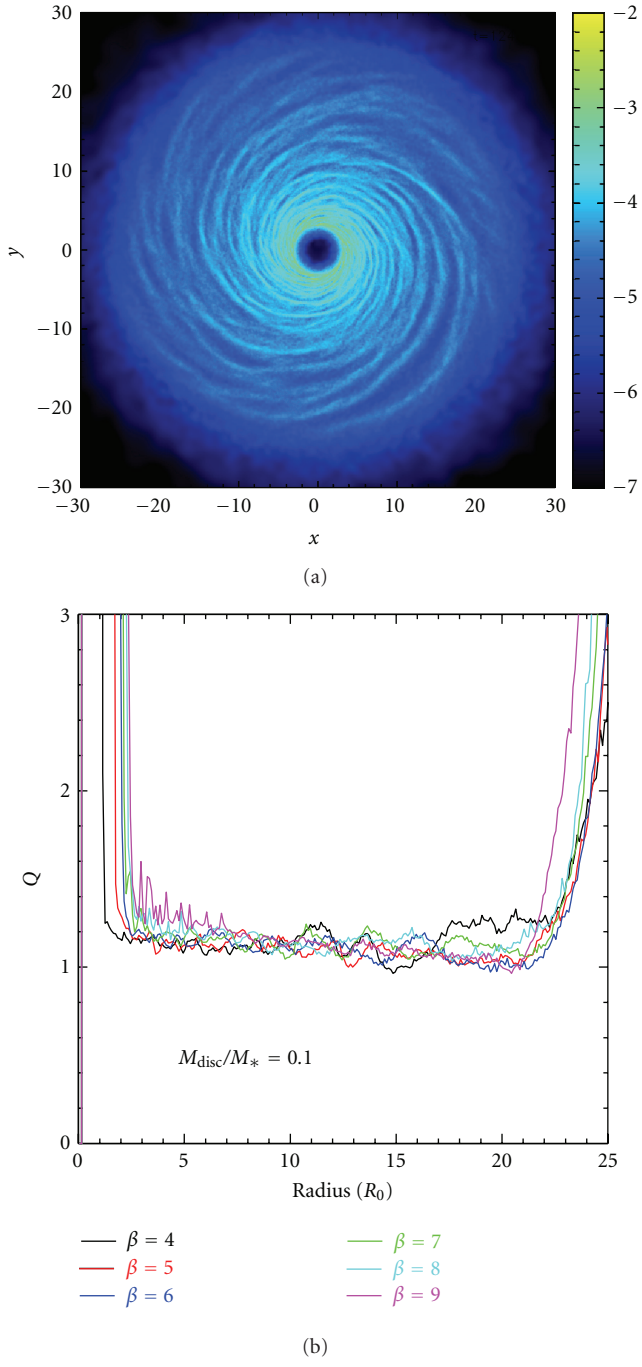


Figure 1: (a) Surface density of a self-gravitating disc, with $M_{\rm disc} = 0.1\,M$ and with $\beta = 10$, where a tightly wound spiral structure is clearly seen. (b) Azimuthally averaged profiles of Q as a function of radius, for several simulations, with varying β, as indicated. From Cossins et al. [13].

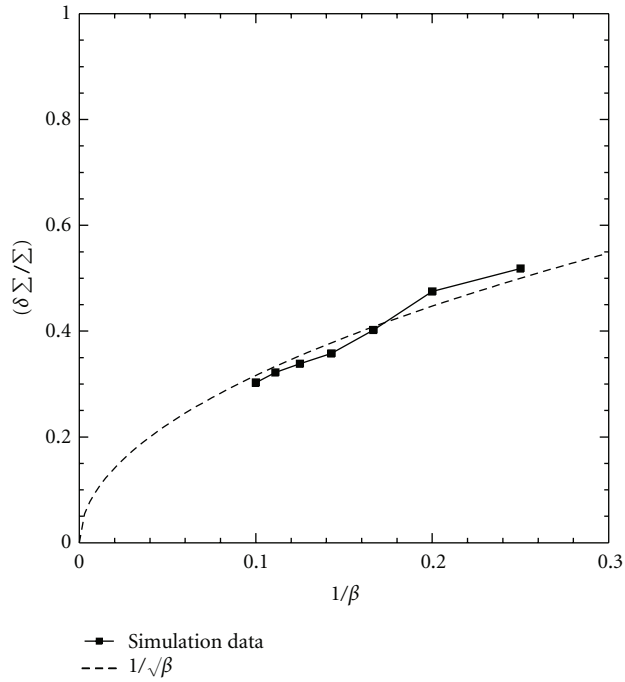


Figure 2: Variation of the radially and azimuthally averaged relative surface density perturbation amplitude $\delta\Sigma/\Sigma$ with the inverse cooling parameter $1/\beta$. From Cossins et al. [13].

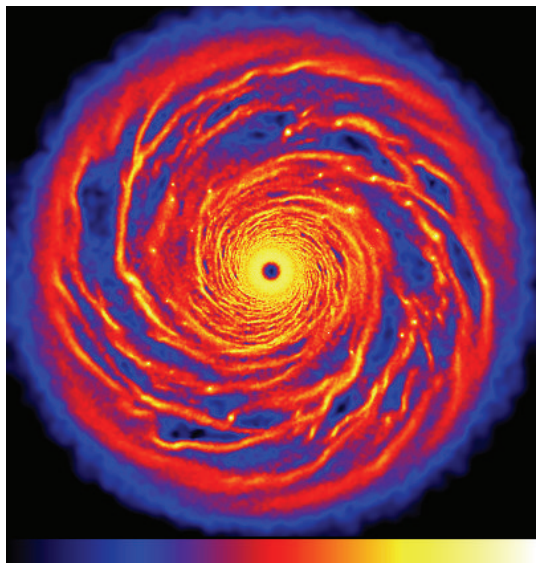


Figure 3: Numerical simulation of a self-gravitating disc with $M_{\rm disc} = 0.1\,M$ and $\beta = 3$. Once unstable, the disc breaks up into numerous gravitationally bound clumps.

the disc does not reach a quasisteady self-regulated state but rather fragments into several bound objects. Figure 3 show the results of a simulation very similar to the one displayed in Figure 1, but where the cooling time is decreased to $\beta = 3$ [24]. The presence of numerous high-density clumps is clearly seen. This result can be understood in the following way, by adopting a local approach to describe the instability. In a gravitationally unstable disc, the typical growth timescale of unstable perturbations is of the order of the dynamical timescale Ω^{-1}. The nonlinear stabilization of the perturbation only works if the heat generated by compression and shocks is not removed too efficiently from the disc through cooling. Since the perturbation grows on the

dynamical timescale, if we want to avoid fragmentation, we require that cooling acts on a longer timescale. Note that the requirement that the cooling timescale be shorter than the dynamical timescale in order to result in fragmentation has been known for several years, even outside the context of disc instability [41, 42]. Note also that the existence of a critical cooling time below which the disc fragments can be easily related, through (6), to a maximum value of the perturbation amplitude that can be sustained by the disc.

The exact value of the threshold for fragmentation does depend somewhat on the specific numerical setup and ranges from $\beta = 3$ to $\beta = 6$ [20, 24, 43]. Recently, the exact value of this threshold has been the subject of intense debate, as it was discovered [44] that the threshold value appeared to increase with increasing resolution in smoothed particle hydrodynamics (SPH) simulations (see also [45, 46]). The same behaviour has also been seen in grid-based simulations [47], and it has been shown that it actually depends on the chosen initial conditions. When carefully chosen initial conditions are used, the threshold value for fragmentation does converge and it turns out to be indeed of the order of $\beta \approx 6$.

We can easily get a reasonable estimate of the mass of the fragments. Indeed, we expect $M_{\rm frag} \approx \pi\Sigma\lambda^2$, where Σ is the local density and $\lambda \approx H$ is the typical wavelength associated with the instability. We thus obtain

$$M_{\rm frag} \approx \pi\Sigma R^2 \left(\frac{H}{R}\right)^2 \approx \left(\frac{H}{R}\right)^3 M, \tag{7}$$

where in the last equality we have used the fact that, for a marginally stable disc, $\pi\Sigma R^2 \approx (H/R)M$. For a typical AGN disc, where $H/R \approx 10^{-3}$–10^{-2} and, say, $M \approx 10^8\, M_\odot$, the fragment mass thus corresponds to 0.1–100 $M_\odot$.

2.3. The Role of $M_{\rm disc}/M$: Global versus Local Dynamics. The issue of locality of the dynamics associated with gravitational instability is essential if one wants to construct simple viscous models for self-gravitating accretion discs [48, 49]. Indeed, it has been long realized that the spiral structure determined by the instability can efficiently transport angular momentum [50], and one may thus suppose that the instability, at the large scales where an AGN disc is unstable (and where probably the disc is too cold to support MHD instabilities, such as the magnetorotational instability, MRI), can produce the required "viscous" torque to allow the accretion of matter from ∼ parsec scales down to the innermost regions where the MRI takes over and releases the accretion fuel down to the SMBH.

In the standard α-prescription for accretion disc viscosity [32], the relevant component of the viscous stress tensor $T_{R\phi}$ is simply parameterized in terms of the local pressure P, such that $T_{R\phi} \approx \alpha P$. The dimensionless parameter α is thus simply a measure of the stress tensor in units of the local pressure. One might thus be tempted to compute the stress tensor resulting from the spiral structure seen in the simulations described above and directly compute an equivalent α parameter associated to the instability. This would be obviously best done for the cases where the disc is self-regulated and the instability saturates at a given perturbation amplitude, as discussed above. However, a fundamental problem arises in this case. This is related to the fact that the gravitational instability, is an intrinsically long-range instability and it is not clear whether the transport of energy and angular momentum associated with it can be simply expressed in terms of a local viscous process [51].

The problem is best understood in terms of a WKB analysis of the energy and angular momentum fluxes associated with the instability [13]. For a local, viscous process the torque exerted on the disc $\dot{\mathcal{L}}_\alpha$ is related to the work done by viscosity $\dot{\mathcal{E}}_\alpha$ via the Keplerian rotation rate Ω, such that

$$\dot{\mathcal{E}}_\alpha = \Omega \dot{\mathcal{L}}_\alpha. \tag{8}$$

A similar but not equal relation governs the case where potentially global effects are mediated through wave transport. In a WKB analysis, the wave angular momentum and energy densities can be obtained [30], and in turn the wave-induced torque $\dot{\mathcal{L}}_w$ and power dissipation $\dot{\mathcal{E}}_w$ are found to be related via [13]

$$\dot{\mathcal{E}}_w = \Omega_p \dot{\mathcal{L}}_w, \tag{9}$$

where the pattern speed of the spiral perturbation is given by $\Omega_p = \omega/m$. The transport properties of gravitationally induced waves are therefore determined not by the rotation rate of the disc material (cf. (8)), but by the pattern speed of the density waves themselves. As these waves are excited or absorbed, the power exchanged with the background flow for a given stress is therefore significantly different than that dissipated by a viscous process that provides the same stress to the extent to which Ω_p is significantly different from Ω. The relative level of global versus local transport can hence be quantified via the parameter ξ, where

$$\xi = \left| \frac{\Omega - \Omega_p}{\Omega} \right|. \tag{10}$$

The analysis of Cossins et al. [13] also allows a spectrally averaged pattern speed to be determined, and thus in turn the nonlocal transport fraction ξ can be measured from the simulations. In agreement with Lodato and Rice [21, 25], this shows that transport by gravitational waves is a predominantly local process for the systems modeled, with $\xi \approx 10\%$ for $M_{\rm disc}/M = 0.1$ and increasing with increasing disc to star mass ratio. This is shown as a function of radius in the left-hand panels of Figure 4, where the increase in nonlocality is clearly seen with q. A corollary of this, seen from the form of (10), is that the waves remain on average close to corotation, $\Omega_p \approx \Omega$.

The right-hand panels of Figure 4 show a further interesting result obtained from the simulations of Cossins et al. [13]—the wave Mach numbers. While the heavy lines shows the values relative to an external inertial frame, the lighter lower lines give the Doppler-shifted Mach numbers $\mathcal{M}$, that is, those relative to a frame corotating with the flow. These Doppler-shifted values are almost exactly unity, implying that the density waves excited by the gravitational instability are only weakly supersonic, and furthermore this

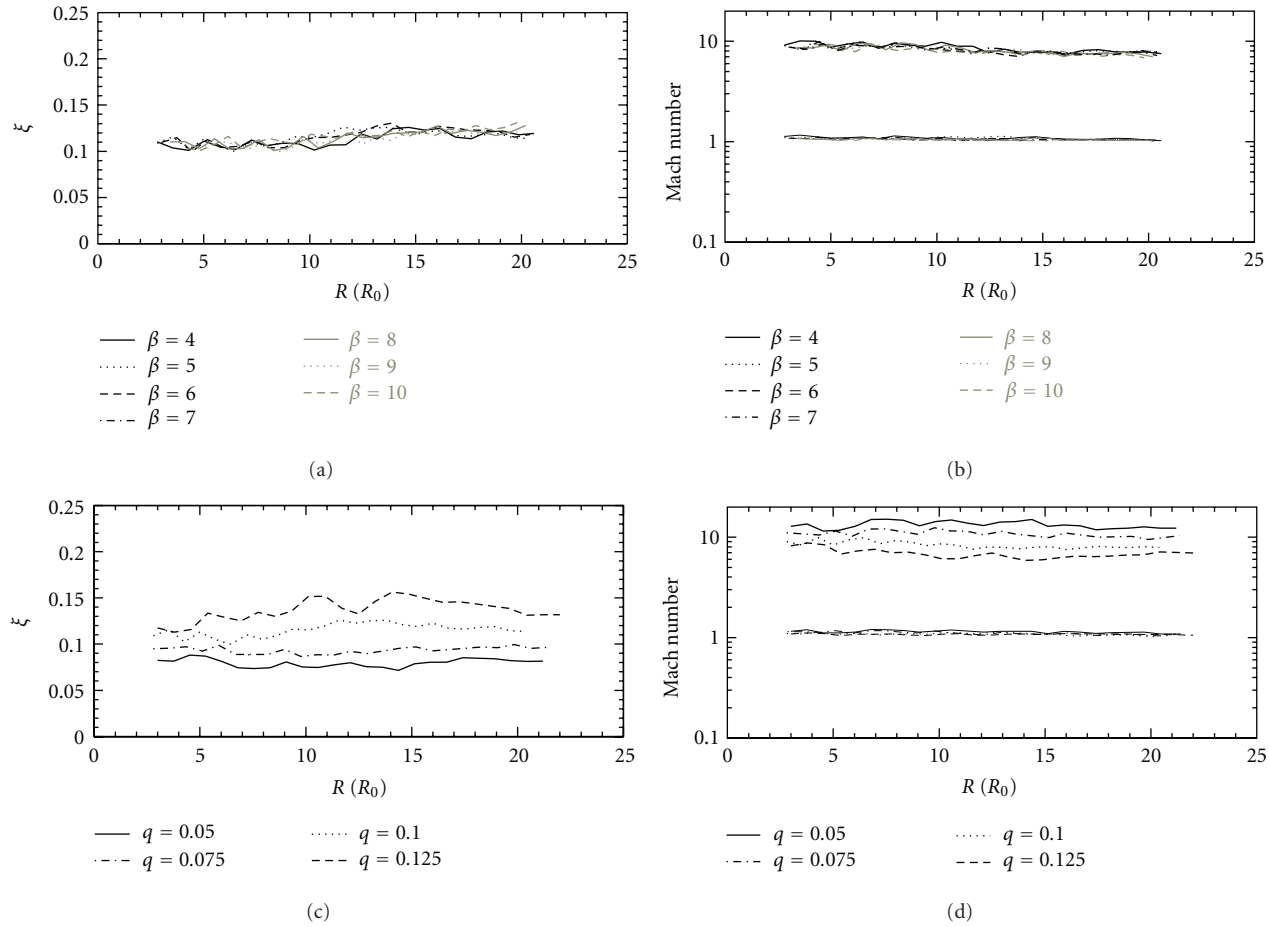


Figure 4: Nonlocal transport parameter ξ (a, c) and Mach Numbers (b, d) for simulations at various values of mass ratio $M_{\rm disc}/M$ and cooing parameter β, where β varies, $M_{\rm disc}/M = 0.1$ and where $M_{\rm disc}/M$ varies, $\beta = 5$. For the Mach number plots, heavy lines denote the Mach number as measured in an inertial (static) frame, whereas light lines show the Doppler-shifted Mach number (measured in a frame corotating with the fluid). Taken from Cossins et al. [13].

result is invariant with either cooling rate or mass ratio. The gravitational instability therefore self-regulates so that not only is $Q \approx 1$, but also $\mathcal{M} \approx 1$. This result is intuitively reasonable in a quasisteady disc—subsonic waves would not impart any net heat to the disc (any compression heating is balanced by the corresponding rarefaction cooling), whereas a strong shock would be highly dissipative, leading to rapid evolution. Furthermore, the only way a fluid element can remain on a circular orbit when passing through an oblique (spiral) shock is if that shock wave has a unit Mach number.

Actually, we can also show that the above-mentioned dependence of the factor ξ on the disc-to-star mass ratio $M_{\rm disc}/M$ can be easily understood from the condition that these waves dissipate where they are almost sonic. In fact, using this sonic condition, we can rewrite (10) as

$$\xi = \left| \frac{\Omega - \Omega_p}{\Omega} \right| \approx \frac{c_s}{v_\phi} = \frac{H}{R} \approx \frac{M_{\rm disc}}{M_\star}, \tag{11}$$

where $v_\phi = \Omega R$ is the azimuthal velocity of the disc, and the last equality holds for marginally stable discs ($Q \approx 1$). This trend can actually be seen in Figure 4(c).

Thus the assumption of a local, viscous-like process for the transport associated with gravitational instabilities is only valid for light discs, where $M_{\rm disc}/M \ll 1$. It is in such cases that one can describe the secular evolution of the disc and the associated angular momentum transport in terms of an effective viscosity, and one can even choose to measure the stress induced by the spiral structure in units of the local pressure, thus obtaining an effective $\alpha_{\rm sg}$ value associated with gravitational instabilities. So, how large is the gravitationally induced $\alpha_{\rm sg}$? Lin and Pringle [6] propose the following parameterization:

$$\alpha_{\rm sg} = \begin{cases} \eta\left(\dfrac{\overline{Q}^2}{Q^2} - 1\right), & Q < \overline{Q}, \\ 0, & Q > \overline{Q}. \end{cases} \tag{12}$$

Here $\overline{Q}$ is the value of Q at which the disc becomes unstable to nonaxisymmetric perturbations and η is a parameter to measure the strength of the induced torques. The above formulation is useful in practical cases, for example, when one wants to incorporate in a simple way the self-regulation mechanism in simple time-dependent models of

self-gravitating discs. However, it lacks one important feature elucidated from the numerical simulations described above. In this picture, α_{sg} only depends on the local value of Q and not on the cooling timescale t_{cool}, which we have seen controls so efficiently the development of the instability. In particular, for self-regulated discs, we expect $Q \approx \overline{Q}$ and the formula above would then produce a negligibly small α_{sg}, while we know that a finite amplitude spiral structure is present in self-regulated discs and indeed it is this spiral structure that provides the heating to balance the imposed cooling rate. On the other hand, we know that the process of self-regulation and the saturation of the gravitational perturbation is fundamentally related to thermal equilibrium in the disc: the saturation amplitude of the instability is such that the power dissipated through shocks in the disc is just enough to balance the imposed cooling (hence the inverse relation described above and displayed in Figure 2). In thermal equilibrium, the value of the viscosity parameter is simply related to the cooling rate [52]:

$$\alpha_{sg} = \left| \frac{\mathrm{d} \ln \Omega}{\mathrm{d} \ln R} \right|^{-2} \frac{1}{\gamma(\gamma-1)\Omega t_{cool}} = \frac{4}{9\gamma(\gamma-1)} \frac{1}{\beta}, \quad (13)$$

where the last equality holds in the case of a Keplerian disc. Indeed, the value of the stress induced by gravitational perturbation as computed directly from simulations of self-regulated discs [21, 23] agrees very well with the value predicted by (13). Thus, in a self-regulated state, not only the fractional amplitude of the density perturbations, but also the induced stress are inversely proportional to the cooling time. Indeed, one can also interpret the fragmentation threshold in terms of α_{sg} rather than in terms of β: there is a maximum value of the stress that can be supported by the disc without fragmenting [24]. Evaluating this critical α_c from (13) using the critical value of β, one finds that $\alpha_c \approx 0.05$–0.1.

Clearly, all this applies in cases where thermal equilibrium is simply established by a balance between the viscous heating and the radiative cooling. In many interesting cases (including the outskirts of AGN discs), irradiation from the central object is going to play a major role in determining the thermal balance. In such cases, (13) should be modified, and an interesting and only rarely discussed issue is what determines fragmentation: is it the stress exceeding the critical value α_c, or is it the cooling time dropping below the critical value [53]?

What happens then for the cases where the disc mass is not much smaller than the central object mass? Here, we already know that we should expect deviations from the analysis discussed above, as transport should become significantly nonlocal. Once again, a change in behaviour has been observed in simulations [25]. The stress computed form the simulations does not agree anymore with (13), exceeding its prediction and peaking at values around unity. Furthermore, in these cases we have a situation where neither self-regulation nor fragmentation occurs. The disc simply cannot find a quasistationary nonlinear saturated, state and it keeps oscillating between periods of high spiral activity, where the stress would correspond to a local α of order unity, to periods of low activity, characterized by a temporarily high value of Q.

A summary of the various possible behaviours of a self-gravitating disc as a function of the three main dimensionless parameters is displayed graphically in Figure 5. Such picture summarizes effectively the various results discussed up to now.

Having discussed the main features of the gravitational instability in gaseous discs, I now turn to the application of the above results to the process that relates to the formation and growth of supermassive black holes in galactic nuclei.

3. The Formation of Supermassive Black Hole Seeds

One of the most important applications of the concepts described in the previous section to the context of supermassive black hole growth is the formation of massive BH seeds from direct gas collapse at high redshift.

This issue has become particularly important due to the recent discovery of active quasars up to redshift $z \sim 6$ [54, 55] and now even to a redshift as high as $z \sim 7$ [56], which indicates that supermassive black holes, with masses up to $10^9 M_\odot$, were already in place when the Universe was only 10^9 years old and beyond. This clearly requires that the black hole growth occurred at very high rates, with an average of $1 M_\odot$/yr. Such a rapid early growth poses serious challenges to models of their formation.

Some models [57–59] assume that the seeds of supermassive black holes are the remnants of the zero-metallicity first stars (the so-called Population III stars), which are expected to be relatively massive [60, 61] and thus produce black holes with a mass of up to $100 M_\odot$. However, unless the efficiency of conversion of matter into energy through the accretion process is very low, it is impossible to grow the seeds to the required masses by $z \sim$ 6-7 through Eddington-limited accretion [62]. The problem here is that when the accretion rate is large, the radiation pressure produced by the accretion luminosity can exceed the gravitational force of the black hole and thus exceeding the Eddington limit. Now, if the accretion efficiency $\epsilon = L/\dot{M}c^2$ exceeds $\approx$0.1 (where L is the accretion luminosity and c is the speed of light), the Eddington limit does not allow the large accretion rates needed to grow the seeds fast enough to become bright AGN by $z \sim 6$ [62]. Note also that the Eddington limit is linearly proportional to the black hole mass, so that the problem of accreting at very high rates is particularly important in the earliest phases of the growth, when the black hole mass is small.

The efficiency is in turn dependent on the spin of the black hole, with high spin producing very large efficiencies $\epsilon \sim 0.5$. Accretion of matter naturally tends to spin up the hole [58] and hence to increase the efficiency, thus exceeding the Eddington limit for relatively low $\dot{M}$ and preventing a fast growth of the hole. While recent calculations [63, 64] show that it is possible to keep the hole spin low if the growth occurs through several small randomly oriented accretion episodes [65], we still have to face the issue of how to produce the high infall rates required.

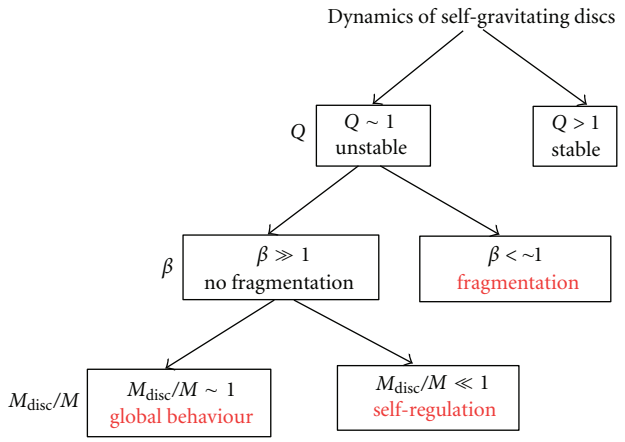


FIGURE 5: A simple diagram showing schematically the possible behaviour of a self-gravitating disc as a function of the three dimensionless parameters discussed here: Q, β, and $M_{\rm disc}/M$.

Alternative models propose the direct formation of more massive seeds with masses of about $10^5\,M_\odot$ directly out of the collapse of dense gas [66–72]. The key limiting factor for these models is the disposal of the angular momentum. Recently, it has been proposed [14, 71–73] that large-scale gravitational instabilities developing during the growth of pregalactic discs is the missing ingredient, able to funnel the required amount of gas into the center of the galaxy.

According to such models, the formation of the seeds of supermassive black holes occurs at a redshift $z \sim 10$–15, when the intergalactic medium had not been yet enriched by metals forming in the first stars. As a consequence, the chemical composition of the gas at this early epoch is essentially primordial, that is, the gas is mostly hydrogen and helium. The cooling properties of this gas are, therefore, relatively simple. In particular, in the absence of molecular hydrogen, the main coolant is provided by atomic hydrogen, for which the cooling timescale becomes extremely long for temperatures smaller than $\sim 10^4$ K, and we thus expect the gas to reach thermal equilibrium at a temperature $T_{\rm gas}$ of the order of 10^4 K.

Now, consider a dark matter halo (modeled, for simplicity, as a truncated singular isothermal sphere) of mass $M_{\rm halo}$ and circular velocity V_h, extending out to $r_h = GM_{\rm halo}/V_h^2$. We also assume that the halo contains a gas mass $M_{\rm gas} = m_d M_{\rm halo}$, where m_d is of the order of the universal baryonic fraction, ≈ 0.1, whose angular momentum is $J_{\rm gas} = j_d J$, where $j_d \sim m_d$. The angular momentum of the dark matter halo J is expressed in terms of its spin parameter $\lambda = J|E|^{1/2}/GM_{\rm halo}^{5/2}$, where E is its total energy. The probability distribution of the spin parameter of dark matter halos can be obtained from cosmological N-body simulations in Warren et al. [74] and is well described by a log-normal distribution peaking at $\lambda = 0.05$.

If the virial temperature of the halo $T_{\rm vir} \propto V_h^2$ is larger than the gas temperature $T_{\rm gas}$, the gas collapses and forms a rotationally supported disc, with circular velocity V_h, determined by the gravitational field of the halo. For low values of the spin parameter λ, the resulting disc can be compact and dense. In this case, during the infall of gas onto the disc, its density rises until the stability parameter Q becomes of the order of unity. At this point, the disc starts developing a gravitational instability, which as we have seen above is able to efficiently redistribute angular momentum and allow accretion. Further infall of gas does not cause the density to rise much further, but rather it promotes an increasingly high accretion rate into the center. This process goes on until infall is over and the disc has attained a surface density low enough to be marginally gravitationally stable, that is, with $Q = \overline{Q}$. It is then possible to calculate what fraction of the infalling mass needs to be transported into the center to make the disc marginally stable, as a function of the main parameters involved. In this way, we get [14, 73]

$$M_{\rm BH} = m_d M_{\rm halo}\left[1 - \sqrt{\frac{8\lambda}{m_d \overline{Q}}\left(\frac{j_d}{m_d}\right)\left(\frac{T_{\rm gas}}{T_{\rm vir}}\right)^{1/2}}\right], \qquad (14)$$

where I have suggestively called $M_{\rm BH}$ the accreted mass, since this mass is the total mass available for the formation of the black hole seed in the center.

However, for large halo mass, the internal torques needed to redistribute the excess baryonic mass become too large to be sustained by the disc, which might then undergo fragmentation. We have seen in the previous sections that the maximum torque that can be delivered by a quasisteady self-regulated disc is of the order of $\alpha_c \approx 0.06$. Since the infall rate of gas from the halo is proportional to $T_{\rm vir}^{3/2}$, we expect fragmentation when the virial temperature exceeds a critical value $T_{\rm max}$, given by (see [73] for details)

$$\frac{T_{\rm max}}{T_{\rm gas}} > \left(\frac{4\alpha_c}{m_d}\frac{1}{1 + M_{\rm BH}/m_d M_{\rm halo}}\right)^{2/3}. \qquad (15)$$

Although it is possible, as mentioned above, that accretion proceeds even for larger values of α in a highly time-variable way when the disc mass is large, and it is also possible that accretion proceeds even in a fragmenting disc, we make here the conservative assumption that all halos that violate, (15), do fragment and do not accrete. Figure 6 illustrates the relationship between halo mass and black hole mass based on (14) for three different values of the spin parameter λ. The red line in Figure 6 corresponds to (15), so that halos on the right of the red line are expected to fragment. We can thus see that the typical mass fed into the center of such pregalactic disc is of the order of $10^3\,M_\odot$ up to $10^5\,M_\odot$. The typical accretion rates during this early epochs is of the order of $10^{-2}\,M_\odot$/yr [73]. If such high masses are assembled as seeds of supermassive black holes at redshift 10–15, it is then easy to grow through Eddington-limited accretion to $10^9\,M_\odot$ by $z = 6$, as required by observations.

Equation (14) provides a powerful link between the properties of dark matter haloes and the mass of massive seed black holes that can grow within them. As shown, the amount of mass that will be concentrated in the central regions of these pregalactic discs depends only on halo properties (such as the spin parameter λ and the fraction of baryonic mass that collapses to the disc m_d), on the ratio

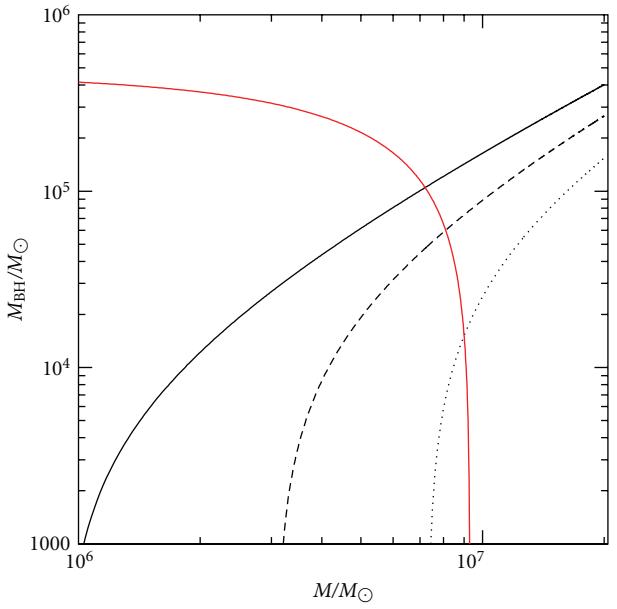


FIGURE 6: Mass available for the formation of the seed of a supermassive black hole in the center of pregalactic discs as a function of the mass of the parent dark matter halo (from [14]). The plots refer to the following choice of parameters: $\overline{Q} = 2$, $T_{\rm gas} = 4000\,\mathrm{K}$, $m_d = j_d = 0.05$, $\lambda = 0.01$ (solid line), $\lambda = 0.015$ (dashed line), and $\lambda = 0.02$ (dotted line). The red curve shows the threshold for fragmentation from (15), with $\alpha_c = 0.06$. Halos on the right of the red line give rise to fragmenting discs.

between gas temperature and halo virial temperature, and on the threshold value of Q, which has a very small range of variation around $\overline{Q} \approx 1$. This simple model has been used to calculate several properties of the black hole population at high redshift. In particular, from the distribution of halo masses and angular momentum, it is straightforward to derive the mass function of the supermassive black hole seeds Lodato and Natarajan [14], which turns out to be strongly peaked at around $10^5\,M_\odot$, as shown Figure 7(a). Furthermore, it is also possible to include such a simple prescription within evolutionary models that track the properties of the black hole population along cosmic time, such as merger tree models [15]. It is then interesting to see that the evolution of such a primordial seed population can naturally account for the current estimates of the density of black holes at low redshift (Figure 7(b)). In addition, an important and testable prediction of such models is that dwarf galaxies, which did not have any progenitor massive enough to seed a black hole, should not host a supermassive black hole. In particular, if the velocity dispersion of the galaxy is below ~ 50 km/sec, the probability of hosting a black hole turns out to be negligibly small [15].

A key requirement for the above model to work is that the gas in the disc is cooling very inefficiently; otherwise, it would rapidly fragment and form stars rather than accreting to the center (see Section 2 above). Indeed, we require the gas to be free of the main coolants such as metals and molecular hydrogen. The process outlined above is thus going to be effective only before the intergalactic medium has been sufficiently enriched by metals and only where the gas is not excessively shielded by a UV background that tends to dissociate molecular hydrogen. This has led some to propose that the above mechanism only works at specific locations which satisfy the above conditions [75, 76]. Others [77] have instead proposed that even if fragmentation does occur, it would produce a compact stellar cluster whose eventual fate is still the formation of a supermassive black hole.

It is also interesting to mention that the process described analytically above has also been simulated numerically [78], and the results appear to be in substantial agreement with the analytical expectations.

Finally, note that the models described in this section only describe how can a substantial reservoir of mass be accumulated in the innermost regions of pregalactic discs. The eventual fate of this large amount of mass is not described here. Most probably, it will form a massive object at the center of the forming galaxy, such as a "quasistar" [79], where a seed black hole grows inside a large gaseous envelope which is accreting at rates which are super-Eddington for the hole, but not for the envelope itself. A detailed discussion of the physics associated with this is clearly beyond the scope of the present contribution.

4. Gravitational Instabilities, Angular Momentum Transport, and Fragmentation in AGN Discs

4.1. AGN Discs: Fragmenting or Nonfragmenting? The situation described above changes dramatically when we consider lower redshift, that is, if we now look at the outer disc in AGNs. As mentioned above, typically, the condition of marginal stability $Q \sim 1$ first occurs at a radius of the order of 0.01 pc from the central black hole. One can easily calculate the cooling rate, and the associated cooling timescale, at this radius, to find that it is typically much smaller than the dynamical time [2, 39]. Stated otherwise, the heating rate needed to keep a marginally stable ($Q \sim 1$) disc in thermal equilibrium is much larger than what can be provided by a viscous disc with reasonable values of α [1, 80]. If we now consult the results of the numerical simulations described in Section 2, we would simply conclude that the fate of such discs is to rapidly fragment into a number of bound objects. This is often interpreted as leading to intense star formation in the disc. However, it is worth noting that the dynamical time at 0.01–1 pc from a $10^8\,M_\odot$ black hole (which is the time needed for the density perturbations to grow under the effect of the gravitational instability) (It is also easy to show that this is also the internal dynamical time of the fragments formed by instability in a $Q \sim 1$ disc.) is of the order of a few to a few thousand years. This is much shorter than the typical timescale associated with star formation in the solar neighbourhood, which is of the order of 10^6 years. Now, clearly, star formation in the Galaxy occurs under significantly different conditions, as the local molecular clouds are much less dense than the fragments

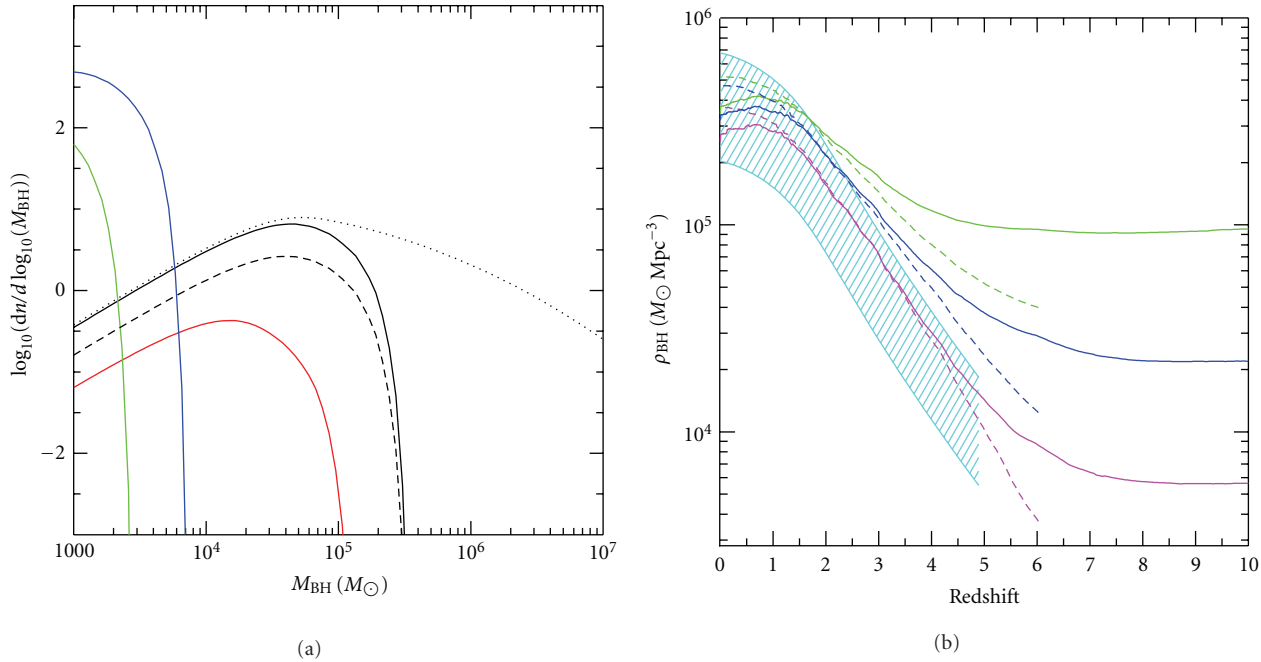


Figure 7: (a) Mass function of seed black holes predicted by the model based on (14) and (15). The black solid line refers to $z = 10$, while the red line refers to $z = 20$. The long-dashed line shows the effect of reducing $\overline{Q}$ from 2 to 1.5. The short dashed line shows the effect of not including the possibility of fragmentation (more details can be found in [14]). (b) The integrated density of black holes predicted from a merger tree evolution of the black hole seed population. The three solid lines refer to three different choices of the parameters (see [15] for details). The dashed area indicates the observationally permitted region from estimates of the black hole density at low redshifts. At high redshift ($z \gtrsim 6$), the density reflects the one attained after the seed formation phase, while the rise at lower redshifts indicates the growth through AGN activity.

produced in a fragmenting disc in this context. However, local star formation can also teach us something: indeed, the relatively long lifetime of molecular clouds in the Galaxy is due to the fact that rather than thermally supported, they are mostly supported by turbulent motions. The same might happen here: a fragmenting disc might produce a number of clouds whose dynamics is controlled by turbulent motions, which prevent their further collapse to form stars. In this case, the effective "cooling time," rather than the radiative timescale, would be the timescale for turbulence decay, which is of the order of the dynamical time and would thus be close to the threshold discussed in Section 2 [81].

A second thing to keep in mind is that if the disc extends to large distances, of the order of a parsec or so, its mass can become a significant fraction of the black hole mass and we might thus enter the regime where the dynamics associated with the gravitational instability is global rather than local. In this case, as already discussed, the energy balance should include some extra "global" terms, [51], arising from wave transport of energy, which might provide the required energy to prevent fragmentation in the outer disc. In this picture, a density wave might remove free rotational energy from the inner disc, but rather than dissipating it locally (as would a standard viscous process do), it might carry it a long way out along the wave and release it at large radii, where the wave is dissipated. As seen above, the evolution of such massive discs is generally highly variable, with episodes of strong accretion and black hole feeding followed by more quiescent periods where the accretion rate is small. Such a time variable accretion model has also been sometimes proposed by Collin and Zahn [82].

From the numerical point of view, simulations of the disc dynamics in this fragmenting regime are the most challenging, as the density in the clumps rapidly rises thus slowing down significantly the simulation. As a result, we still do not know in detail what is the fate of the disc: how much mass is turned into "stars" and what fraction of the disc mass is able to accrete to the central black hole rather than onto the forming stars [83]. Clearly, if most of the disc mass ends up in stars, it would be disastrous for black hole feeding, and we would thus conclude that only gas with very low angular momentum, which would circularize within the $Q \sim 1$ radius at 0.01 pc would accrete onto the hole. This is the basic assumption behind the chaotic accretion scenario proposed recently [84, 85].

The presence of a significant stellar component within the disc can also in principle significantly affect the overall disc dynamics. Indeed, it has been proposed [86] that even if it constitutes a minor fraction of the overall disc mass, a stellar component in the disc is able to excite low-m global spiral modes, even in a relatively low mass disc, and would thus provide a significant source of angular momentum transport, thus allowing accretion even from distances of order of several parsecs [19].

In the context delineated above, an important role is played by the evidence that has been gathered in the last few years, which points to the presence of a large number of young stars very close to the supermassive black hole at the center of our own Milky Way [87, 88]. In particular, most of these stars appear to belong to two distinct stellar discs orbiting at roughly the same distance to the black hole, that is, at a distance of 0.05–0.5 pc [8, 89]. The most likely explanation for the origin of these stars in that they formed *in situ* and in particular from the fragmentation of a self-gravitating accretion disc [89, 90]. Such observations thus fit naturally in the context described above, since we know that at parsec distances an AGN accretion disc would be self-gravitating and its cooling time is expected to be short enough to induce fragmentation. The conditions in the Galactic Center might be typical of other galaxies, where a nuclear starburst can be a result of the very same mechanism [91, 92].

4.2. Hints from Maser Dynamics. As mentioned above, there is clear observational evidence of the presence of significant mass in gas at parsec scales from the central black hole, in the form of maser emitting clumps. Such maser spots can effectively be used as a probe of the disc dynamics, as we can infer their rotation curve and hence probe the potential in the galactic nucleus. In most cases, as for example, the case of NGC 4258, the resulting rotation curve is very close to Keplerian [10], and it thus allows a very precise determination of the mass of the central BH, which for NGC 4258 is $3.6 \times 10^7\, M_\odot$ (see also the recent compilation of Keplerian rotation curves obtained through maser emission by [93]).

However, in many other cases the rotation curve, while still displaying a smooth declining profile, as would be expected for a rotating disc, does not follow exactly Kepler's law. This is, for example, the case of NGC 1068 [11, 94], of the Circinus galaxy [95], and of NGC 3079 [96]. In particular, for the case of NGC 1068, the maser data are consistent with a circular velocity $v_\phi \propto r^{-0.31}$ [11]. Given the discussion above, which shows that at a scale of a fraction of a parsec, where the maser spots are detected, the disc can be self-gravitating, it is then tempting to attribute such (often small) deviation from Keplerian rotation to the contribution of the disc self-gravity.

A detailed fit to the circular velocity traced by water masers in NGC 1068 with a model which incorporates both the gravitational field of the black hole and that of the disc has been performed by Lodato and Bertin [33], by using self-regulated models of massive discs. The resulting black hole mass is $M = (8.0 \pm 0.3)\, 10^6\, M_\odot$ and the disc mass is approximately equal to the black hole mass. From the required disc surface density, it is then possible to obtain $\dot{M} = (28.1 \pm 0.2)\, \alpha M_\odot/\mathrm{yr}$. The mass accretion rate $\dot{M}$ can be estimated, for example, from the bolometric luminosity as $\dot{M} \approx 0.23\, M_\odot/\mathrm{yr}$, and we thus obtain also an estimate of $\alpha \approx 8.3 \times 10^{-3}$, which is of the right order of magnitude as would be expected from the transport induced by gravitational instabilities.

4.3. Gravitational Instabilities and the Process of Binary Black Holes Merger. A related issue is connected to the process of black hole mergers. Black hole pairs are a natural by-product of hierarchical galaxy formation, as a consequence of the merger of two galaxies each containing a nuclear black hole. Stellar dynamical processes are able to shrink the binary down to separations of the order of 1 pc [97]. Additional gas dynamical processes can reduce the separation down to 0.1 pc or so [98, 99]. Below 0.001 pc, the emission of gravitational waves can shrink the binary further and lead to the merger of the two black holes. Such black hole mergers are indeed expected to be a primary source of gravitational radiation (and a prime target for gravitational wave detectors, such as LISA). However, an outstanding question is how to reduce the binary separation from 0.1 pc to 0.001 pc. Given the essential lack of observational evidence for sub-pc black hole binaries, we know that the process needs to be fast. It has been frequently suggested that the role of gaseous discs at sub-pc scales can provide the necessary torques to produce such fast evolution.

The problem is in several ways connected with the issues discussed above. Indeed, both in the case of mass accretion to feed a single central black hole and in the case of reducing the separation of a black hole binary, the problem is how to dispose of the large orbital angular momentum. The internal torques within an accretion disc (whether "viscous" and thus local, or globally related to gravitational instabilities) could be the natural way to remove the excess angular momentum also in the case of a binary. In reality, in the case of a binary, the angular momentum transfer process is mediated by disc tides. A secondary black hole carves an annular gap within an accretion disc. It is the gravitational force between the disc and the satellite to remove the angular momentum from the satellite and reduce the binary separation. Viscous torques within the disc are then essential in redistributing the angular momentum taken up by the gas and transport it to large radii. The circumbinary disc then evolves subject to a source of angular momentum from its inner edge, in a way that has been termed a "decretion" disc [100]. The binary evolution timescale in this case is given by [16, 101, 102]

$$t_{\mathrm{shrink}} = \frac{M_d(a) + M_s}{M_d(a)}\, t_\nu, \tag{16}$$

where t_ν is the disc viscous timescale, $M_{\mathrm{disc}}(a) = 4\pi\Sigma(a)a^2$ is a measure of the local disc mass at the binary separation a, and M_s is the mass of the secondary black hole. The dependency on t_ν indicates the fact that viscous torques in the disc are ultimately responsible for the removal of angular momentum, while the factor depending on the relative mass of the disc and of the secondary indicates that if the inertia of the secondary black hole is much larger than the disc, then the shrinking must necessarily take much longer. At 1 pc, the viscous timescale is already of the order of 10^8 years, and we thus see that if the disc mass is much less than the secondary mass, the shrinking timescale rapidly grows and can become exceedingly long for the merger to take place. Disc-assisted merger then requires large disc masses, comparable to the

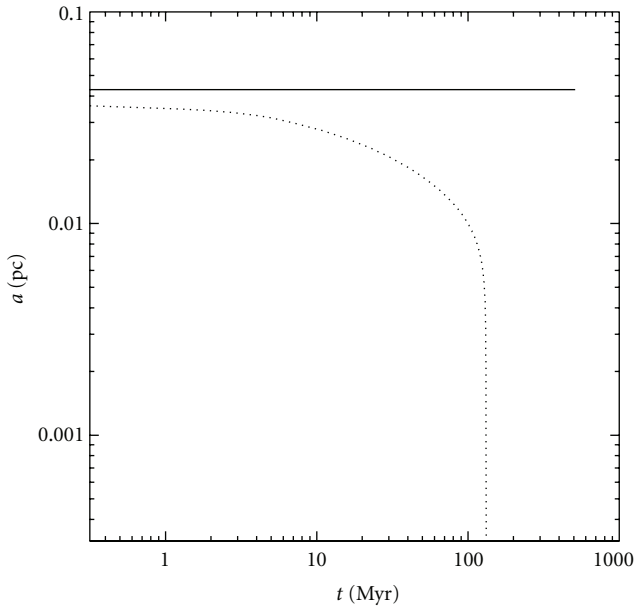


Figure 8: Evolution of the separation of a black hole binary with mass ratio $q = 0.1$. The initial separation here was $a_0 = 0.05$ pc and the disc/secondary mass ratio is $M_d/M_s = 1$. The solid curve shows the case where the disc is allowed to fragment and form stars, whereas the dotted line shows where star formation ignored. Star formation severely depletes the disc, and the remaining disc mass is not large enough to induce a black hole merger within a Hubble time. The figure taken from Lodato et al. [16].

secondary black hole, which are thus likely to be subject to gravitational instabilities.

Once again then, the dynamics of the gas disc at ~ 0.01 pc from the central black hole is essential in order to understand the evolution of the system. Lodato et al. [16] have studied the evolution of such systems. In particular, they have considered the case of a $10^8\,M_\odot$ primary black hole and of a secondary black hole with a mass ratio $q = 0.1$. They have shown that, when neglecting the possible fragmentation associated with gravitational instabilities, disc torques are able to shrink the binary and allow the merger of the two black holes within a Hubble time. On the contrary, when the disc is subject to fragmentation in the self-gravitating regime (assuming that in the self-gravitating state, enough mass is turned into stars in order to keep the disc marginally stable), the shrinking stalls and the merger does not take place. One such calculations is shown in Figure 8. Here the initial separation of the binary was 0.05 pc and the mass ratio was $q = 0.1$. The two lines refer, respectively, to the case where the disc is allowed to fragment and form stars (solid line) and where fragmentation is ignored (dotted line). The effect of fragmentation is to inhibit completely the merger process.

Fragmentation thus has a severe impact on the ability of the disc to induce a black hole merger in a reasonable time. The rapid effect of fragmentation sets an upper limit to the effective mass of a gaseous disc. To avoid the disc becoming self-gravitating, the disc must have $Q \gtrsim 1$ and so must have a mass at most $M_d/M_p \lesssim H/R$ (see Section 2). On the other hand, we must also require that the disc mass be at least comparable to the secondary mass in order for the viscous torques to be able to remove the secondary angular momentum. With a typical aspect ratio H/R of order of a few times 10^{-3}, one can conclude that disc-assisted mergers only work for mass ratios $q \lesssim 0.001$.

5. Conclusions

AGN discs become self-gravitating at a distance of about 0.01 pc from their central black hole. The development of gravitational instabilities in the disc can be both beneficial and detrimental for the process of black hole growth. Indeed, a beneficial effect is provided by the ability of gravitational instabilities to redistribute angular momentum within the disc and thus promote accretion. A detrimental effect is instead produced by the possibility of fragmentation, which could in principle turn most of the gas mass into star and thus remove it from the accretion flow. In this contribution, I have reviewed the recent progresses that we have made in the last ten years in our understanding of the nonlinear evolution of gravitational instabilities in gaseous discs, and in particular, on the parameters (most importantly, the disc cooling rate), that determine whether the instability saturates at a finite value—thus providing a quasisteady source of angular momentum transport in a self-regulated way—or rather fragments into bound objects.

Despite the impressive progresses made from the theoretical point of view, the application to AGN discs is not straightforward. Indeed, it turns out that the cooling timescale at the radius where the disc is self-gravitating is way to short to support a self-regulated state. This has led some authors to propose that AGN discs are effectively truncated at the self-gravitating radius. On the other hand, the evolution of a fragmenting disc is still not well understood, and it is not at all clear that the onset of fragmentation would totally preclude accretion. On the contrary, observational evidence such as the presence of Keplerian or quasi-Keplerian maser emitting gas at parsec scales in several AGNs, and the lack of observed black hole binaries at sub-pc scales—which in turn require the effective dynamical presence of a massive gaseous disc—hints to the importance of discs in the self-gravitating regime. Finally, it is worth noting that the presence of young stars in our own Galactic Center is indicative of the fact that star formation does effectively take place in the discs surrounding nuclear black holes.

A relatively better-understood evolution occurs at high redshifts, $z \sim 10$ or so, where the intergalactic medium was still not significantly polluted with metals and thus the cooling properties of the gas were significantly different. In these cases, fragmentation would be inhibited and the transport of angular momentum due to gravitational torques would naturally provide a way to accumulate large amounts of gas (up to $10^5\,M_\odot$) in the central regions of pregalactic discs, potentially opening the way to the rapid formation of black hole seeds by direct gas collapse.

References

[1] G. Bertin and G. Lodato, "Thermal stability of self-gravitating, optically thin accretion disks," *Astronomy and Astrophysics*, vol. 370, no. 1, pp. 342–350, 2001.

[2] J. Goodman, "Self-gravity and quasi-stellar object discs," *Monthly Notices of the Royal Astronomical Society*, vol. 339, no. 4, pp. 937–948, 2003.

[3] P. I. Kolykhalov and R. A. Sunyaev, "Disk formation through accretion of stellar wind," *Soviet Astronomy Letters*, vol. 5, pp. 180–183, 1979.

[4] S. Collin and J. P. Zahn, "Star formation and evolution in accretion disks around massive black holes: star formation and evolution in accretion disks," *Astronomy and Astrophysics*, vol. 344, no. 2, pp. 433–449, 1999.

[5] D. N. C. Lin and J. E. Pringle, "A viscosity prescription for a self-gravitating accretion disc," *Monthly Notices Royal Astronomical Society*, vol. 225, pp. 607–613, 1987.

[6] D. N. C. Lin and J. E. Pringle, "The formation and initial evolution of protostellar disks," *Astrophysical Journal*, vol. 358, no. 2, pp. 515–524, 1990.

[7] I. Shlosman, J. Frank, and M. C. Begelman, "Bars within bars: a mechanism for fuelling active galactic nuclei," *Nature*, vol. 338, no. 6210, pp. 45–47, 1989.

[8] T. Paumard, R. Genzel, F. Martins et al., "The two young star disks in the central parsec of the galaxy: properties, dynamics, and formation," *Astrophysical Journal*, vol. 643, no. 2 I, article 033, pp. 1011–1035, 2006.

[9] R. Genzel, F. Eisenhauer, and S. Gillessen, "The Galactic center massive black hole and nuclear star cluster," *Reviews of Modern Physics*, vol. 82, no. 4, pp. 3121–3195, 2010.

[10] M. Miyoshi, J. Moran, J. Herrnstein et al., "Evidence for a black hole from high rotation velocities in a sub-parsec region of NGC4258," *Nature*, vol. 373, no. 6510, pp. 127–129, 1995.

[11] L. J. Greenhill, C. R. Gwinn, R. Antonucci, and R. Barvainis, "VLBI imaging of water maser emission from the nuclear torus of NGC 1068," *Astrophysical Journal*, vol. 472, no. 1, pp. L21–L24, 1996.

[12] P. T. Kondratko, L. J. Greenhill, and J. M. Moran, "Discovery of water maser emission in five AGNs and a possible correlation between water maser and nuclear 2-10 keV luminosities," *Astrophysical Journal*, vol. 652, no. 1 I, pp. 136–145, 2006.

[13] P. Cossins, G. Lodato, and C. J. Clarke, "Characterizing the gravitational instability in cooling accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 4, pp. 1157–1173, 2009.

[14] G. Lodato and P. Natarajan, "The mass function of high-redshift seed black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L64–L68, 2007.

[15] M. Volonteri, G. Lodato, and P. Natarajan, "The evolution of massive black hole seeds," *Monthly Notices of the Royal Astronomical Society*, vol. 383, no. 3, pp. 1079–1088, 2008.

[16] G. Lodato, S. Nayakshin, A. R. King, and J. E. Pringle, "Black hole mergers: can gas discs solve the "final parsec" problem?" *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 3, pp. 1392–1402, 2009.

[17] L. Ciotti and J. P. Ostriker, "Radiative feedback from massive black holes in elliptical galaxies: agn flaring and central starburst fueled by recycled gas," *Astrophysical Journal*, vol. 665, no. 2 I, pp. 1038–1056, 2007.

[18] L. Ciotti, J. P. Ostriker, and D. Proga, "Feedback from central black holes in elliptical galaxies. III. Models with both radiative and mechanical feedback," *Astrophysical Journal*, vol. 717, no. 2, pp. 708–723, 2010.

[19] P. F. Hopkins and E. Quataert, "How do massive black holes get their gas?" *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 3, pp. 1529–1564, 2010.

[20] C. F. Gammie, "Nonlinear outcome of gravitational instability in cooling, gaseous disks," *Astrophysical Journal*, vol. 553, no. 1, pp. 174–183, 2001.

[21] G. Lodato and W. K. M. Rice, "Testing the locality of transport in self-gravitating accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 2, pp. 630–642, 2004.

[22] A. C. Mejia, R. H. Durisen, M. K. Pickett, and K. Cai, "The thermal regulation of gravitational instabilities in protoplanetary disks. II. Extended simulations with varied cooling rates," *Astrophysical Journal*, vol. 619, no. 2 I, pp. 1098–1113, 2005.

[23] A. C. Boley, A. C. Mejía, R. H. Durisen, K. Cai, M. K. Pickett, and P. D'Alessio, "The thermal regulation of gravitational instabilities in protoplanetary disks. III. Simulations with radiative cooling and realistic opacities," *Astrophysical Journal*, vol. 651, no. 1 I, pp. 517–534, 2006.

[24] W. K.M. Rice, G. Lodato, and P. J. Armitage, "Investigating fragmentation conditions in self-gravitating accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 364, no. 1, pp. L56–L60, 2005.

[25] G. Lodato and W. K. M. Rice, "Testing the locality of transport in self-gravitating accretion discs - II. The massive disc case," *Monthly Notices of the Royal Astronomical Society*, vol. 358, no. 4, pp. 1489–1500, 2005.

[26] G. Lodato, "Self-gravitating accretion discs," *Nuovo Cimento Rivista Serie*, vol. 30, no. 7, pp. 293–353, 2007.

[27] R. H. Durisen, A. P. Boss, L. Mayer, A. F. Nelson, T. Quinn, and W. K. M. Rice, "Gravitational instabilities in gaseous protoplanetary disks and implications for giant planet formation," in *Protostars and Planets V*, B. Reipurth, D. Jewitt, and K. Keil, Eds., pp. 607–622, University of Arizona Press, Tucson, Ariz, USA, 2007.

[28] G. Bertin and G. Lodato, "A class of self-gravitating accretion disks," *Astronomy and Astrophysics*, vol. 350, no. 2, pp. 694–704, 1999.

[29] A. Toomre, "On the gravitational stability of a disk of stars," *The Astrophysical Journal*, vol. 139, p. 1217, 1964.

[30] G. Bertin, *Dynamics of Galaxies*, Cambridge University Press, Cambridge, UK, 2000.

[31] J. P. Ostriker and P. J. E. Peebles, "A Numerical study of the stability of flattened galaxies: or, can cold galaxies survive?" *Astrophysical Journal*, vol. 186, pp. 467–480, 1973.

[32] N. I. Shakura and R. A. Sunyaev, "Black holes in binary systems. Observational appearance," *Astronomy & Astrophysics*, vol. 24, pp. 337–355, 1973.

[33] G. Lodato and G. Bertin, "Non-Keplerian rotation in the nucleus of NGC 1068: evidence for a massive accretion disk?" *Astronomy and Astrophysics*, vol. 398, no. 2, pp. 517–524, 2003.

[34] B. Paczyński, "A model of selfgravitating accretion disk with a hot corona," *Acta Astronomica*, vol. 28, no. 3, pp. 241–251, 1978.

[35] G. Bertin, "Self-regulated accretion disks," *Astrophysical Journal*, vol. 478, no. 2, pp. L71–L74, 1997.

[36] G. Laughlin and P. Bodenheimer, "Nonaxisymmetric evolution in protostellar disks," *Astrophysical Journal*, vol. 436, no. 1, pp. 335–354, 1994.

[37] J. A. Sellwood and R. G. Carlberg, "Spiral instabilities provoked by accretion and star formation," *Astrophysical Journal*, vol. 282, pp. 61–74, 1984.

[38] L. Mayer, G. Lufkin, T. Quinn, and J. Wadsley, "Fragmentation of gravitationally unstable gaseous protoplanetary disks with radiative transfer," *Astrophysical Journal*, vol. 661, no. 1, pp. L77–L80, 2007.

[39] B. M. Johnson and C. F. Gammie, "Nonlinear outcome of gravitational instability in disks with realistic cooling," *Astrophysical Journal*, vol. 597, no. 1 I, pp. 131–141, 2003.

[40] P. Cossins, G. Lodato, and C. Clarke, "The effects of opacity on gravitational stability in protoplanetary discs," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 4, pp. 2587–2598, 2010.

[41] M. J. Rees, "Opacity-limited hierarchical fragmentation and the masses of protostars," *Monthly Notices of the Royal Astronomical Society*, vol. 176, p. 483, 1976.

[42] J. Silk, "On the fragmentation of cosmic gas clouds. II—opacity-limited star formation," *Astrophysical Journal*, vol. 214, pp. 152–160, 1977.

[43] C. J. Clarke, E. Harper-Clark, and G. Lodato, "The response of self-gravitating protostellar discs to slow reduction in cooling time-scale: the fragmentation boundary revisited," *Monthly Notices of the Royal Astronomical Society*, vol. 381, no. 4, pp. 1543–1547, 2007.

[44] F. Meru and M. R. Bate, "Non-convergence of the critical cooling time-scale for fragmentation of self-gravitating discs," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 1, pp. L1–L5, 2011.

[45] G. Lodato and C. J. Clarke, "Resolution requirements for smoothed particle hydrodynamics simulations of self-gravitating accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 4, pp. 2735–2740, 2011.

[46] G. Lodato and P. J. Cossins, "Smoothed Particle Hydrodynamics for astrophysical flows. The dynamics of protostellar discs," *The European Physical Journal Plus*, vol. 126, no. 4, 2011.

[47] S.-J. Paardekooper, C. Baruteau, and F. Meru, "Numerical convergence in self-gravitating disc simulations: initial conditions and edge effects," *Monthly Notices of the Royal Astronomical Society*, vol. 416, no. 1, pp. L65–L69, 2011.

[48] C. J. Clarke, "Pseudo-viscous modelling of self-gravitating discs and the formation of low mass ratio binaries," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 2, pp. 1066–1074, 2009.

[49] R. R. Rafikov, "Properties of gravitoturbulent accretion disks," *Astrophysical Journal*, vol. 704, no. 1, pp. 281–291, 2009.

[50] D. Lynden-Bell and A. J. Kalnajs, "On the generating mechanism of spiral structure," *Monthly Notices of the Royal Astronomical Society*, vol. 157, p. 1, 1972.

[51] S. A. Balbus and J. C. B. Papaloizou, "On the dynamical foundations of α disks," *Astrophysical Journal*, vol. 521, no. 2, pp. 650–658, 1999.

[52] J. E. Pringle, "Accretion discs in astrophysics," *Annual Review of Astronomy and Astrophysics*, vol. 19, pp. 137–160, 1981.

[53] W. K. M. Rice, P. J. Armitage, G. Mamatsashvili, G. Lodato, and C. J. Clarke, "Stability of self-gravitating discs under irradiation," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 2, pp. 1356–1362, 2011.

[54] X. Fan, J. F. Hennawi, G. T. Richards et al., "A survey of z > 5.7 quasars in the sloan digital sky survey. III. Discovery of five additional quasars," *Astronomical Journal*, vol. 128, no. 2, pp. 515–522, 2004.

[55] X. Fan, M. A. Strauss, G. T. Richards et al., "A survey of z > 5.7 quasars in the sloan digital sky survey. IV. Discovery of seven additional quasars," *Astronomical Journal*, vol. 131, no. 3, pp. 1203–1209, 2006.

[56] D. J. Mortlock, S. J. Warren, B. P. Venemans et al., "A luminous quasar at a redshift of z = 7.085," *Nature*, vol. 474, no. 7353, pp. 616–619, 2011.

[57] Z. Haiman and A. Loeb, "Observational signatures of the first quasars," *Astrophysical Journal*, vol. 503, no. 2, pp. 505–517, 1998.

[58] M. Volonteri, P. Madau, E. Quataert, and M. J. Rees, "The distribution and cosmic evolution of massive black hole spins," *Astrophysical Journal*, vol. 620, no. 1 I, pp. 69–77, 2005.

[59] J. S.B. Wyithe and A. Loeb, "Constraints on the process that regulates the growth of supermassive black holes based on the intrinsic scatter in the $M_{bh-\sigma sph}$ relation," *Astrophysical Journal*, vol. 634, pp. 910–920, 2005.

[60] T. Abel, G. L. Bryan, and M. L. Norman, "The formation and fragmentation of primordial molecular clouds," *Astrophysical Journal*, vol. 540, no. 1, pp. 39–44, 2000.

[61] V. Bromm, P. S. Coppi, and R. B. Larson, "The formation of the first stars. I. The primordial star-forming cloud," *Astrophysical Journal*, vol. 564, no. 1 I, pp. 23–51, 2002.

[62] M. Volonteri and M. J. Rees, "Rapid growth of high-redshift black holes," *Astrophysical Journal*, vol. 633, no. 2 I, pp. 624–629, 2005.

[63] A. R. King, S. H. Lubow, G. I. Ogilvie, and J. E. Pringle, "Aligning spinning black holes and accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 363, no. 1, pp. 49–56, 2005.

[64] G. Lodato and J. E. Pringle, "The evolution of misaligned accretion discs and spinning black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 3, pp. 1196–1208, 2006.

[65] A. R. King and J. E. Pringle, "Growing supermassive black holes by chaotic accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 1, pp. L90–L92, 2006.

[66] M. G. Haehnelt and M. J. Rees, "The formation of nuclei in newly formed galaxies and the evolution of the quasar population," *Monthly Notices Royal Astronomical Society*, vol. 263, no. 1, pp. 168–178, 1993.

[67] M. Umemura, A. Loeb, and E. L. Turner, "Early cosmic formation of massive black holes," *Astrophysical Journal*, vol. 419, no. 2, pp. 459–468, 1993.

[68] A. Loeb and F. A. Rasio, "Collapse of primordial gas clouds and the formation of quasar black holes," *Astrophysical Journal*, vol. 432, no. 1, pp. 52–61, 1994.

[69] D. J. Eisenstein and A. Loeb, "Origin of quasar progenitors from the collapse of low-spin cosmological perturbations," *Astrophysical Journal*, vol. 443, no. 1, pp. 11–17, 1995.

[70] V. Bromm and A. Loeb, "Formation of the first supermassive black holes," *Astrophysical Journal*, vol. 596, no. 1 I, pp. 34–46, 2003.

[71] S. M. Koushiappas, J. S. Bullock, and A. Dekel, "Massive black hole seeds from low angular momentum material," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 1, pp. 292–304, 2004.

[72] M. C. Begelman, M. Volonteri, and M. J. Rees, "Formation of supermassive black holes by direct collapse in pre-galactic haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 1, pp. 289–298, 2006.

[73] G. Lodato and P. Natarajan, "Supermassive black hole formation during the assembly of pre-galactic discs," *Monthly*

Notices of the Royal Astronomical Society, vol. 371, no. 4, pp. 1813–1823, 2006.

[74] M. S. Warren, P. J. Quinn, J. K. Salmon, and W. H. Zurek, "Dark halos formed via dissipationless collapse. I. Shapes and alignment of angular momentum," *Astrophysical Journal*, vol. 399, no. 2, pp. 405–425, 1992.

[75] M. Dijkstra, Z. Haiman, A. Mesinger, and J. S. B. Wyithe, "Fluctuations in the high-redshift Lyman-Werner background: close halo pairs as the origin of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 391, no. 4, pp. 1961–1972, 2008.

[76] C. Shang, G. L. Bryan, and Z. Haiman, "Supermassive black hole formation by direct collapse: keeping protogalactic gas H_2 free in dark matter haloes with virial temperatures $T_{\rm vir}$ > rsim 10^4 K," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 2, pp. 1249–1262, 2010.

[77] B. Devecchi and M. Volonteri, "Formation of the first nuclear clusters and massive black holes at high redshift," *Astrophysical Journal*, vol. 694, no. 1, pp. 302–313, 2009.

[78] L. Mayer, S. Kazantzidis, A. Escala, and S. Callegari, "Direct formation of supermassive black holes via multi-scale gas inflows in galaxy mergers," *Nature*, vol. 466, no. 7310, pp. 1082–1083, 2010.

[79] M. C. Begelman, E. M. Rossi, and P. J. Armitage, "Quasi-stars and the cosmic evolution of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 387, no. 4, pp. 1649–1659, 2008.

[80] E. Sirko and J. Goodman, "Spectral energy distributions of marginally self-gravitating quasi-stellar object discs," *Monthly Notices of the Royal Astronomical Society*, vol. 341, no. 2, pp. 501–508, 2003.

[81] M. C. Begelman and I. Shlosman, "Angular momentum transfer and lack of fragmentation in self-gravitating accretion flows," *Astrophysical Journal*, vol. 702, no. 1, pp. L5–L8, 2009.

[82] S. Collin and J. P. Zahn, "Star formation in accretion discs: from the Galactic center to active galactic nuclei," *Astronomy and Astrophysics*, vol. 477, no. 2, pp. 419–435, 2008.

[83] S. Nayakshin, J. Cuadra, and V. Springel, "Simulations of star formation in a gaseous disc around Sgr A^*—a failed active galactic nucleus," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 1, pp. 21–33, 2007.

[84] A. R. King and J. E. Pringle, "Fuelling active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 1, pp. L25–L28, 2007.

[85] A. R. King, J. E. Pringle, and J. A. Hofmann, "The evolution of black hole mass and spin in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 3, pp. 1621–1627, 2008.

[86] P. F. Hopkins and E. Quataert, "An analytic model of angular momentum transport by gravitational torques: from galaxies to massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 2, pp. 1027–1050, 2011.

[87] R. Genzel, R. Schödel, T. Ott et al., "The stellar cusp around the supermassive black hole in the Galactic center," *Astrophysical Journal*, vol. 594, no. 2 I, pp. 812–832, 2003.

[88] A. M. Ghez, S. Salim, S. D. Hornstein et al., "Stellar orbits around the galactic center black hole," *Astrophysical Journal*, vol. 620, no. 2 I, pp. 744–757, 2005.

[89] Y. Levin and A. M. Beloborodov, "Stellar disk in the Galactic center: a remnant of a dense accretion disk?" *Astrophysical Journal*, vol. 590, no. 1, pp. L33–L36, 2003.

[90] S. Nayakshin and R. Sunyaev, "The "missing" young stellar objects in the central parsec of the Galaxy: evidence for star formation in a massive accretion disc and a top-heavy initial mass function," *Monthly Notices of the Royal Astronomical Society*, vol. 364, no. 1, pp. L23–L27, 2005.

[91] S. Nayakshin, "Massive stars in subparsec rings around galactic centres," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 1, pp. 143–150, 2006.

[92] Y. Levin, "Starbursts near supermassive black holes: young stars in the Galactic Centre, and gravitational waves in LISA band," *Monthly Notices of the Royal Astronomical Society*, vol. 374, no. 2, pp. 515–524, 2007.

[93] C. Y. Kuo, J. A. Braatz, J. J. Condon et al., "The megamaser cosmology project. III. Accurate masses of seven supermassive black holes in active galaxies with circumnuclear megamaser disks," *Astrophysical Journal*, vol. 727, no. 1, 2011.

[94] L. J. Greenhill and C. R. Gwinn, "VLBI imaging of water maser emission from a nuclear disk in NGC 1068," *Astrophysics and Space Science*, vol. 248, no. 1-2, pp. 261–267, 1997.

[95] L. J. Greenhill, P. T. Kondratko, J. E. J. Lovell et al., "The discovery of H_2O maser emission in seven active galactic nuclei and at high velocities in the Circinus galaxy," *Astrophysical Journal*, vol. 582, no. 1, pp. L11–L14, 2003.

[96] P. T. Kondratko, L. J. Greenhill, and J. M. Moran, "Evidence for a geometrically thick self-gravitating accretion disk in NGC 3079," *Astrophysical Journal*, vol. 618, no. 2 I, pp. 618–634, 2005.

[97] M. Milosavljević and D. Merritt, "Formation of galactic nuclei," *Astrophysical Journal*, vol. 563, no. 1, pp. 34–62, 2001.

[98] M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruszkowski, and F. Haardt, "Dual black holes in merger remnants - II. Spin evolution and gravitational recoil," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 682–690, 2010.

[99] M. Dotti, M. Colpi, F. Haardt, and L. Mayer, "Supermassive black hole binaries in gaseous and stellar circumnuclear discs: orbital dynamics and gas accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 3, pp. 956–962, 2007.

[100] J. E. Pringle, "Self-induced warping of accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 281, no. 1, pp. 357–361, 1996.

[101] D. Syer and C. J. Clarke, "Satellites in disks: regulating the accretion luminosity," *Monthly Notices of the Royal Astronomical Society*, vol. 277, p. 758, 1995.

[102] P. B. Ivanov, J. C. B. Papaloizou, and A. G. Polnarev, "The evolution of a supermassive binary caused by an accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 307, no. 1, pp. 79–90, 1999.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 486750, 9 pages
doi:10.1155/2012/486750

Review Article

Testing the No-Hair Theorem with Sgr A*

Tim Johannsen

Physics Department, University of Arizona, 1118 E. 4th Street, Tucson, AZ 85721, USA

Correspondence should be addressed to Tim Johannsen, timj@physics.arizona.edu

Received 16 May 2011; Accepted 5 July 2011

Academic Editor: Francesco Shankar

The no-hair theorem characterizes the fundamental nature of black holes in general relativity. This theorem can be tested observationally by measuring the mass and spin of a black hole as well as its quadrupole moment, which may deviate from the expected Kerr value. Sgr A*, the supermassive black hole at the center of the Milky Way, is a prime candidate for such tests thanks to its large angular size, high brightness, and rich population of nearby stars. In this paper, I discuss a new theoretical framework for a test of the no-hair theorem that is ideal for imaging observations of Sgr A* with very long baseline interferometry (VLBI). The approach is formulated in terms of a Kerr-like spacetime that depends on a free parameter and is regular everywhere outside of the event horizon. Together with the results from astrometric and timing observations, VLBI imaging of Sgr A* may lead to a secure test of the no-hair theorem.

1. Introduction

According to the no-hair theorem, black holes are uniquely characterized by their masses and spins and are described by the Kerr metric [1–6]. Mass M and spin J are the first two multipole moments of the Kerr spacetime, and all higher-order moments can be expressed in terms of these two [7, 8]. The no-hair theorem, then, naturally leads to the expectation that all astrophysical black holes are Kerr black holes. To date, however, a definite proof for the existence of such black holes is still lacking despite a wealth of observational evidence (see discussion in, e.g., [9]).

Tests of the no-hair theorem have been suggested using observations in either the gravitational-wave [10–21] or the electromagnetic spectrum [22–31]. Both approaches are based on parametric frameworks that contain one or more free parameters in addition to mass and spin which measure potential deviations from the Kerr metric [18–20, 32–34]. If no deviation is detected, then the compact object is indeed a Kerr black hole. However, since such deviations can have a significant impact on the observed signals, the no-hair theorem may be tested in a twofold manner: if a deviation is measured to be nonzero and if general relativity is assumed, the object cannot be a black hole [18, 35]. Alternatively, if the object is otherwise known to possess an event horizon, it is a black hole, but different from a Kerr black hole. In the latter case, the no-hair theorem would be falsified [22].

Sgr A*, the supermassive black hole at the center of the Milky Way, is a prime target for testing strong-field gravity and the no-hair theorem with electromagnetic observations (see [36] for a review). Monitoring the orbits of stars around this compact object for more than a decade has led to precise mass and distance measurements making Sgr A* the black hole with the largest angular size in the sky [37, 38]. In addition, very long baseline interferometric observations have resolved Sgr A* on event horizon scales [39]. On the theoretical side, there have been significant advances recently in the development of a framework within which the search for violations of the no-hair theorem can be carried out.

In this paper, I review this framework as well as the prospects for an observational test of the no-hair theorem with Sgr A*.

2. An Ideal Framework for Testing the No-Hair Theorem

Spacetimes of rotating stellar objects in general relativity have been studied for several decades. Due to the nonlinearity of Einstein field equations, the construction of such metrics is plagued with sometimes incredible technical challenges.

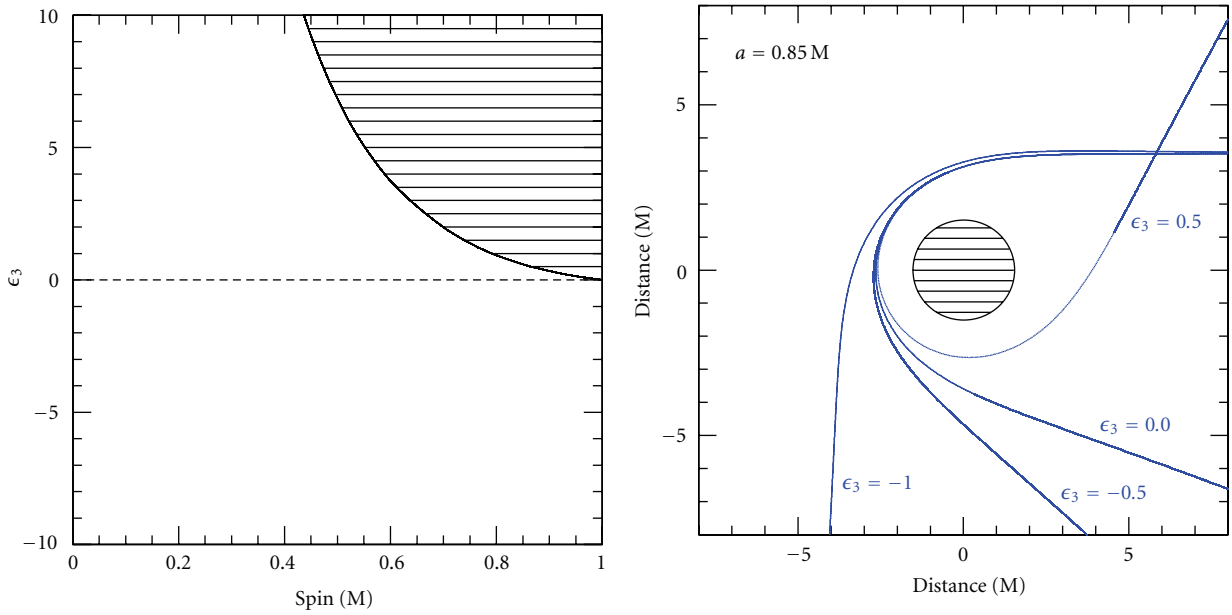


FIGURE 1: (a) Values of the parameter ϵ_3 versus the spin a, for which the central object is a black hole. The shaded region marks the excluded part of the parameter space where this object is a naked singularity. The dashed line corresponds to a Kerr black hole [34]. (b) Trajectories of photons lensed by a black hole with a (counterclockwise) spin $a = 0.85\,\mathrm{M}$ for several values of the parameter ϵ_3. The shaded region corresponds to the event horizon of a Kerr black hole of equal spin.

Following the discovery of the Schwarzschild [40] and Kerr metrics [41] in 1916 and 1963, respectively, Hartle and Thorne [42, 43] constructed a metric for slowly rotating neutron stars that is appropriate up to the quadrupole order. Tomimatsu and Sato [44, 45] found a discrete family of spacetimes in 1972 that contain the Kerr metric as a special case. After a full decade of research, Manko and Novikov [32] found two classes of metrics in 1992 that are characterized by an arbitrary set of multipole moments. Many exact solutions of the Einstein field equations are now known [46]. Of particular interest is the subclass of stationary, axisymmetric, vacuum (SAV) solutions of the Einstein equations, and especially those metrics within this class that are also asymptotically flat. Once an explicit SAV has been found, all SAVs can in principle be generated by a series of HKX-transformations ([47, 48] and references therein), which form an infinite-dimensional Lie group [49, 50]. Each SAV is fully and uniquely specified by a set of scalar multipole moments [51, 52] and can also be generated from a given set of multipole moments [53, 54]. These solutions, however, are generally very complicated and often unphysical. For some astrophysical applications, such as the study of neutron stars, it is oftentimes more convenient to resort to a numerical solution of the field equations [55–59].

To date, there exist seven different approaches that model parametric deviations from the Kerr metric. Ryan [10–12] studied the motion of test particles in the equatorial plane of compact objects with a general expansion in Geroch-Hansen multipoles. Collins and Hughes [18], Vigeland and Hughes [19], and Vigeland et al. [33] constructed Schwarzschild and Kerr metrics with perturbations in the form of Weyl sector bumps. Glampedakis and Babak [20] designed a metric starting from the Hartle-Thorne metric [42, 43] that deviates from the Kerr metric by an independent quadrupole moment. Gair et al. [21] applied a similar technique to the Manko-Novikov metric [32] affecting the quadrupole as well as higher-order moments. Sopuerta and Yunes [60] found a metric for a slowly rotating black hole that violates parity. Vigeland et al. [33] designed parametric deviations from the Kerr metric that possess four integrals of the motion and, hence, allow for the full separability of the Hamilton-Jacobi equations. Finally, Johannsen and Psaltis constructed a metric of a rapidly rotating Kerr-like black hole [34]. Other metrics of static black holes in alternative theories of gravity have also been found (e.g., [61–64]).

Due to the no-hair theorem, the Kerr metric is the only asymptotically flat SAV in general relativity with an event horizon but no closed timelike loops [1–6]. Consequently, any parametric deviation within general relativity has to violate at least one of these prerequisites and introduces either singularities or regions with closed timelike loops outside of the event horizon, which usually occur very near to the central object at radii $r \lesssim 2\,\mathrm{M}$ [65]. Otherwise, these metrics would render the no-hair theorem false. The relevance of this kind of pathologies depends on the astrophysical application. They play no role for tests of the no-hair theorem that only involve the orbits of objects at large distances from the horizon, as is the case for extreme mass-ratio inspirals or the motion of stars or pulsars around a black hole. They are, however, critical for the study

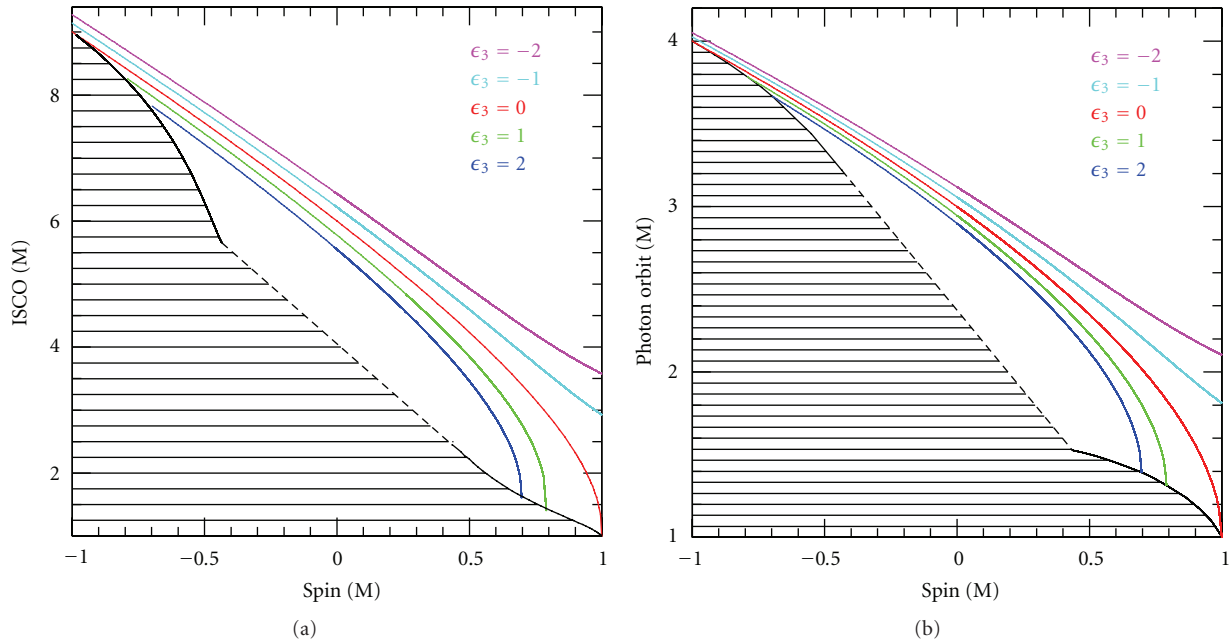


FIGURE 2: Radius of (a) the ISCO and of (b) the circular photon orbit as a function of the spin a for several values of the parameter ϵ_3. The radius of the ISCO and the circular photon orbit decrease with increasing values of the parameter ϵ_3. The shaded region marks the excluded part of the parameter space [34].

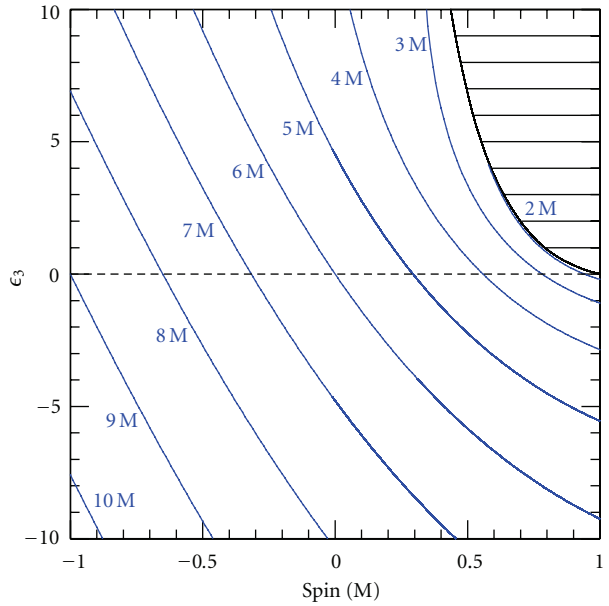


FIGURE 3: Contours of constant radius of the ISCO for values of the spin $-1 \leq a/\mathrm{M} \leq 1$ and of the parameter $-10 \leq \epsilon_3 \leq 10$. The radius of the ISCO decreases for increasing values of the spin and the parameter ϵ_3. The shaded region marks the excluded part of the parameter space. The dashed line corresponds to the parameter space for a Kerr black hole [34].

of accretion flows around black holes [34], because the electromagnetic radiation originates predominantly from the immediate vicinity of the event horizon.

For this reason, the emission from accretion flows around black holes is most interesting for strong-field tests of the no-hair theorem with observations across the electromagnetic spectrum ranging from X-ray observations of quasiperiodic variability, fluorescent iron lines, or continuum disk spectra [22, 24] to sub-mm imaging of supermassive black holes with VLBI [22, 23]. All of these observation techniques critically depend on the location of either the circular photon orbit or the innermost stable circular orbit (ISCO), because these orbits dominate the characteristics of the received signals.

These strong-field tests of the no-hair theorem require a very careful modeling of the inner region of the spacetime of black holes. Due to the pathologies of previously known parametric deviations, it has been necessary to impose an artificial cutoff at some radius outside of the event horizon that encloses all of the above pathologies and, thereby, shields them from the observer. Therefore, the application of parametric frameworks to such tests of the no-hair theorem in the electromagnetic spectrum has, so far, been limited to only slowly to moderately spinning black holes, for which the circular photon orbit and ISCO are still located outside of the cutoff radius [34, 65].

Recently [34], we constructed a black hole metric that is regular everywhere outside of the event horizon for all values of the spin within the allowable range and that depends on a set of free parameters in addition to mass and spin. In the case when all parameters vanish, our metric reduces smoothly to the Kerr metric. Our metric is a vacuum solution of a more general set of field equations, but otherwise fulfills all of the prerequisites of the no-hair theorem and, therefore, preserves these essential properties even if the deviation parameters from the Kerr metric are nonzero. At present, our metric constitutes the only known black hole spacetime of this kind and serves as an ideal framework for the study of

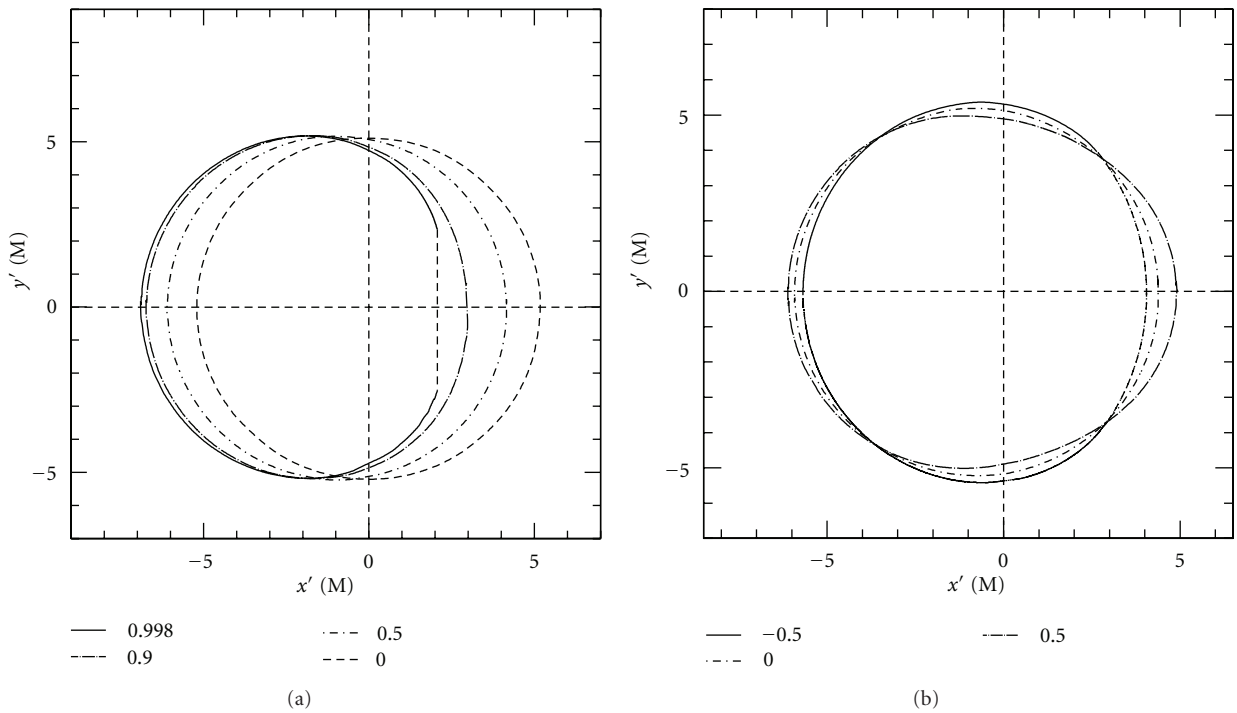


Figure 4: Images of rings of light of (a) a Kerr and (b) a quasi-Kerr black hole at an inclination $\cos i = 0.25$. Increasing values of the spin cause a displacement of the ring in the image plane, but the ring remains (nearly) circular for values of the spin $a \lesssim 0.9\,\mathrm{M}$ (a). Nonzero values of the deviation parameter ϵ (b) lead to an asymmetric ring image [23].

the signatures of a possible violation of the no-hair theorem from astrophysical phenomena near the event horizon of a black hole and, in particular, of Sgr A*. For the case of one additional parameter ϵ_3, our metric in Boyer-Lindquist coordinates is given by the expression [34]

$$\begin{aligned} ds^2 = &-[1+h(r,\theta)]\left(1-\frac{2Mr}{\Sigma}\right)dt^2 - \frac{4aMr\sin^2\theta}{\Sigma} \\ &\times [1+h(r,\theta)]dtd\phi \\ &+\frac{\Sigma[1+h(r,\theta)]}{\Delta+a^2\sin^2\theta h(r,\theta)}dr^2+\Sigma d\theta^2 \\ &+\left[\sin^2\theta\left(r^2+a^2+\frac{2a^2Mr\sin^2\theta}{\Sigma}\right)\right. \\ &\left.+h(r,\theta)\frac{a^2(\Sigma+2Mr)\sin^4\theta}{\Sigma}\right]d\phi^2, \end{aligned} \tag{1}$$

where

$$\begin{aligned} \Delta &\equiv r^2 - 2Mr + a^2, \\ \Sigma &\equiv r^2 + a^2\cos^2\theta, \\ h(r,\theta) &\equiv \epsilon_3\frac{M^3 r}{\Sigma^2}, \end{aligned} \tag{2}$$

and $a \equiv J/M$ is the spin parameter.

In [34], we analyzed several of the key properties of our black hole metric as a function of the mass M, the spin a, and the parameter ϵ_3. The left panel in Figure 1 shows the range of the spin and the parameter ϵ_3, for which our metric describes a black hole. The shaded region marks the part of the parameter space where the event horizon is no longer closed, and the black hole becomes a naked singularity. The right panel of Figure 1 shows the gravitational lensing experienced by photons on an orbit in the equatorial plane that approach the black hole closely for several values of the deviation parameter ϵ_3.

In Figure 2, we plot the radius of the ISCO and of the circular photon orbit, respectively, as a function of the spin for several values of the parameter ϵ_3. The location of both orbits decreases with increasing values of the spin and of the parameter ϵ_3. In Figure 3, we plot contours of constant ISCO radius as a function of the spin and the parameter ϵ_3. The location of these orbits depends significantly on the value of the deviation parameter ϵ_3.

3. Testing the No-Hair Theorem with VLBI Imaging of Sgr A*

In [23], we explored in detail the effects of a violation of the no-hair theorem for VLBI imaging using a quasi-Kerr metric [20]. This metric can be used to accurately describe Kerr-like black holes up to a spin of about $a \leq 0.4\,\mathrm{M}$. For Sgr A*, this spin range might already be sufficient ($a \leq 0.3\,\mathrm{M}$; [66, 67]).

The location of the circular photon orbit determines the size of the shadow of Sgr A* (see [68, 69]). VLBI observations are expected to image the shadow of Sgr A* and to measure the mass, spin, and inclination of this black hole (e.g., [69–73]). In addition to these parameters, the shape of the

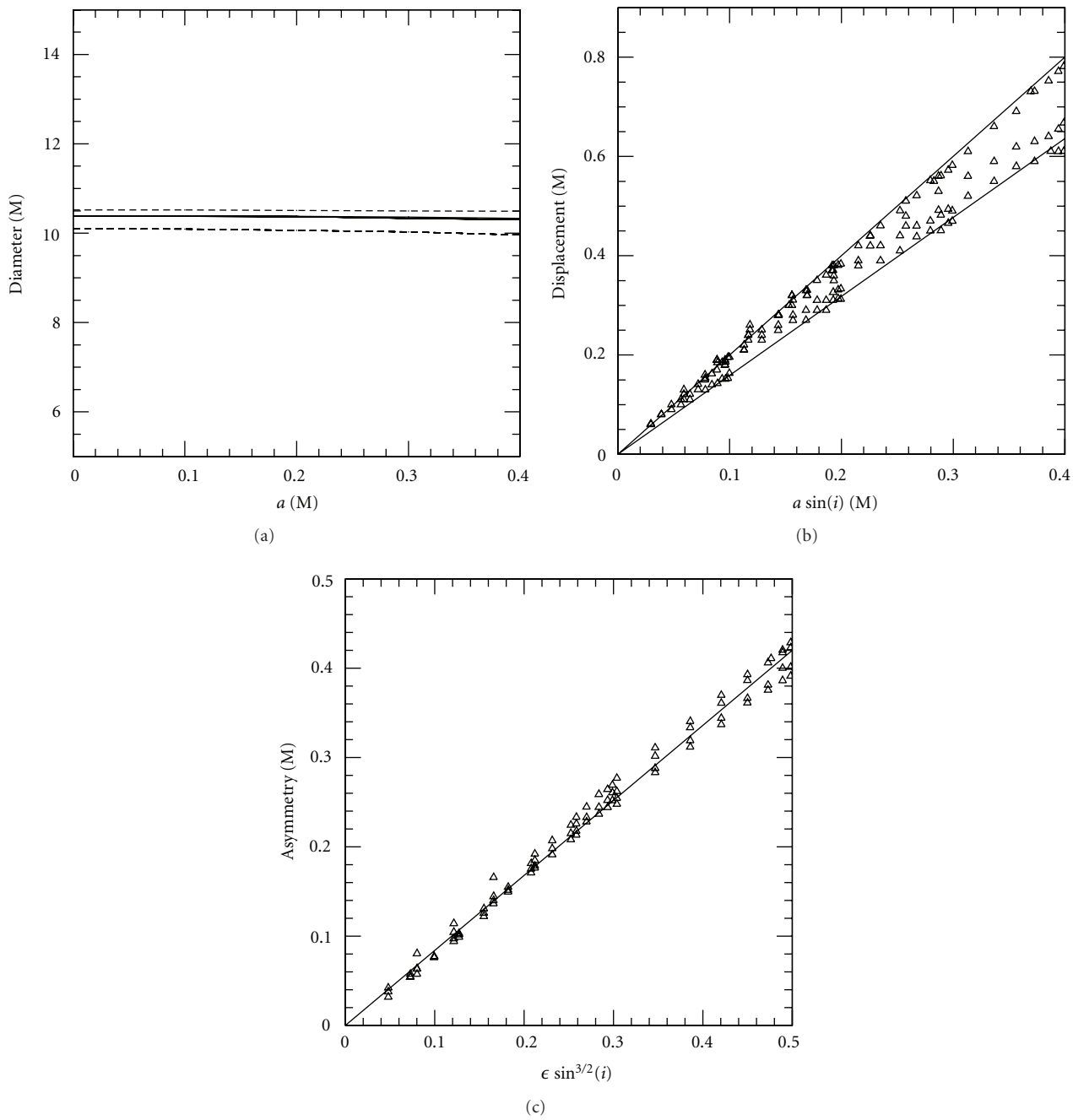


FIGURE 5: (a) The ring diameter versus spin for inclinations $17^\circ \leq i \leq 86^\circ$ for a Kerr black hole (solid lines) and for a quasi-Kerr black hole with a value of the deviation parameter $\epsilon = 0.5$ (dashed lines). The diameter is practically independent of the spin, inclination, and deviation parameter with a constant value of $\simeq 10.4\,\mathrm{M}$ for a Kerr black hole. (b) The displacement of the ring of light as a function of $a \sin i$ for various values of the parameter $0 \leq \epsilon \leq 0.5$. The displacement depends only weakly on the parameter ϵ. (c) The ring asymmetry versus $\epsilon \sin^{3/2} i$ for various inclinations $17^\circ \leq i \leq 86^\circ$ and $0.0 \leq a/\mathrm{M} \leq 0.4$. The asymmetry is nearly independent of the spin and hence provides a direct measure of a violation of the no-hair theorem [23].

shadow also depends uniquely on the value of the deviation parameter [23]. In practice, however, these measurements will be model dependent (e.g., [74]) and affected by finite telescope resolution (e.g., [69, 75]). Therefore, VLBI imaging may have to be complemented by additional observations such as a multiwavelength study of polarization ([73]; see also [76, 77]).

In an optically thin accretion flow such as the one around Sgr A* at sub-mm wavelengths (e.g., [66]), photons can orbit around the black hole several times before they are detected by a distant observer. This produces an image of a ring that can be significantly brighter than the underlying flow thanks to the long optical path of the contributing photons (e.g., [78]). In [23], we showed that the shape and location of this

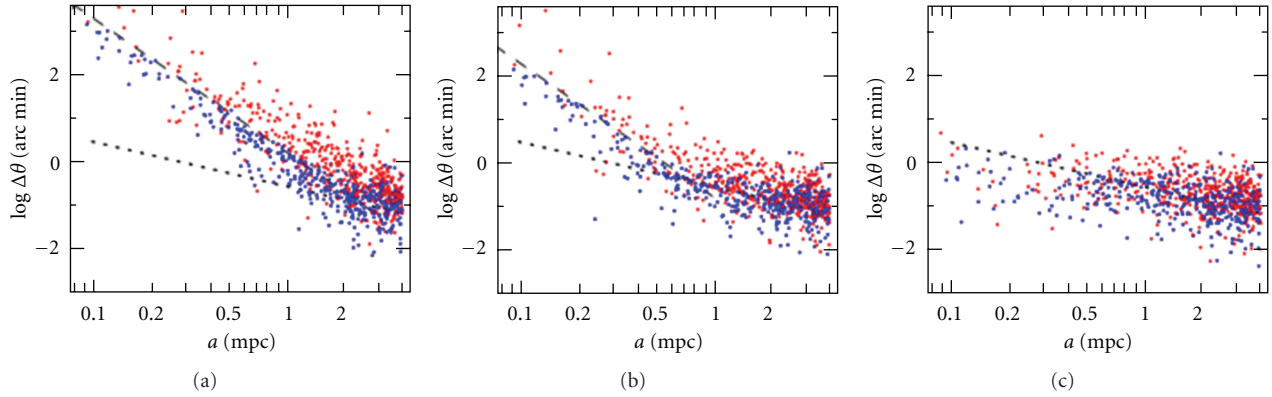


FIGURE 6: Evolution of the orbital angular momenta of stars around Sgr A* due to frame-dragging (dashed lines) and stellar perturbations (dotted lines) as measured by the angle $\Delta\theta$ as a function of the orbital semimajor axis a. The three panels correspond to a Kerr black hole with a spin (a) $a = \mathrm{M}$, (b) $a = 0.1\,\mathrm{M}$, and (c) $a = 0$ [30].

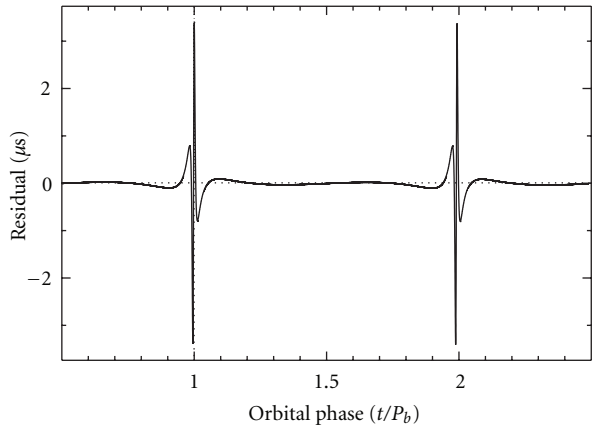


FIGURE 7: Typical timing residuals for a radio pulsar in an orbit around a black hole with a mass of $10^4\,\mathrm{M}_\odot$ due to the presence of a nonvanishing quadrupole moment [31]. For Sgr A*, such residuals will have a similar shape but with a larger amplitude.

"ring of light" depends directly on the mass, spin, inclination, and the deviation parameter of the black hole (see Figure 4).

The diameter of the ring of light as observed by a distant observer depends predominantly on the mass of the black hole and is nearly constant for all values of the spin and disk inclination as well as for small values of the deviation parameter. For nonzero values of the spin of the black hole, the ring is displaced off center in the image plane. In all cases, the ring of a Kerr black hole remains nearly circular except for very large values of the spin $a \gtrsim 0.9\,\mathrm{M}$. However, if Sgr A* is not a Kerr black hole, the ring becomes asymmetric in the image plane. This asymmetry is a direct measure for a violation of the no-hair theorem (see Figure 5).

4. Combining Strong-Field with Weak-Field Tests of the No-Hair Theorem

In addition to a strong-field test of the no-hair theorem with VLBI imaging of Sgr A*, there exist two other promising possibilities for performing such a test in the weak-field regime. The presence of a nonzero spin and quadrupole moment independently leads to a precession of the orbit of stars around Sgr A* at two different frequencies, which can be studied with parameterized post-Newtonian dynamics [29, 79]. Merritt et al. [30] showed that the effect of the quadrupole moment on the orbit of such stars is masked by the effect of the spin for the group of stars known to orbit Sgr A*. However, if a star can be detected within ~1000 Schwarzschild radii of Sgr A* and if it can be monitored over a sufficiently long period of time, this technique may also measure the spin (see Figure 6; [30]) and even the quadrupole moment [30, 80] together with the already obtained mass [37, 38]. Future instruments, such as GRAVITY [81], may be able to resolve the orbits of such stars providing an independent test of the no-hair theorem.

Yet another weak-field test can be performed by the observation of a radio pulsar on an orbit around Sgr A*. If present, timing observations may resolve characteristic spin-orbit residuals that are induced by the quadrupole moment and infer its magnitude (see Figure 7; [31]). Recent surveys set an upper limit for the existence of up to 90 pulsars within the central parsec of the galaxy [82] making this technique a promising third approach for testing the no-hair theorem with Sgr A*.

The fundamental properties of the black hole in the center of our galaxy can be probed with three different techniques. The combination of the results of all three approaches will lead to a secure test of the no-hair theorem with Sgr A*.

References

[1] W. Israel, "Event horizons in static vacuumspace-times," *Physical Review*, vol. 164, no. 5, pp. 1776–1779, 1967.

[2] W. Israel, "Event horizons in static electrovac space-times," *Communications in Mathematical Physics*, vol. 8, no. 3, pp. 245–260, 1968.

[3] B. Carter, "Axisymmetric black hole has only two degrees of freedom," *Physical Review Letters*, vol. 26, no. 6, pp. 331–333, 1971.

[4] S. W. Hawking, "Black holes in general relativity," *Communications in Mathematical Physics*, vol. 25, p. 152, 1972.

[5] B. Carter, *Black Holes*, Gordon and Breach, New York, NY, USA, 1973.

[6] D. C. Robinson, "Uniqueness of the Kerr black hole," *Physical Review Letters*, vol. 34, no. 14, pp. 905–906, 1975.

[7] R. Geroch, "Multipole moments. II. Curved space," *Journal of Mathematical Physics*, vol. 11, no. 8, pp. 2580–2588, 1970.

[8] R. O. Hansen, "Multipole moments of stationary space–times," *Journal of Mathematical Physics*, vol. 15, no. 1, pp. 46–53, 1974.

[9] D. Psaltis, *Compact Stellar X-Ray Sources*, Cambridge University Press, Cambridge, Mass, USA, 2006.

[10] F. D. Ryan, "Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments," *Physical Review D*, vol. 52, no. 10, pp. 5707–5718, 1995.

[11] F. D. Ryan, "Accuracy of estimating the multipole moments of a massive body from the gravitational waves of a binary inspiral," *Physical Review D*, vol. 56, no. 4, pp. 1845–1855, 1997.

[12] F. D. Ryan, "Scalar waves produced by a scalar charge orbiting a massive body with arbitrary multipole moments," *Physical Review D*, vol. 56, no. 12, pp. 7732–7739, 1997.

[13] L. Barack and C. Cutler, "LISA capture sources: approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy," *Physical Review D*, vol. 69, no. 8, Article ID 082005, 2004.

[14] L. Barack and C. Cutler, "Using LISA extreme-mass-ratio inspiral sources to test off-Kerr deviations in the geometry of massive black holes," *Physical Review D*, vol. 75, no. 4, Article ID 042003, 2007.

[15] J. Brink, "Spacetime encodings. I. A spacetime reconstruction problem," *Physical Review D*, vol. 78, Article ID 102001, 8 pages, 2008.

[16] C. Li and G. Lovelace, "Generalization of Ryan's theorem: probing tidal coupling with gravitational waves from nearly circular, nearly equatorial, extreme-mass-ratio inspirals," *Physical Review D*, vol. 77, Article ID 064022, 10 pages, 2008.

[17] T. A. Apostolatos, G. Lukes-Gerakopoulos, and G. Contopoulos, "How to observe a non-kerr spacetime using gravitational waves," *Physical Review Letters*, vol. 103, no. 11, Article ID 111101, 2009.

[18] N. A. Collins and S. A. Hughes, "Towards a formalism for mapping the spacetimes of massive compact objects: bumpy black holes and their orbits," *Physical Review D*, vol. 69, Article ID 124022, 2004.

[19] S. J. Vigeland and S. A. Hughes, "Spacetime and orbits of bumpy black holes," *Physical Review D*, vol. 81, no. 2, Article ID 024030, 2010.

[20] K. Glampedakis and S. Babak, "Mapping spacetimes with LISA: inspiral of a test body in a "quasi-Kerr" field," *Classical and Quantum Gravity*, vol. 23, no. 12, article 013, pp. 4167–4188, 2006.

[21] J. R. Gair, C. Li, and I. Mandel, "Observable properties of orbits in exact bumpy spacetimes," *Physical Review D*, vol. 77, no. 2, Article ID 024035, 23 pages, 2008.

[22] T. Johannsen and D. Psaltis, "Testing the no-hair theorem with observations in the electromagnetic spectrum. I. Properties of a Quasi-Kerr spacetime," *Astrophysical Journal*, vol. 716, no. 1, pp. 187–197, 2010.

[23] T. Johannsen and D. Psaltis, "Testing the no-hair theorem with observations in the electromagnetic spectrum. II. Black hole images," *Astrophysical Journal*, vol. 718, no. 1, p. 446, 2010.

[24] T. Johannsen and D. Psaltis, "Testing the no-hair theorem with observations in the electromagnetic spectrum. III. Quasi-periodic variability," *Astrophysical Journal*, vol. 726, no. 1, p. 11, 2011.

[25] T. Johannsen and D. Psaltis, "Testing the no-hair theorem with observations of black holes in the electromagnetic spectrum," *Advances in Space Research*, vol. 47, p. 528, 2011.

[26] D. Psaltis and T. Johannsen, "A ray-tracing algorithm for spinning compact object spacetimes with arbitrary quadrupole moments. I. Quasi-kerr black holes," *Astrophysical Journal*. In press.

[27] C. Bambi and E. Barausse, "Constraining the quadrupole moment of stellar-mass black hole candidates with the continuum fitting method," *Astrophysical Journal*, vol. 731, p. 121, 2011.

[28] C. Bambi, "Constraint on the quadrupole moment of supermassive black hole candidates from the estimate of the mean radiative efficiency of AGN," *Physical Review D*, vol. D 83, Article ID 103003, 4 pages, 2011.

[29] C. M. Will, " Testing the general relativistic "no-hair" theorems using the galactic center black hole Sgr A*," *Astrophysical Journal*, vol. 674, p. L25, 2008.

[30] D. Merritt, T. Alexander, S. Mikkola, and C. M. Will, "Testing properties of the Galactic center black hole using stellar orbits," *Physical Review D*, vol. 81, no. 6, Article ID 062002, 2010.

[31] N. Wex and S. M. Kopeikin, "Frame dragging and other precessional effects in black hole pulsar binaries," *Astrophysical Journal*, vol. 514, no. 1, pp. 388–401, 1999.

[32] V. S. Manko and I. D. Novikov, "Generalizations of the Kerr and Kerr-Newman metrics possessing an arbitrary set of mass-multipole moments," *Classical and Quantum Gravity*, vol. 9, no. 11, article 013, pp. 2477–2487, 1992.

[33] S. J. Vigeland, N. Yunes, and L. C. Stein, "Bumpy black holes in alternative theories of gravity," *Physical Review D*, vol. 83, Article ID 104027, 16 pages, 2011.

[34] T. Johannsen and D. Psaltis, "Metric for rapidly spinning black holes suitable for strong-field tests of the no-hair theorem," *Physical Review D*, vol. 83, Article ID 124015, 16 pages, 2011.

[35] S. A. Hughes, "(Sort of) testing relativity with extreme mass ratio inspirals," in *Proceedings of the AIP Conference*, vol. 873, pp. 233–240, November 2006.

[36] D. Psaltis and T. Johannsen, "Sgr A*: the optimal testbed of strong-field gravity," *Journal of Physics*, vol. 283, Article ID 102030, 2011.

[37] A. M. Ghez, S. Salim, N. N. Weinberg et al., "Measuring distance and properties of the milky way's central supermassive black hole with stellar orbits," *Astrophysical Journal*, vol. 689, no. 2, pp. 1044–1062, 2008.

[38] S. Gillessen, F. Eisenhauer, S. Trippe et al., "Monitoring stellar orbits around the massive black hole in the galactic center," *Astrophysical Journal*, vol. 692, p. 1075, 2009.

[39] S. S. Doeleman, J. Weintroub, A. E. E. Rogers et al., "Event-horizon-scale structure in the supermassive black hole candidate at the Galactic Centre," *Nature*, vol. 455, no. 7209, pp. 78–80, 2008.

[40] K. Schwarzschild, "Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* , vol. 1, pp. 189–196, 1916.

[41] R. P. Kerr, "Gravitational field of a spinning mass as an example of algebraically special metrics," *Physical Review Letters*, vol. 11, no. 5, pp. 237–238, 1963.

[42] J. B. Hartle, "Slowly Rotating Relativistic Stars. I. Equations of Structure," *Astrophysical Journal*, vol. 150, p. 1005, 1967.

[43] J. B. Hartle and K. S. Thorne, "Slowly Rotating Relativistic Stars. II. Models for Neutron Stars and Supermassive Stars," *Astrophysical Journal*, vol. 153, p. 807, 1968.

[44] A. Tomimatsu and H. Sato, "New exact solution for the gravitational field of a spinning mass," *Physical Review Letters*, vol. 29, no. 19, pp. 1344–1345, 1972.

[45] A. Tomimatsu and H. Sato, "New series of exact solutions for gravitational fields of spinning masses," *Progress of Theoretical Physics*, vol. 50, no. 1, pp. 95–110, 1973.

[46] H. Stephani, D. Kramer, M. A. H. MacCallum, C. Hoenselaers, and E. Herlt, *Exact Solutions of Einstein's Field Equations*, Cambridge University Press, Cambridge, Mass, USA, 2003.

[47] I. Hauser and F. J. Ernst, "Proof of a Geroch conjecture," *Journal of Mathematical Physics*, vol. 22, no. 5, pp. 1051–1063, 1981.

[48] C. Hoenselaers, W. Kinnersley, and B. C. Xanthopoulos, "Symmetries of the stationary Einstein–Maxwell equations. VI. Transformations which generate asymptotically flat space-times with arbitrary multipole moments," *Journal of Mathematical Physics*, vol. 20, no. 8, p. 2530, 1979.

[49] R. Geroch, "A method for generating solutions of Einstein's equations," *Journal of Mathematical Physics*, vol. 12, no. 6, pp. 918–924, 1971.

[50] R. Geroch, "A method for generating new solutions of Einstein's equation. II," *Journal of Mathematical Physics*, vol. 13, no. 3, pp. 394–404, 1972.

[51] R. Beig and W. Simon, "Proof of a multipole conjecture due to Geroch," *Communications in Mathematical Physics*, vol. 78, p. 75, 1980.

[52] R. Beig and W. Simon, "On the multipole expansion for stationary space-times," *Proceedings of the Royal Society A*, vol. 376, pp. 333–341, 1981.

[53] N. R. Sibgatullin, *Oscillations and Waves in Strong Gravitational and Electromagnetic Fields*, Springer, Berlin, Germany, 1991.

[54] V. S. Manko and N. R. Sibgatullin, "Construction of exact solutions of the Einstein-Maxwell equations corresponding to a given behaviour of the Ernst potentials on the symmetry axis," *Classical and Quantum Gravity*, vol. 10, no. 7, article 014, pp. 1383–1404, 1993.

[55] E. M. Butterworth and J. R. Ipser, "On the structure and stability of rapidly rotating fluid bodies in general relativity. I - The numerical method for computing structure and its application to uniformly rotating homogeneous bodies," *Astrophysical Journal*, vol. 204, pp. 200–223, 1976.

[56] N. Stergioulas and J. L. Friedman, "Comparing models of rapidly rotating relativistic stars constructed by two numerical methods," *Astrophysical Journal*, vol. 444, no. 1, pp. 306–311, 1995.

[57] W. G. Laarakkers and E. Poisson, "Quadrupole moments of rotating neutron stars," *Astrophysical Journal*, vol. 512, no. 1, pp. 282–287, 1999.

[58] E. Berti, F. White, A. Maniopoulou, and M. Bruni, "Rotating neutron stars: an invariant comparison of approximate and numerical space-time models," *Monthly Notices of the Royal Astronomical Society*, vol. 358, no. 3, pp. 923–938, 2005.

[59] C. Cadeau, S. M. Morsink, D. Leaky, and S. S. Campbell, "Light curves for rapidly rotating neutron stars," *Astrophysical Journal*, vol. 654, no. 1 I, pp. 458–469, 2007.

[60] C. F. Sopuerta and N. Yunes, "Extreme- and intermediate-mass ratio inspirals in dynamical Chern-Simons modified gravity," *Physical Review D*, vol. 80, no. 6, Article ID 064006, 2009.

[61] D.-C. Dai and D. Stojkovic, "Analytic solution for a static black hole in RSII model," submitted to *General Relativity and Quantum Cosmology*.

[62] N. Yunes and L. C. Stein, "Effective gravitational wave stress-energy tensor in alternative theories of gravity," *Physical Review D*, vol. 83, p. 4002, 2011.

[63] E. Barausse, T. Jacobson, and T. P. Sotiriou, "Black holes in Einstein-aether and Horava-Lifshitz gravity," *General Relativity and Quantum Cosmology*, vol. 83, Article ID 124043, 2011.

[64] P. Figueras and T. Wiseman, "Gravity and large black holes in Randall-Sundrum II braneworlds," submitted to *High Energy Physics*.

[65] T. Johannsen et al., in preparation.

[66] A. E. Broderick, V. L. Fish, S. S. Doeleman, and A. Loeb, "Estimating the parameters of sagittarius A*'s accretion flow via millimeter vlbi," *Astrophysical Journal*, vol. 697, no. 1, pp. 45–54, 2009.

[67] A. E. Broderick, V. L. Fish, S. S. Doeleman, and A. Loeb, "Evidence for low black hole spin and physically motivated accretion models from millimeter-VLBI observations of sagittarius A*," *Astrophysical Journal*, vol. 735, p. 110, 2011.

[68] J. M. Bardeen, *Black Holes*, Gordon and Breach, New York, NY, USA, 1973.

[69] H. Falcke, F. Melia, and E. Agol, "Viewing the shadow of the black hole at the Galactic center," *Astrophysical Journal*, vol. 528, no. 1, pp. L13–L16, 2000.

[70] A. E. Broderick and A. Loeb, "Imaging bright-spots in the accretion flow near the black hole horizon of Sgr A*," *Monthly Notices of the Royal Astronomical Society*, vol. 363, no. 2, pp. 353–362, 2005.

[71] A. E. Broderick and A. Loeb, "Imaging optically-thin hotspots near the black hole horizon of Sgr A* at radio and near-infrared wavelengths," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 3, pp. 905–916, 2006.

[72] V. L. Fish and S. S. Doeleman, *IAU Symp. 261, Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis*, Cambridge University Press, Cambridge, UK, 2009.

[73] A. E. Broderick and A. Loeb, "Imaging optically-thin hotspots near the black hole horizon of Sgr A* at radio and near-infrared wavelengths," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 3, pp. 905–916, 2006.

[74] A. E. Broderick and A. Loeb, "Imaging the black hole silhouette of M87: implications for jet formation and black hole spin," *Astrophysical Journal*, vol. 697, no. 2, pp. 1164–1179, 2009.

[75] R. Takahashi, "Shapes and positions of black hole shadows in accretion disks and spin parameters of black holes," *Astrophysical Journal*, vol. 611, no. 2, pp. 996–1004, 2004.

[76] J. D. Schnittman and J. H. Krolik, "X-ray polarization from accreting black holes: the thermal state," *Astrophysical Journal*, vol. 701, no. 2, pp. 1175–1187, 2009.

[77] J. D. Schnittman and J. H. Krolik, "X-ray polarization from accreting black holes: coronal emission," *Astrophysical Journal*, vol. 712, no. 2, pp. 908–924, 2010.

[78] K. Beckwith and C. Done, "Extreme gravitational lensing near rotating black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 359, no. 4, pp. 1217–1228, 2005.

[79] C. M. Will, *Theory and Experiment in Gravitational Physics*, Cambridge University Press, Cambridge, Mass, USA, 1993.

[80] L. Sadeghian and C. M. Will, "Testing the black hole no-hair theorem at the galactic center: perturbing effects of stars in

the surrounding cluster," submitted to *General Relativity and Quantum Cosmology*.

[81] H. Bartko, G. Perrin, W. Brandner et al., "GRAVITY: astrometry on the galactic center and beyond," *New Astronomy Reviews*, vol. 53, no. 11-12, pp. 301–306, 2009.

[82] J.-P. MacQuart, N. Kanekar, D. A. Frail, and S. M. Ransom, "A high-frequency search for pulsars within the central parsec of Sgr A*," *Astrophysical Journal*, vol. 715, no. 2, pp. 939–946, 2010.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 940568, 14 pages
doi:10.1155/2012/940568

Review Article

Massive Black Hole Binaries: Dynamical Evolution and Observational Signatures

M. Dotti,[1] A. Sesana,[2] and R. Decarli[3]

[1] *Dipartimento di Fisica G. Occhialini, Università di Milano Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*
[2] *Galaxy and Cosmology Department, Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institut, 14476 Golm, Germany*
[3] *Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany*

Correspondence should be addressed to R. Decarli, decarli@mpia.de

Received 18 August 2011; Accepted 13 October 2011

Academic Editor: Francesco Shankar

The study of the dynamical evolution of massive black hole pairs in mergers is crucial in the context of a hierarchical galaxy formation scenario. The timescales for the formation and the coalescence of black hole binaries are still poorly constrained, resulting in large uncertainties in the expected rate of massive black hole binaries detectable in the electromagnetic and gravitational wave spectra. Here, we review the current theoretical understanding of the black hole pairing in galaxy mergers, with a particular attention to recent developments and open issues. We conclude with a review of the expected observational signatures of massive binaries and of the candidates discussed in literature to date.

1. Introduction

Understanding the formation and evolution of massive black holes (MBHs) is one of the most exciting goals of contemporary astrophysics and cosmology. It is now well established that MBHs are ubiquitous in nearby spheroids (e.g., [1]), most of them lurking in a quiescent accretion state, while during the cosmic history million-to-billion solar mass MBHs powered quasars. These objects have become in the last years central building blocks for all the proposed scenarios of galaxy formation (e.g., [2]), playing a major role in shaping galaxies through feedback processes springel [3]. Massive black hole binaries (MBHBs), formed during the galaxy merging process [4], promise to be among the most luminous gravitational wave (GW) sources for future space-borne interferometers like the proposed new gravitational wave observatory (NGO) (https://lisa-light.aei.mpg.de/bin/view/) and ongoing pulsar timing array (PTA) campaigns [5]. Theoretical modeling of MBHB dynamics is essential in addressing a number of fundamental astrophysical questions (such as the merger-quasar connection or the MBH-host relations) and in identifying putative signatures that may serve as a guidance for present and future observational campaigns.

Early stages of MBH pairing have been observed, from the initial phases of galaxy mergers, where two distinct but gravitationally bound galaxies are observable at separations of: $\sim$100 kpc [6–11], down to unbound pairs of MBHs at separations of $\lesssim$1 kpc embedded in a single galaxy remnant [12, 13].

During this initial stage, the MBH pairing is driven by dynamical friction acting on the host galaxies. The two MBHs (here after M_1 and M_2 for the primary and secondary MBH, resp.) bind in a binary if they reach a relative separation

$$a_{\rm BHB} \sim \frac{{\rm G}M_{\rm BHB}}{2\sigma^2} \sim 0.2\, M_{\rm BHB,6}\, \sigma_{100}^{-2}\ {\rm pc}, \quad (1)$$

where σ is the velocity dispersion of the host galaxy, $M_{\rm BHB} = M_1 + M_2$ is the total mass of the binary, and σ_{100} and $M_{\rm BHB,6}$ are in units of 100 km s^{-1} and $10^6\,{\rm M}_\odot$, respectively. If, in galaxy mergers, the mass of the MBHB scales with σ following the MBH mass versus σ relation (see, e.g., [14, 15] and references therein), (1) implies $a_{\rm BHB} \sim 0.5\, M_{\rm BHB,6}^{1/2}$ pc.

The efficiency of the process depends on the mass ratio between the merging galaxies. While equal mass (major) mergers result in a fast formation of a MBHB, in very unequal mass (minor) mergers, the satellite, tidally disrupted

along the course of the encounter, leaves its *naked* MBH wandering in the outskirts of the most massive galaxy (see, e.g., [16]). Recent numerical studies have addressed the efficiency of the formation of a MBHB as a function of the mass ratio between the two galaxies, the gas fraction in the merging galaxies, and the redshift of the merger [17]. In particular, the presence of a significant gas component in the satellite helps the MBHB formation: during the first pericenters, the interaction between the two galaxies promotes the formation of bars, that convey a large fraction of the available gas in the center of the merging galaxies (already noticed in lower resolution merger simulations, see, e.g., [18, 19]). This new nuclear gas overdensity deepens the potential well of the secondary nucleus and prevents its tidal disruption. Callegari et al. [17] found a critical galaxy mass ratio for the formation of a MBHB of $\sim$1/10 for gas-rich galaxies and $\gtrsim$1/4 for gas poor galaxies. In zero-order approximation, assuming the MBH mass versus bulge mass relation [20, 21], we expect similar mass ratios $q = M_2/M_1$ ($\leq$1) in MBHBs. However, q can be increased for gas-rich galaxies. The strong gas inflows in the satellite (more perturbed) galaxy can result in a faster growth of the smaller MBH and in higher values of the expected MBH mass ratio ($q \gtrsim 1/3$ [22]). (We caution that these ranges in q implicitly assume similar morphology in the two merging galaxies, although minor mergers can involve very different morphological types. A study of MBH pairing and the formation of MBHBs in mixed mergers (e.g., disks merging with ellipticals) is not available to date.)

As a consequence, efficient dynamical friction promotes the formation of MBH binaries with similar mass ratios. The expected number of observable binaries and the rate of MBH coalescences, however, depend on the dynamical evolution after the binary formation. In order to coalesce through GW emission in less than an Hubble time, the two MBHs have to reach a separation

$$a_{\rm GW} \approx 2 \times 10^{-3} f(e)^{1/4} \frac{q^{1/4}}{(1+q)^{1/2}} \left(\frac{M_{\rm BHB}}{10^6 {\rm M}_\odot} \right)^{3/4} {\rm pc}, \quad (2)$$

where $f(e) = [1 + (73/24)e^2 + (37/96)e^4](1 - e^2)^{-7/2}$ is a function of the binary eccentricity e [23]. Circular equal mass binaries can coalesce shrinking by a factor

$$\frac{a_{\rm BHB}}{a_{\rm GW}} \sim 350 \left(\frac{M_{\rm BHB}}{10^6\ {\rm M}_\odot} \right)^{-1/4}, \quad (3)$$

while this factor decreases to 1 for 10^6 M$_\odot$ binary with $e \approx 0.999$. In order to understand the final fate of a MBHB and to constrain theoretically its observability, it is fundamental to study at the same time the orbital decay and the eccentricity evolution of the binary. The final fate of the MBHs, that is, if they will coalesce in a single object or not, strongly depends on the amount of matter (stars and gas) they can interact with after the binary formation. A definite answer is not present to date. In this paper, different scenarios will be discussed, depending on the nature of the galaxy mergers and on the properties of the galaxy nuclei.

This paper is organized as follow. In Section 2, we review the dynamical evolution of MBHB in gas poor environments, while the effect of gas is discussed in Section 3. In Section 4 we describe the MBHB candidates observed to date. Finally, our conclusions are drawn in Section 5.

2. Dynamical Evolution in Gas-Poor Environment

In systems where dynamical friction is efficient in dragging the two MBHs to the center of the merger remnant, the now bound MBHB is inevitably embedded in a gas- and star-rich environment. Such rich ambient provides a variety of physical mechanisms to efficiently extract the energy and angular momentum of the MBHB, promoting its final coalescence. In this section, we focus on dynamical processes involving interactions with stars. MBHBs in pure stellar environments were the first to be examined (the basics going back to [4]), for the obvious reason that stars can be considered as point particles, affected by gravitational forces only. The MBHB-star interactions are therefore adequately described by Newton's laws only, without all the complications involved in gas dynamics. Nonetheless, a single star-binary interaction is, by definition, a three-body problem, and the dynamics of the system is inevitably chaotic. Therefore, no simple analytical solutions are viable, and numerical studies (both involving three body scatterings and full N-body simulations) have been massively exploited to tackle the problem. The fate of the MBHB is determined by its semimajor axis and eccentricity evolution (see Section 1); in the following, we discuss them separately.

2.1. Shrinking of the Binary Semimajor Axis. The basic physical process driving the MBHB evolution in presence of stars is the slingshot mechanism. A star intersecting the MBHB orbit undergoes a complex three-body interaction being eventually ejected at infinity, carrying away energy and angular momentum from the binary. Extensive three-body scattering experiments [24–26] have shown that ejected stars carry away an energy per unit mass of the order of

$$\Delta E_{\rm BHB} \approx \frac{3}{2} \frac{G\mu}{a_{\rm BHB}}, \quad (4)$$

where $\mu = M_1 M_2 / M_{\rm BHB}$ is the reduced mass of the binary. Assuming a classical interaction rate given by $\Gamma = \rho/m \Sigma v$, where Σ is the binary cross section, $\rho/m = n$ is the number density of the ambient stars, and v is their typical velocity "at infinity" with respect to the binary (i.e., the velocity dispersion σ of the stellar system). Quinlan [25] showed that the evolution of the binary semimajor axis is simply given as

$$\frac{da_{\rm BHB}}{dt} = -\frac{a_{\rm BHB}^2 G\rho}{\sigma} H, \quad (5)$$

where H is a dimensionless *hardening rate*. If $a_{\rm BHB} < GM_2/(4\sigma^2)$, $H \approx 15$ independently on the mass, mass ratio and eccentricity of the system. In principle, given a stellar system with density ρ and velocity dispersion σ, (5) predicts efficient coalescence of the MBHB.

However, the above simple treatment ignores the concept of loss cone depletion. In an extended stellar system, only

a tiny fraction of the stellar phase space allows orbits intersecting the MBHB, commonly referred as "binary loss cone." As stars are ejected, the loss cone is depleted, and the binary evolution is governed by the rate at which new stars are fed into the loss cone [27]. In typical stellar systems, the mass in stars in the binary loss cone is of the order of few times μ [28], insufficient to reach $a_{\rm GW}$ in most of the cases [29]. This is the origin of the "last parsec problem" [30]. In a spherical stellar system, the loss cone refilling proceeds on a two-body relaxation timescale [31], which is usually much longer than the Hubble time. In the last decade, this fact has been confirmed in N-body simulations [27, 30, 32]. In such simulations, after loss cone depletion, further hardening was provided by two-body relaxation. This is a process that depends on the "granularity" of the systems, and the result is an N-dependent hardening rate, with the binary evolution slowing down as the number of particle in the simulation increases. Extrapolating these results to a realistic N representative of a galactic bulge, the binary evolution would have stalled.

In recent years, evidence has emerged that the "last parsec problem" might be an artificial product of the "spherical cow" approximation which is often exploited in astronomy. Basically, the spherical systems studied in the simulations represent a worse case (and unrealistic) scenario. MBHBs are infact produced in galaxy mergers, in which the resulting stellar bulge is rotating, triaxial, and likely to undergo bar-like instabilities. In a triaxial potential, an orbiting star does not conserve any of its angular momentum components [31]. As a result, there is a vast family of orbits (called centrophilic) that are allowed to get arbitrarily close to the binary [33–35], keeping the loss cone full during the MBHB hardening process. Recent N-body simulations have confirmed this scenario. Berczik et al. [36] studied the evolution of a MBHB in a rotating bulge. In this case, the stellar system experiences a bar instability resulting in a triaxial potential. The binary hardening rate was found to be N-independent; a proof that the hardening was not proceeding because of spurious two-body relaxation. More recently, the advent of GPU computing made possible to simulate "ab initio" the evolution of two interacting stellar bulges hosting MBHs; a first step toward a realistic galactic merger scenario. Several simulations were performed by Khan et al. [37], Preto et al. [38], and Gualandris and Merritt [39]. In all cases, the stellar remnant was triaxial and rotating, and the hardening rate, given by triaxiality driven loss cone replenishment, was found to be independent on N, implying coalescence timescales of $\approx 10^8$ yr. Remarkably, when normalized to the merging galaxy properties, the binary hardening rates found in these simulations follow (5) where $H \approx 20$ [39]. This is a consequence of the fact that, whatever is the geometry of the system, the average "quantum" of energy taken away from an interacting star is always the same and the evolution of the system is determined by the star-binary interaction rate only.

2.2. Eccentricity Evolution. As pointed out in the introduction, eccentricity plays an important role in driving the binary coalescence. However, addressing the eccentricity evolution of the system is more complicated because the $\Delta e_{\rm BHB}$ caused by each individual interaction depends on a combination of energy and angular momentum exchanges. The angular momentum distribution of the interacting stars is therefore crucial. The eccentricity evolution can be described as

$$\frac{de_{\rm BHB}}{d\ln(1/a_{\rm BHB})} = K. \tag{6}$$

Here K is a dimensionless parameter that, differently than H, depends on the binary mass ratio and eccentricity itself [25, 26]. In general K is a positive number in the range 0–0.3 (the peak value occurs at $e_{\rm BHB} \approx 0.6$), meaning that the binary eccentricity grows during the shrinking process. Sesana [40] constructed complete binary evolutionary tracks by coupling three-body scattering experiments of bound and unbound stars to an analytical description of the stellar distribution and on the loss cone refilling. As a general trend, quasi-circular-equal mass MBHBs experience just a mild eccentricity growth, while systems which are already eccentric at the moment of pairing, or with q significantly lower than 1, can evolve up to $e_{\rm BHB} > 0.9$.

The eccentricity evolution in stellar environments has been tackled by several authors by means of full N-body simulations. However, the limited number of particles ($N < 10^6$) in such simulations results in very noisy behavior for the binary eccentricity, and it is difficult to draw conclusions about the general trends behind the numerical noise. Milosavljevic and Merritt [30] carried out numerical integration of equal MBHBs embedded in two merging isothermal cusps ($\rho \propto r^{-\gamma}$ with $\gamma = 2$). Starting with circular orbits, they find a mild eccentricity increase to a value of $\lesssim$0.2 during the stellar-driven hardening phase. Merritt et al. [32] considered equal MBHBs embedded in Dehnen density profiles [41] with $\gamma = 1.2$ with different initial eccentricities. Again, they find that circular binaries tend to stay circular, while eccentric binaries tend to increase their eccentricities in reasonable agreement with the prediction of scattering experiments. Simulations starting before the formation of MBHBs carried by Hemsendorf et al. [42] and Aarseth [43] produce binaries with $e_0 \approx 0.8$ at the moment of pairing, with e subsequently increasing up to $\gtrsim$0.95. Amaro-Seoane et al. [44] focused on binaries of intermediate MBHs ($M \sim 10^3\,{\rm M}_\odot$) in massive star clusters. Coupling full N-body simulations to three-body scattering experiments, they find binaries with significant eccentricity (~0.5-0.6) at the moment of pairing, growing up to >0.9 during the hardening phase; similar conclusions are reported by Amaro-Seoane et al. [45]. On the small q side, simulations were performed by Baumgardt et al. [46] and Matsubayashi et al. [47], assuming a stellar density profile $\gamma = 1.75$, motivated by the analytical equilibrium solution for a dense relaxed stellar cusp around a massive object [48]. When properly rescaled, the eccentricity increase found in both papers agrees remarkably well with predictions based on the hybrid model by Sesana et al. [49]. Iwasawa et al. [50] investigated in detail the angular momentum exchanges between the binary and the stars responsible for the eccentricity growth, in bound stellar cusps. In particular,

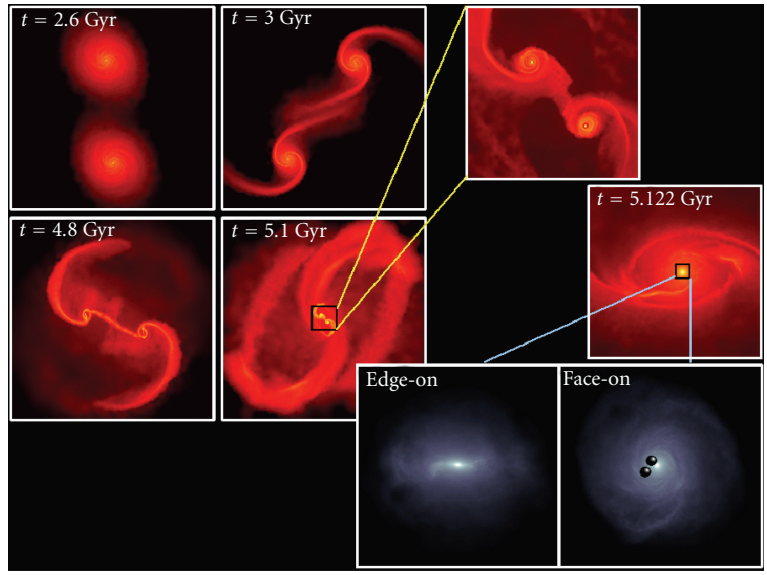


Figure 1: Different stages of an equal mass gas-rich merger (from [54]). The color code refers to gas density. The four panels to the left show the large-scale evolution at different times. The strong inflow of gas onto the two galactic centers is observable after the first pericenter. The boxes are 120 kps on a side (top) and 60 kpc on a side (bottom). The panels on the upper right are zoom-ins (8 kpc on a side) of the central region during the merger of the two galactic cores. The two bottom panels show the circumnuclear disk forming in the center of the galaxy remnant. The MBHs are shown as black spheres.

they showed that stars counterrotating with the binary tend to extract a lot of angular momentum from the MBHB, causing the eccentricity growth, whereas corotating stars do not. This is a simple consequence of angular momentum conservation during the ejection process, as shown by Sesana et al. [51].

The evolution of the binary eccentricity can be extremely different for nonisotropic systems. For example, Dotti et al. [52] showed that at large scales before the formation of a binary dynamical friction exerted by rotationally supported stellar disks tend to circularize the orbit of a MBH pair. At smaller separations, Sesana et al. [51] demonstrated, that in a rotating stellar system, the eccentricity evolution of unequal MBHB is dramatically affected by the level of co/counterrotation of the stellar distribution with respect to the binary, with corotating distributions promoting circularization rather than eccentricity growth. Nonetheless, most of the simulations involving rotating bulges [36, 53] or merging systems [37–39] find quite eccentric binaries at the moment of pairing (ranging from 0.4 to 0.8), and the subsequent evolution leads to a general eccentricity growth, in good agreement of what predicted for an isotropic stellar distribution. This may be because the binary evolution is mostly driven by loss cone refilling of unbound stars on almost radial orbits, with negligible initial angular momentum.

Overall, the emerging general picture favors efficient coalescence of MBHBs in dense stellar merger remnants. The triaxial and rotating nature of the stellar distribution promotes efficient loss cone refilling, while large eccentricities (especially in unequal mass systems) shorten the gap between the binary pairing and the efficient GW emission stage. In the near future, massive N-body simulations with several million particles will offer a unique opportunity to confirm this scenario.

3. Dynamical Evolution in Gas-Rich Environment

3.1. Formation of a MBH Binary in a Circumnuclear Gas Disks. As discussed in the Introduction, in comparable mass, gas-rich galaxy mergers, the gravitational interaction drives strong inflows of gas toward the galactic centers. The numerical multiscale investigation of an equal mass galaxy merger discussed in Mayer et al. [54] revealed that, in advanced stages of the galaxy merger, the two MBHs, orbiting in the central 100 pc of the merger remnant, are embedded in a dense, rotationally supported, gas disk (see Figure 1). This *circumnuclear* disk is self-gravitating and can be up to ∼500 times more massive than the MBH pair [54]. The dynamical evolution of the two MBHs is driven by dynamical friction, and, since the circumnuclear disk is the densest structure in the remnant nucleus, it is the main cause of their orbital decay. Mayer et al. [54] followed the evolution of the two MBHs from the initial stages of the galaxy merger down to $\lesssim$5 pc where they form a binary, as the mass of gas enclosed within their separation is less than the mass of the binary.

The evolution of MBHs in circumnuclear disks has been studied in details in dedicated simulations, in which the former evolution of the MBHs at distances $\gtrsim$100 pc is not explored. This allows to achieve a better resolution in the central region of the remnant and to study the latter

MBH pairing. Similar independent investigations discussed in Escala et al. [55] and Dotti et al. [52, 56] agree with Mayer et al. [54] on a rapid (on a timescale of $\lesssim 10^6$ yr) formation of a MBHB at parsec separations. The higher resolution in these studies allow for a further decay of the binary, that reaches sub-pc separations comparable to the spatial resolution in these numerical studies.

During this intermediate stage ($100\,\mathrm{pc} \gtrsim a \gtrsim 0.1\,\mathrm{pc}$), the interaction with the circumnuclear disk strongly affects the eccentricity of the MBH pair. Dotti et al. [52] first noted that, for MBHs with an initial significant eccentricity, the decay phase is preceded by a circularization of the orbit. This is due to the dynamical friction exerted by a rotating background onto the BHs at their apocenters. Since the circumnuclear disk is rotationally supported, the BHs at their apocenter move more slowly than the local gas and are dragged in the direction of their motion. This positive torque is exerted only in proximity of the apocenter, where the angular momentum of the MBHs can be maximally increased. This effect is a general feature of dynamical friction exerted by a rotating background. Dotti et al. [52] demonstrated that the same effect is present for MBHBs orbiting in a stellar circumnuclear disk (as discussed in Section 2.2), and Callegari et al. [17] found the same effect acting in unequal mass galaxy mergers at larger scales, where the disk of the primary galaxy circularizes the orbit of the satellite.

3.2. Evolution of Close MBH Binaries in Circumbinary Gas Disks. After the formation of a close binary, the MBHs, acting as a source of angular momentum, exerts a tidal torque that inhibits the gas from drifting inside its orbit. This creates a hollow density region, called gap that surrounds the binary (e.g., [57–61]).

As a consequence of disk clearance, corotation and inner Lindblad resonances are reduced in power, drastically changing the dynamical evolution of the binary. The same transfer of angular momentum that keeps the disk from infalling onto the MBHs drives the shrinking of the binary. Analytical and numerical studies agree in predicting that the interaction between the binary and the surrounding material reduces the semimajor axis of the binary and increases the eccentricity of an initially low eccentricity binary (e.g., [61–70]).

The MBHs/circumbinary disk interaction can be studied in two-limit cases: (i) assuming that the MBHB is embedded in a virtually infinite disk, as in the case in which it is continuously refilled by a long-lived larger-scale structure, or (ii) assuming that the disk has a finite mass (and, as a consequence, a finite reservoir of energy and angular momentum).

In the first limit of a constantly fuelled disk, analytical models of the evolution of the binary are available. The orbital decay timescale is

$$t_{\mathrm{decay}} = \frac{M_{\mathrm{disk}}(a) + M_2}{M_{\mathrm{disk}}(a)} t_{\mathrm{visc}}, \tag{7}$$

where $M_{\mathrm{disk}}(a)$ is an estimate of the disk mass inside the orbit of the secondary MBH. As long as $M_{\mathrm{disk}}(a)$ is greater than M_2, the MBH behaves as a fluid element, decaying on the viscous timescale t_{visc} onto the primary. As the binary orbit shrinks, $M_{\mathrm{disk}}(a)$ decreases. When the enclosed mass becomes comparable with the secondary its decay timescale increases to resulting in very large migration timescales at small separations. However, as first noticed in Ivanov et al. [65] in a continuously refilled disk, the stalling of the binary would cause a steady increase of the density of the inner edge of the circumbinary disk, until the mass close to the binary becomes again comparable to the secondary. At this stage, fast migration starts again. Following Ivanov et al. [65] and Cuadra et al. [70] estimated that, for a disk on the verge to undergo fragmentation (i.e., as dense as possible), the coalescence timescale of a $3 \times 10^6\,\mathrm{M}_\odot$ binary with $q = 1/3$, starting from an initial separation of $<0.05\,\mathrm{pc}$, is $<10^9$ yr. This timescale decreases with the decrease of the binary mass. Promisingly, the initial separation assumed here is close to the limit achieved in the larger scale simulations discussed in Section 3.1. Note, however, that such a timescale is comparable with the age of the Universe at $z \gtrsim 7$. If the migration in a dense circumbinary disk is the fastest process driving the MBHs coalescence, no coalescences of binary with $M_{\mathrm{BHB}} \gtrsim 10^6\,\mathrm{M}_\odot$ are expected at $z \gtrsim 7$.

In a similar way, the study of the evolution of e is possible. For simple alpha disks [71], Artymowicz and Lubow [59] found that the outer edge of the gap (i.e., the inner edge of the circumbinary disk) depends on the binary eccentricity, with more eccentric binaries opening larger gaps. However, the eccentricity itself increases as a consequence of the interaction with the expanding disk [62, 64, 67, 68, 70] resulting in a steady expansion of the gap. These two effect together would result in a limiting eccentricity of ~ 1, leading to a fast coalescence due to efficient gravitational wave emissions (see (2)). A smaller limiting eccentricity may result from a less efficient coupling between the gas and the binary, as expected if the disk moves farther out. Since the evolution of e in an expanding gap has been performed with finite mass disks, we postpone its discussion to the following.

If the circumbinary disk is limited in mass, the evolution of the orbital separation and eccentricity can be quite different. Since, in this scenario, the disk is not continuously refilled from the outside, the gas mass initially available in the disk is the key parameter.

(i) If the disk is $\gg M_2$, the evolution is similar to the infinite-mass disk, with the binary coalescing on a short time-scale.

(ii) If the mass in the circumbinary disk is less or of the order of M_2, the interaction with the binary forces the whole circumbinary disk to move outward in few orbital periods. This expansion of the circumnuclear disk has been observed in simulations in which the components of the binary have similar masses (e.g., [62, 70]). In this case, the gas reaches distances $\gtrsim 4a$, at which the interaction with the binary is not efficient anymore. A small amount of gas can then fall again closer to the binary because of orbital angular momentum exchange with the bulk of the disk, but most of the gas would never get close

enough to significantly alter the evolution of the binary. Note that the expansion of the circumbinary disk is most effective for eccentric binaries. Binaries do not coalesce because of the interaction with too small ($\lesssim M_2$), nonrefuelled disks.

In this simple description, the mass in gas within the disk is either accreted onto the MBHs or conserved. However, the gas in such a dense environment could be consumpted by star formation, decreasing the effective mass of the circumbinary disk. As a consequence, the disk could be initially $\gg M_2$, but decreasing in mass with time, and could possibly fail in bringing the binary to the final coalescence. This scenario has been recently discussed in Lodato et al. [72]. In this investigation, stars are allowed to form in the disk whenever it becomes gravitationally unstable. The rate of new star formation is obtained requiring that they would inject enough energy to keep the disk on the verge of fragmentation (i.e., providing an heating term exactly equal to the cooling losses in the disk). Even considering such a simple "thermal" feedback from the newly formed stars, the disk loses so much mass that the binary cannot reach the final coalescence, unless its initial separation is $\lesssim$0.01 pc (for the MBH masses considered in the paper, $M_1 = 10^8\,\mathrm{M}_\odot$ and $M_2 = 10^7\,\mathrm{M}_\odot$).

Other possible feedback terms, not included in this model, that can help in preventing such a strong gas consumption have been suggested by Lodato and collaborators, such as momentum feedback from stellar winds and supernovae explosions. Furthermore, the interaction between the MBHs and the forming gas clumps and stars, not considered in the investigation, could help the binary decay. The consumption of gas may not be a problem if large inflows of gas are present, as in the case of a continuously refuelled disk discussed above.

In the finite mass disk scenario, the existence of a limiting eccentricity has been studied in Roedig et al. [62] through a suite of high-resolution SPH simulations. They find a critical value $e_{\mathrm{crit}} \approx 0.6 - 0.8$. In these simulations, the initial ratio δ between the gap size and the semimajor axis of the binary is 2 and can increase during the runs up to more than 4, when the interaction efficiency drops. The analytical model presented in the paper agrees with the simulations, predicting the limiting eccentricity to be

$$e_{\mathrm{crit}} = 0.66\sqrt{\ln(\delta - 0.65)} + 0.19. \quad (8)$$

The initial choice of $\delta = 2$ is somewhat arbitrary. In reality, the feeding of a MBH binary forming in a gas-rich galaxy merger can be a very dynamic process, and the interaction with a single circumbinary disk could be too idealized a picture. Larger scale simulations show episodic gas inflows due to the dynamical evolution of the nucleus of the remnant (see, e.g., [73, 74]). In this scenario, the binary can still interact with a disk and excavate a gap, but the size of it would be time dependent (as in the simulations presented here) and would depend on the angular momentum distribution of the inflowing streams, resulting in a range of e_{crit}.

Note that the discussion above implicitly assumes that the MBHB and the circumnuclear disk corotate with each other. This is the natural outcome of a evolutionary sequence in a gas-rich galaxy mergers, in which the two MBHs orbit in a large scale circumnuclear disk, are forced to corotate with it [56] and open a gap in the very central region of the gas distribution. This picture, however, could not apply in gas-poorer mergers, or even in a gas-rich scenario, if the circumnuclear disk formed during the merger fails in bringing the two MBHs to the final coalescence before it is consumed by star formation and/or MBH feedback. In one of these cases, an occasional small inflow of gas could happen with a random angular momentum and could form a retrograde circumbinary disk, counterrotating with respect to the binary.

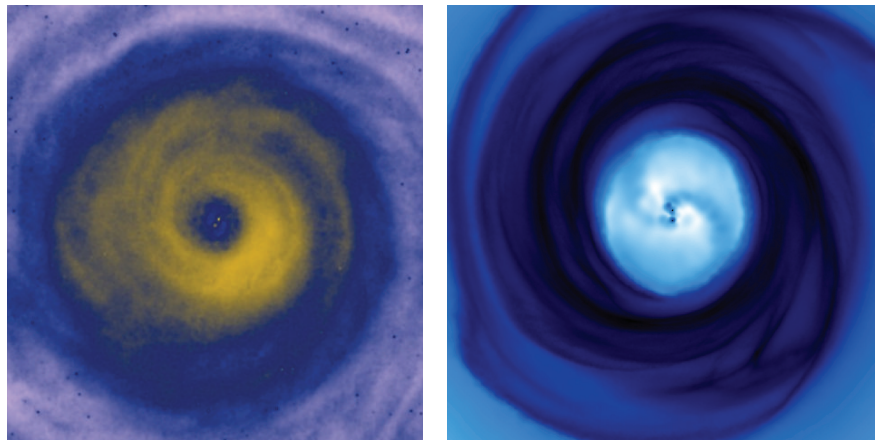
The evolution of a MBHB in a retrograde disk has been discussed in Nixon et al. [75]. In this case, the gravitational interaction between the binary and the gas brakes both the components, so that, unlike in the prograde scenario, here the torques responsible for the binary shrinking and its eccentricity evolution cause the edge of the disk to move inwards. The binary and the secondary MBH in particular, experience the presence of a closer distribution of gas, which would imply a faster evolution, moving at a higher relative velocity, that results in a less effective interaction. Nixon and collaborators show that (i) if the binary is initially not exactly circular ($e < e_{\mathrm{crit}} \sim H/R$, where H/R is the aspect ratio of the disk) and (ii) if it interacts with the disk mainly at the apocenter, then the secondary evolves onto an almost radial orbit after interacting with a gas mass comparable to its own. (Note that assuming the interaction to take place mainly at the apocenter is in agreement with the circumbinary disk to form after the MBHs bind in a close binary, as a consequence of a randomly oriented accretion event.)

The increase of the eccentricity to $e \sim 1$ in the retrograde case is due to direct accretion of linear momentum from counterrotating material. Since the secondary has null radial velocity at the apocenter, before and after the interaction, and far from the apocenter, the secondary is assumed to move on an unperturbed Keplerian orbit, the apocenter is constant. Interacting with counterrotating gas, the secondary decreases its angular momentum, reducing its semimajor axis (of up to a factor of 2) and, most importantly, increasing its eccentricity (up to 1). At very high eccentricities the emission of gravitational waves can bring the binary to the final coalescence in less than an Hubble time (see (2)).

Note that this scenario suffers of the same disk consumption problem as the prograde one. If the disk is consumed by star formation or evacuated by MBH or supernovae feedbacks, the process stops. This makes this process particularly interesting for very unequal mass binaries ($M_2/M_1 \lesssim H/R$, less likely to form from galaxy mergers, as discussed in the Introduction). Fast inflows of gas, on timescales shorter than the consumption time, would help the coalescence in both the prograde and the retrograde scenario.

4. Binary Candidates in the Realm of Observations

Despite being a natural outcome of galaxy mergers, MBH pairs are still elusive. Less than 20 systems with separations



(a) (b)

Figure 2: Two consecutive stages of the MBHB decay in the nucleus of a gas-rich galaxy merger remnant. (a) Central region (∼5 pc) of a large-scale (100 pc) massive ($10^8\,M_\odot$) circumnuclear disk, embedding a sub-pc MBHB (from one of the simulations in [56]). The color code refers to the gas density (high densities in yellow). (b) an eccentric MBHB embedded in a sub-pc circumbinary disk (from [62]). Here, darker blue refers to higher densities.

of ∼10 pc to ∼10 kpc pairs are of this kind are known to date. MBHs orbit in the common postmerger stellar environment, in-spiralling because of dynamical friction. They appear as a single galaxy (eventually, with disturbed morphology) with two active nuclei. Examples are the prototypical case of NGC 6240 [12]; the radio galaxy 3C75 [76]; the spiral galaxy NGC 3341 [77]; the ULIRG Mrk463 [78]; the interacting galaxy COSMOS J100043.15+020637.2 [79]; the quasar pair J1254+0846 [80]. All these objects have been discovered because of the presence of two resolved X-ray sources wandering in the merged galaxy. In order to look for these systems, an alternative approach is to search for objects with two systems of narrow lines at slightly different redshifts [81, 82]. Large spectroscopic surveys, like the Sloan Digital Sky Survey (SDSS [83]), have been used to search for these systems. Follow-up observations were then used to discriminate between dual AGN and single AGN with complex gas dynamics in the narrow line region [81, 84].

At separations of $\lesssim$10 pc, the two MBHs start experiencing their own gravitational interaction, binding in a binary. These systems cannot be spatially resolved in optical and X-ray observations, and radio interferometry is required. This has been successfully done only in the case of 0402+379 [85, 86]. The two flat-spectrum radio sources, corresponding to the two components of the candidate MBHB, have a projected separation of ∼7 pc (few milliarcsec at $z = 0.055$). We note, however, that radio interferometry at very high spatial resolution is not an efficient technique to search for rare objects as MBHBs, because of the limited field of views and the requirement that both the MBHs are radio-luminous (see, f.i., [87]).

Another approach to look for MBHBs is to study periodic variations in the luminosity of some AGN. The only MBHB candidate selected on these bases up to now is the BL Lac object OJ287 (see [88], and references therein). It shows a 11 yr periodic flaring. In the MBHB scenario, the secondary MBH orbits on a tilted plane with respect to the accretion disk of the primary. The flares are associated to the passages of the secondary MBH through the nodes. However, this is not the only available explanation of the peculiar behavior of this source [89].

More promisingly, signatures of MBHBs have been searched for in optical and NIR spectroscopic databases. According to the MBHB hypothesis, if at least one of the MBHs is active, the broad lines (BLs) emitted by gas bound to it may be red- or blue-shifted with respect to their host galaxy systemic recessional velocity, as a consequence of the Keplerian motion of the binary [4]. Therefore, looking for quasars with significant velocity shifts (>few hundreds km s^{-1}) can be a valid approach to systematically search for MBHB candidates over large fields. This technique does not suffer of any angular resolution limitations. Actually, the closer (and more massive) the binary is, the more shifted/deformed the BLs are. Five MBHB candidates have been individually found in this way: SDSS J0927+2943 [90–92], J1536+0441 [93], J1050+3456 [94], J1000+2233 [95], and J0932+0318 [96]. Tsalmantza et al. [97] recently applied this technique in a systematic way over the whole spectroscopic sample of SDSS, using a method developed for searching unresolved gravitational lenses [98]. This analysis resulted in 4 new MBHB candidates (J1012+2613, J1154+0134, J1539+3333, and J1714+3327) and the confirmation of all the previously known objects (see Figure 3). In a similar study, Eracleous et al. [99] searched for objects with anomalous Hβ profiles in the SDSS quasar catalogue [100]. Among them, they identified 88 sources showing velocity shifts between the broad Hβ line peak and the rest frame of the narrow emission lines.

It should be noted that alternative interpretations for the spectral properties of the known candidates are available.

(i) Modest line shifts ($\lesssim$500 km s^{-1}) are often observed in "normal" AGN [101].

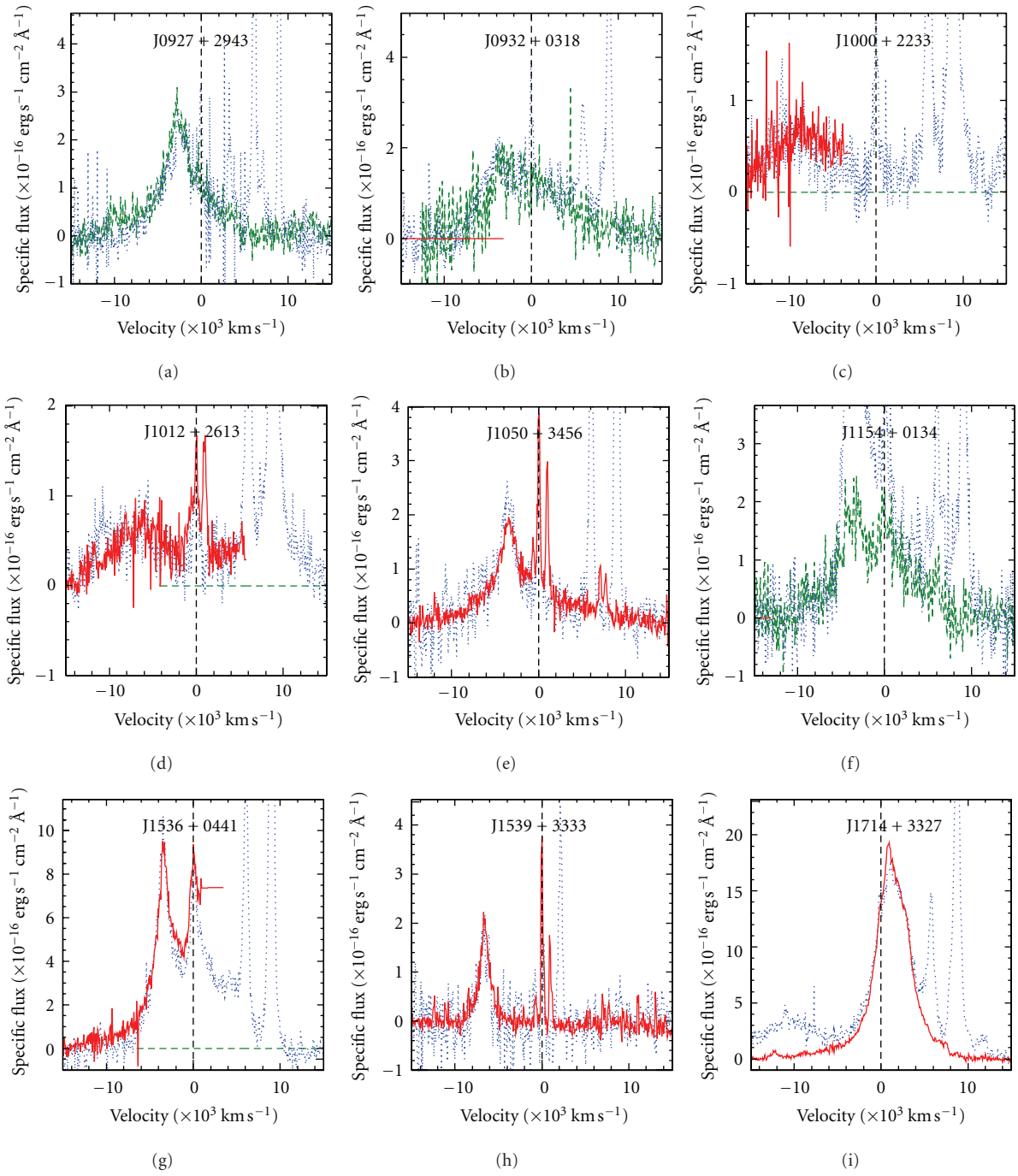


FIGURE 3: Velocity diagrams of Hα (red, solid lines), Hβ (blue, dotted lines), and MgII (green, dashed lines) for all the spectroscopically identified MBHB candidates. The flux of Hβ is scaled up to match the one of Hα or MgII. Adapted from Tsalmantza et al. [97].

(ii) Similarly, small velocity shifts (<4000 km s^{-1}) can be associated to the remnant of a binary coalescence, recoiling because of anisotropic gravitational wave [90]. (However, if the galaxy merger is gas rich, the maximum recoil velocity is expected to be <100 km s^{-1} [102–105].)

(iii) An unobscured MBHB with both MBHs active could resemble the spectrum of a double peaked emitter (see, e.g., [106]), where broad double-peaked lines are emitted because of the almost edge-on, disk-like structure of the broad line region of a single MBH.

(iv) Chance superposition of two AGNs (or an AGN-galaxy superposition) within the angular resolution of the used spectrograph can also mimic velocity shifts of different line systems [107].

The simplest way to discriminate between these scenarios and the MBHB hypothesis would be to look for a periodic

oscillation of the broad line shifts around the host galaxy redshift. However, the orbital period of the binary could be too long to be easily observed [4]. Noticeably, Eracleous et al. [99] observed a variation in the shifts at two different epochs in 14 out of 88 candidates with resulting accelerations between −120 and +120 km/s/yr. Longer temporal baselines are needed to prove the MBH binary interpretation for these objects.

In order to increase the number of known MBHB candidates, and to confirm their interpretation, it is therefore of fundamental importance to identify new signatures of MBHBs. The simultaneous observation of various MBHB signatures could represent the only way to firmly validate the MBHB scenario in the known candidates. A possibility is to look for peculiar flux ratios between high- and low-ionization broad lines. The broad line region of each MBH in a binary can be perturbed and disrupted by the gravitational potential of the companion. External shells of the broad line region (where most of the low-ionization line flux is emitted) are affected first, resulting in peculiar flux ratios [108]. This criterion is particularly interesting for quasars at high redshift ($z \sim 2$), where high- and low-ionization lines are observable in large surveys as the SDSS, while the most prominent narrow lines, needed to measure a shift related to the orbital motions, are not present in the SDSS spectra anymore.

At even closer separations between the two MBHs, when the size of the BLR is significantly larger than the semimajor axis of the binary ($a \lesssim 0.01$ pc), the optical and UV spectral features discussed above become more complex and not directly related with the period of the binary [109]. However, in this case, typical MBHB periods are $\lesssim$10 yr, opening the interesting opportunity of directly detect periodic variability of the system, related to the periodicity of the accretion flows [110–112]. Such close separations are particularly interesting, since they will be proven by the ongoing and future pulsar timing arrays (PTAs,[5]). In this context, Sesana et al. [112] estimate that up to few hundred MBHBs contributing to the GW signal in the PTA band may be identified through their periodicity in future X-ray all-sky surveys. Among those, few exceptionally bright sources may be resolved both in the GW and in the electromagnetic window through the detection of peculiar double K-α fluorescence lines, offering a unique multimessenger astronomy opportunity. An alternative possibility, suggested by Tanaka et al. [113], is that the presence of a circumbinary cavity results in a suppression of the UV soft-X emission. MBHBs close to coalescence may therefore be identified as AGNs with exceptionally faint UV X-ray continuum.

5. Conclusions

We reviewed the most recent findings about the dynamical evolution of MBHBs and their detectability. Regarding the binary dynamics, in the last few years we recognized the importance of the medium/large scale galactic structures ($\gtrsim$100 pc) in the dynamical evolution of the binaries. In gas free environments, the shape of the bulge potential is directly related to the possibility of a MBHB to reach the final coalescence. In a spherical system, the stars that can interact with the binary (i.e., that orbit within its loss-cone) are evacuated from the center before the binary can coalesce (e.g., [30]). A fast refilling of the loss cone, that can result in the merger of the two MBHs, is possible in triaxial systems, in which the angular momentum components of the orbiting stars are not conserved (e.g., [33]). Recently, thanks to the advent of GPU computing, large-scale galaxy mergers proved the occurrence of such a replenishment in more realistic scenarios [37–39]. The evolution of the binary eccentricity e depends on the dynamical properties of the core as well. If the MBHs are embedded in a nonrotating stellar system, the general trend is towards an increase of e with time (e.g., [25]). This promotes the coalescence of the MBHs, since gravitational wave emission is more efficient for eccentric binaries. In rotating systems, on the other hand, the evolution of e depends on the orientation of the binary: a binary corotating with the stellar cusp tends to decrease its eccentricity, while, in the counterrotating case, e grows up to $\gtrsim$0.95 [51]. This clear cut scenario could be modified by the interaction with stars on quasiradial, centrophilic orbits. A study of the orbital properties of these stars has not been presented to date.

Galaxy mergers can easily promote strong inflows of gas towards the center of the galaxy remnant (e.g., [54]). Hence, it is fundamental to understand how the presence of massive gas structures in the cores modify the dynamical evolution of the forming MBHBs. It has been proven that the interaction with circumnuclear (∼100 pc) disks can result in a fast ($\lesssim 10^7$ yr) formation of a MBHB (e.g., [55]). After this fast transient, the binary is thought to open a gap in the central gas distribution (e.g., [57]), and any further evolution is mediated by the interaction between the MBHs and the inner edge of the circumbinary disk.

Simulations and analytical studies about the interaction of MBHBs and circumbinary disks have improved our knowledge of the physical processes in play and their effect onto the binary evolution. However, a complete understanding of the binary/disk interaction is still to come. For example, in many investigations, the circumbinary disk is assumed to be cylindrically symmetric, that is, the study is reduced to an effective one-dimensional problem. The assumption of cylindrical symmetry removes any possibility of studying the effects of structures in the disk and, most importantly, of gas streams periodically inflowing from the disk onto the two MBHs, routinely observed in simulations (e.g., [62, 70, 110–112]). The torques exerted by these inflowing streams have not been studied in detail yet (with the notable exception of [114]) and could provide additional help (or resistance) in bringing the binary to the final coalescence.

The final fate of a binary embedded in such a circumbinary disk is still debated. If the disk is continuously refueled from any larger-scale gas distribution, a fast coalescence can easily be achieved. However, if the binary cannot interact with enough gas (e.g., because it turns in stars), the circumbinary disk gets evacuated and fails in bringing the MBHs to coalescence (e.g., [72]). As a consequence, as in the stellar scenario, the final fate of the binary depends

on the properties of larger scale structures and its ability to efficiently refuel the proximity of the MBHB with fresh gas. In principle, in presence of an intense inflow toward the center, the binary could fail in opening a gap at all, and the interaction between MBHs and a closer/denser gas structure could result in a faster coalescence of the MBHB [55]. Furthermore, if the angular momentum of gas can be efficiently reshuffled, inflowing streams could form counterrotating circumbinary disks, that can also promote the binary coalescence [75]. Only recently, the formation of a gap has been observed in large scale simulations [55, 115], in which the evolution of an extended ($\sim$100 pc) massive disk is followed (as massive as the binary in Escala et al., up to $\gtrsim 10\, M_{\rm BHB}$ in Dotti et al., see Figure 2). The spatial and mass resolution of these simulations do not allow yet a detailed study of the sub-pc evolution of the binary, down to a possible coalescence. Simulating the evolution of a sub-pc binary starting from large scale initial conditions, that can constrain the properties of the nuclear inflows together with the evolution of the binary, is the fundamental improvement needed to build a coherent picture of MBHBs in gas-rich environment.

To summarize the recent findings present in literature, we can draw a comparison between orbital decay timescales obtained considering different scenarios, for equal mass binaries.

(i) *In Dense Stellar Environments:* if the loss cone of the MBHB is constantly refilled (see Section 2.1), a binary of $10^6\,M_\odot$ with an initial separation of $a_0 \approx 1$ pc will coalesce in $\approx 3 \times 10^7$–10^8 years, while a binary 100 times more massive will inspire for about $\approx 10^8$–10^9 years before reaching the final coalescence [38, 40].

(ii) *In Gas Rich Environments:* if the interaction with a steady, long-lived, *corotating*, maximally massive circumbinary disk is responsible for the MBHB orbital decay (see Section 3.2), the orbital decay timescale for a $10^6\,M_\odot$ equal mass binary is less than the age of the Universe if the binary starts at an initial separation $a_0 \approx 0.03$ pc ($\sim$10 times smaller than the separation at which the MBHs bind in a binary $a_{\rm BHB}$) and is 2 orders of magnitude shorter for $a_0 \approx 0.01$ pc [70]. For a $10^8\,M_\odot$ equal mass binary, this timescale is less than the age of the Universe if $a_0 \approx 0.05$ pc [70], $\sim$100 times less than $a_{\rm BHB}$.

(iii) *In Gas Rich Environments:* if the MBHB interacts with a continuous sequence of counterrotating accretion disks with an accretion rate corresponding to the MBH Eddington limit, the orbital decay timescale is of the order of 10^8 yr, regardless the binary mass, for $a_0 < a_{\rm BHB}$. In this case, the coalescence timescale increases linearly with the inverse of the accretion rate.

Note that the estimates in gas-rich environment should be considered as lower limits, since they assume continuous accretion and, as stressed above, a continuous refuelling of the disks from larger scales. A single, not refuelled disk with enough mass and angular momentum to bring the MBHs to the final coalescence form $a_0 \lesssim 0.1$ pc would undergo fragmentation and star formation, as discussed in [72].

The presence of gas close to the binary is necessary to its detection. If at least one of the two MBHs is active, the orbital velocity of the binary can result in a frequency shift between the broad emission lines and the narrow emission lines (e.g., [4]). This shift is expected to change periodically, on the orbital period. Using this technique, few tens of MBHB candidates have been selected (e.g., [97, 99]). For all the candidates discussed to date, possible explanations other than a MBHB have been proposed (see Section 4). Moreover, the orbital period expected for such binaries is often too long to be observed, thus the periodic variation of the velocity shift cannot be used (yet) to determine the real nature of the candidates. It is therefore fundamental to couple this with other (independent) signatures of MBHBs, to confirm their nature.

In the near future, space-based interferometers like NGO will detect the GWs emitted in the late inspiral and final coalescence of MBHBs, opening the fascinating prospective of multimessenger astronomy. The identification of an electromagnetic counterpart to a gravitational wave detection of merging MBHs can teach us about relativity, accretion physics, galaxy formation and evolution, and, more in general, cosmology. A large variety of potential EM signatures have been proposed (see, e.g., [116] and references therein). In particular, thanks to the recent quick progresses in numerical relativity, simulations of the coalescence process in presence of matter is now becoming possible [117–120] and in the coming years may provide valuable insights about the nature of the associated electromagnetic signals. However, most of the counterparts discussed to date depend on some assumptions on the distribution of gas and/or stars in the immediate vicinities of the MBHB. A theoretical comprehension of the dynamics of matter close to the binary, necessary to understand the fate of a MBHB, is also fundamental to constrain the observability of an electromagnetic counterpart.

Tracing the formation, evolution, and final fate of MBHBs is certainly one of the open challenges of contemporary astrophysics. A better understanding of the interaction between MBHs and gas and the prediction of new peculiar observational features of MBHBs are needed to unambiguously constrain their properties and demography, adding an important missing piece to the galaxy evolution puzzle.

Acknowledgment

The authors are grateful to Jorge Cuadra, Monica Colpi, and Paraskevi Tsalmantza for useful discussions and comments.

References

[1] L. Ferrarese and H. Ford, "Supermassive black holes in galactic nuclei: past, present and future research," *Space Science Reviews*, vol. 116, no. 3-4, pp. 523–624, 2005.

[2] G. Kauffmann and M. G. Haehnelt, "A unified model for the evolution of galaxies and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 311, no. 3, pp. 576–588, 2000.

[3] V. Springel, T. Di Matteo, and L. Hernquist, "Modelling feedback from stars and black holes in galaxy mergers," *Monthly Notices of the Royal Astronomical Society*, vol. 361, no. 3, pp. 776–794, 2005.

[4] M. C. Begelman, R. D. Blandford, and M. J. Rees, "Massive black hole binaries in active galactic nuclei," *Nature*, vol. 287, no. 5780, pp. 307–309, 1980.

[5] G. Hobbs et al., "Gravitational waves and pulsar timing: stochastic background, individual sources and parameter estimation," *Classical and Quantum Gravity*, vol. 27, no. 8, article 084016, 2010.

[6] J. F. Hennawi, M. A. Strauss, M. Oguri et al., "Binary quasars in the sloan digital sky survey: evidence for excess clustering on small scales," *Astronomical Journal*, vol. 131, no. 1, pp. 1–23, 2006.

[7] A. D. Myers, R. J. Brunner, G. T. Richards, R. C. Nichol, D. P. Schneider, and N. A. Bahcall, "Clustering analyses of 300,000 photometrically classified quasars. II. The excess on very small scales," *Astrophysical Journal*, vol. 658, no. 1 I, pp. 99–106, 2007.

[8] A. D. Myers, G. T. Richards, R. J. Brunner et al., "Quasar clustering at 25 h-1 kpc from a complete sample of binaries," *Astrophysical Journal*, vol. 678, no. 2, pp. 635–646, 2008.

[9] G. Foreman, M. Volonteri, and M. Dotti, "Double quasars: probes of black hole scaling relationships and merger scenarios," *The Astrophysical Journal*, vol. 693, no. 2, article 1554, 2009.

[10] Y. Shen, J. F. Hennawi, F. Shankar et al., "Binary quasars at high redshift. II. Sub-Mpc clustering at z ~ 3-4," *Astrophysical Journal Letters*, vol. 719, no. 2, pp. 1693–1698, 2010.

[11] E. P. Farina, R. Falomo, and A. Treves, "A study of six low-redshift quasar pairs," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 4, pp. 3163–3167, 2011.

[12] S. Komossa, V. Burwitz, G. Hasinger, P. Predehl, J. S. Kaastra, and Y. Ikebe, "Discovery of a binary active galactic nucleus in the ultraluminous infrared galaxy NGC 6240 using Chandra," *Astrophysical Journal*, vol. 582, no. 1, pp. L15–L19, 2003.

[13] M. Colpi and M. Dotti, "Massive binary black holes in the cosmic landscape," *Advanced Science Letters*, vol. 4, no. 2, pp. 181–204, 2011.

[14] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal*, vol. 698, no. 1, pp. 198–221, 2009.

[15] A. W. Graham, C. A. Onken, E. Athanassoula, and F. Combes, "An expanded M_{bh}-σ diagram, and a new calibration of active galactic nuclei masses," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2211–2228, 2011.

[16] F. Governato, M. Colpi, and L. Maraschi, "The fate of central black holes in merging galaxies," *Monthly Notices of the Royal Astronomical Society*, pp. 271–317, 1994.

[17] S. Callegari, L. Mayer, S. Kazantzidis et al., "Pairing of supermassive black holes in unequal-mass galaxy mergers," *Astrophysical Journal*, vol. 696, no. 1, pp. L89–L92, 2009.

[18] J. E. Barnes and L. Hernquist, "Transformations of galaxies. II. Gasdynamics in merging disk galaxies," *Astrophysical Journal*, vol. 471, no. 1, pp. 115–142, 1996.

[19] J. E. Barnes, "Formation of gas discs in merging galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 333, no. 3, pp. 481–494, 2002.

[20] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *Astrophysical Journal Letters*, vol. 604, no. 2 II, pp. L89–L92, 2004.

[21] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *Astrophysical Journal*, vol. 589, no. 1, pp. L21–L24, 2003.

[22] S. Callegari, S. Kazantzidis, L. Mayer et al., "Growing massive black hole pairs in minor mergers of disk galaxies," *Astrophysical Journal Letters*, vol. 729, no. 2, article 85, 2011.

[23] P. C. Peters, "Gravitational radiation and the motion of two point masses," *Physical Review B*, vol. 136, no. 4, pp. B1224–B1232, 1964.

[24] S. Mikkola and M. J. Valtonen, "Evolution of binaries in the field of light particles and the problem of two black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 259, no. 1, pp. 115–120, 1992.

[25] G. D. Quinlan, "The dynamical evolution of massive black hole binaries: I. Hardening in a fixed stellar background," *New Astronomy*, vol. 1, no. 1, pp. 35–56, 1996.

[26] A. Sesana, F. Haardt, and P. Madau, "Interaction of massive black hole binaries with their stellar environment. I. Ejection of hypervelocity stars," *Astrophysical Journal*, vol. 651, no. 1 I, pp. 392–400, 2006.

[27] J. Makino and Y. Funato, "Evolution of massive black hole binaries," *Astrophysical Journal*, vol. 602, no. 1 I, pp. 93–102, 2004.

[28] D. Merritt and M. Milosavljevic, "Massive black hole binary evolution," *Living Reviews in Relativity*, vol. 8, no. 8, pp. 1–65, 2005.

[29] A. Sesana, F. Haardt, and P. Madau, "Interaction of massive black hole binaries with their stellar environment. II. Loss cone depletion and binary orbital decay," *Astrophysical Journal*, vol. 660, no. 1 I, pp. 546–555, 2007.

[30] M. Milosavljevic and D. Merritt, "Formation of galactic nuclei," *Astrophysical Journal*, vol. 563, no. 1, pp. 34–62, 2001.

[31] J. Binney and S. Tremaine, *Galactic Dynamics*, Princeton University Press, Princeton, NJ, USA, 1988.

[32] D. Merritt, S. Mikkola, and A. Szell, "Long-term evolution of massive black hole binaries. III. binary evolution in collisional nuclei," *Astrophysical Journal*, vol. 671, no. 1, pp. 53–72, 2007.

[33] M. Y. Poon and D. Merritt, "A self-consistent study of triaxial black hole nuclei," *Astrophysical Journal*, vol. 606, no. 2 I, pp. 774–787, 2004.

[34] D. Merritt and M. Y. Poon, "Chaotic loss cones and black hole fueling," *Astrophysical Journal*, vol. 606, no. 2 I, pp. 788–798, 2004.

[35] D. Merritt and E. Vasiliev, "Orbits around black holes in triaxial nuclei," *Astrophysical Journal Letters*, vol. 726, no. 2, article 61, 2011.

[36] P. Berczik, D. Merritt, R. Spurzem, and H. P. Bischof, "Efficient merger of binary supermassive black holes in non-axisymmetric galaxies," *Astrophysical Journal*, vol. 642, no. 1, pp. L21–L24, 2006.

[37] F. M. Khan, A. Just, and D. Merritt, "Efficient merger of binary supermassive black holes in merging galaxies," *Astrophysical Journal Letters*, vol. 732, no. 2, article 89, 2011.

[38] M. Preto, I. Berentzen, P. Berczik, and R. Spurzem, "Fast coalescence of massive black hole binaries from mergers of galactic nuclei: implications for low-frequency gravitational-wave astrophysics," *Astrophysical Journal Letters*, vol. 732, no. 2, article L26, 2011.

[39] A. Gualandris and D. Merritt, "Long-term evolution of massive black hole binaries. IV. Mergers of galaxies with

collisionally relaxed nuclei," *The Astrophysical Journal*. In press.

[40] A. Sesana, "Self consistent model for the evolution of eccentric massive black hole binaries in stellar environments: implications for gravitational wave observations," *Astrophysical Journal*, vol. 719, no. 1, pp. 851–864, 2010.

[41] W. Dehnen, "A family of patential-density pairs for spherical galaxies and bulges," *Monthly Notices of the Royal Astronomical Society*, vol. 265, no. 1, p. 250, 1993.

[42] M. Hemsendorf, S. Sigurdsson, and R. Spurzem, "Collisional dynamics around binary black holes in galactic centers," *Astrophysical Journal*, vol. 581, no. 2 I, pp. 1256–1270, 2002.

[43] S. J. Aarseth, "Black hole binary dynamics," *Astrophysics and Space Science*, vol. 285, no. 2, pp. 367–372, 2003.

[44] P. Amaro-Seoane, M. C. Miller, and M. Freitag, "Gravitational waves from eccentric intermediate-mass black hole binaries," *Astrophysical Journal*, vol. 692, no. 1, pp. L50–L53, 2009.

[45] P. Amaro-Seoane, C. Eichhorn, E. K. Porter, and R. Spurzem, "Binaries of massive black holes in rotating clusters: dynamics, gravitational waves, detection and the role of eccentricity," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 4, pp. 2268–2284, 2010.

[46] H. Baumgardt, A. Gualandris, and S. Portegies Zwart, "Ejection of hypervelocity stars from the Galactic Centre by intermediate-mass black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 1, pp. 174–182, 2006.

[47] T. Matsubayashi, J. Makino, and T. Ebisuzaki, "Orbital evolution of an IMBH in the galactic nucleus with a massive central black hole," *Astrophysical Journal*, vol. 656, no. 2 I, pp. 879–896, 2007.

[48] J. N. Bahcall and R. A. Wolf, "Star distribution around a massive black hole in a globular cluster," *Astrophysical Journal*, vol. 209, pp. 214–232, 1976.

[49] A. Sesana, F. Haardt, and P. Madau, "Interaction of massive black hole binaries with their stellar environment. III. Scattering of bound stars," *Astrophysical Journal*, vol. 686, no. 1, pp. 432–447, 2008.

[50] M. Iwasawa, S. An, T. Matsubayashi, Y. Funato, and J. Makino, "Ices in the quiescent ic 5146 dense cloud," *The Astrophysical Journal*, vol. 731, no. 1, article 9, 2011.

[51] A. Sesana, A. Gualandris, and M. Dotti, "Massive black hole binary eccentricity in rotating stellar systems," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 415, no. 1, pp. L35–L39, 2011.

[52] M. Dotti, M. Colpi, F. Haardt, and L. Mayer, "Supermassive black hole binaries in gaseous and stellar circumnuclear discs: orbital dynamics and gas accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 3, pp. 956–962, 2007.

[53] I. Berentzen, M. Preto, P. Berczik, D. Merritt, and R. Spurzem, "Binary black hole merger in galactic nuclei: post-newtonian simulations," *The Astrophysical Journal*, vol. 695, no. 1, pp. 455–468, 2009.

[54] L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. Quinn, and J. Wadsley, "Rapid formation of supermassive black hole binaries in galaxy mergers with gas," *Science*, vol. 316, no. 5833, pp. 1874–1877, 2007.

[55] A. Escala, R. B. Larson, P. S. Coppi, and D. Mardones, "The role of gas in the merging of massive black holes in galactic nuclei. II. Black hole merging in a nuclear gas disk," *Astrophysical Journal*, vol. 630, no. 1 I, pp. 152–166, 2005.

[56] M. Dotti, M. Ruszkowski, L. Paredi, M. Colpi, M. Volonteri, and F. Haardt, "Dual black holes in merger remnants—I. Linking accretion to dynamics," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 3, pp. 1640–1646, 2009.

[57] D. N. C. Lin and J. Papaloizou, "Tidal torques on accretion discs in binary systems with extreme mass ratios," *Monthly Notices of the Royal Astronomical Society*, vol. 186, no. L799, 1979.

[58] D. N. C. Lin and J. Papaloizou, "On the structure of circumbinary accretion disks and the tidal evolution of commensurable satellites," *Monthly Notices of the Royal Astronomical Society*, vol. 188, no. L191, 1979.

[59] P. Artymowicz and S. H. Lubow, "Dynamics of binary-disk interaction. I. Resonances and disk gap sizes," *Astrophysical Journal*, vol. 421, no. 2, pp. 651–667, 1994.

[60] D. Syer and J. C. Clarke, "Satellites in discs: regulating the accretion luminosity," *Monthly Notices of the Royal Astronomical Society*, vol. 277, no. 3, pp. 758–766, 1995.

[61] A. Gould and H. W. Rix, "Binary black hole mergers from planet-like migrations," *Astrophysical Journal*, vol. 532, no. 1, pp. L29–L32, 2000.

[62] C. Roedig, M. Dotti, A. Sesana, J. Cuadra, and M. Colpi, "Limiting eccentricity of subparsec massive black hole binaries surrounded by self-gravitating gas discs," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 4, pp. 3033–3041, 2011.

[63] J. Papaloizou and J. E. Pringle, "Tidal torques on accretion discs in close binary systems," *Monthly Notices of the Royal Astronomical Society*, vol. 181, p. 441, 1977.

[64] P. Goldreich and S. Tremaine, "Disk-satellite interactions," *Astrophysical Journal*, vol. 241, pp. 425–441, 1980.

[65] P. B. Ivanov, J. C. B. Papaloizou, and A. G. Polnarev, "The evolution of a supermassive binary caused by an accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 307, no. 1, pp. 79–90, 1999.

[66] P. J. Armitage and P. Natarajan, "Accretion during the merger of supermassive black holes," *Astrophysical Journal*, vol. 567, no. 1, pp. L9–L12, 2002.

[67] P. Goldreich and R. Sari, "Eccentricity evolution for planets in gaseous disks," *Astrophysical Journal*, vol. 585, no. 2 I, pp. 1024–1037, 2003.

[68] P. J. Armitage and P. Natarajan, "Eccentricity of supermassive black hole binaries coalescing from gas-rich mergers," *Astrophysical Journal*, vol. 634, no. 2 I, pp. 921–927, 2005.

[69] Z. Haiman, B. Kocsis, and K. Menou, "The population of viscosity- and gravitational wave-driven supermassive black hole binaries among luminous active galactic nuclei," *Astrophysical Journal*, vol. 700, no. 2, pp. 1952–1969, 2009.

[70] J. Cuadra, P. J. Armitage, R. D. Alexander, and M. C. Begelman, "Massive black hole binary mergers within subparsec scale gas discs," *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 4, pp. 1423–1432, 2009.

[71] N. I. Shakura and R. A. Sunyaev, "Black holes in binary systems. Observational appearance," *Astronomy and Astrophysics*, vol. 24, pp. 337–355, 1973.

[72] G. Lodato, S. Nayakshin, A. R. King, and J. E. Pringle, "Black hole mergers: can gas discs solve the "final parsec" problem?" *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 3, pp. 1392–1402, 2009.

[73] A. Escala, "On the fueling of massive black holes and the properties of their host spheroids," *Astrophysical Journal*, vol. 648, no. 1, pp. L13–L16, 2006.

[74] P. F. Hopkins and E. Quataert, "How do massive black holes get their gas?" *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 3, pp. 1529–1564, 2010.

[75] C. J. Nixon, P. J. Cossins, A. R. King, and J. E. Pringle, "Retrograde accretion and merging supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 3, pp. 1591–1598, 2011.

[76] D. S. Hudson, T. H. Reiprich, T. E. Clarke, and C. L. Sarazin, "X-ray detection of the proto supermassive binary black hole at the centre of Abell 400," *Astronomy and Astrophysics*, vol. 453, no. 2, pp. 433–446, 2006.

[77] A. J. Barth, M. C. Bentz, J. E. Greene, and L. C. Ho, "An offset seyfert 2 nucleus in the minor merger system ngc 3341," *Astrophysical Journal*, vol. 683, no. 2, pp. L119–L122, 2008.

[78] S. Bianchi, M. Chiaberge, E. Piconcelli, M. Guainazzi, and G. Matt, "Chandra unveils a binary active galactic nucleus in Mrk 463," *Monthly Notices of the Royal Astronomical Society*, vol. 386, no. 1, pp. 105–110, 2008.

[79] J. M. Comerford, R. L. Griffith, B. F. Gerke et al., "1.75 h -1 kpc separation dual active galactic nuclei at z = 0.36 in the cosmos field," *The Astrophysical Journal Letters*, vol. 702, no. 1, pp. L82–L86, 2009.

[80] P. J. Green, A. D. Myers, W. A. Barkhouse et al., "SDSS J1254+0846: a binary quasar caught in the act of merging," *Astrophysical Journal*, vol. 710, no. 2, pp. 1578–1588, 2010.

[81] J.-M. Wang, Y.-M. Chen, C. Hu, W.-M. Mao, S. Zhang, and W.-H. Bian, "Active galactic nuclei with double-peaked narrow lines: are they dual active galactic nuclei?" *Astrophysical Journal Letters*, vol. 705, no. 1, pp. L76–L80, 2009.

[82] X. Liu, J. E. Greene, Y. Shen, and M. A. Strauss, "Discovery of four kpc-scale binary active galactic nuclei," *Astrophysical Journal Letters*, vol. 715, no. 1, pp. L30–L34, 2010.

[83] D. G. York et al., "The sloan digital sky survey: technical summary," *The Astronomical Journal*, vol. 120, pp. 1579–1587, 2000.

[84] H. Fu, L. Yan, A. D. Myers et al., "The nature of double-peaked [O III] active galactic nuclei," Submitted to *The Astrophysical Journal*.

[85] H. L. Maness, G. B. Taylor, R. T. Zavala, A. B. Peck, and L. K. Pollack, "Breaking all the rules: the compact symmetric object 0402+379," *Astrophysical Journal*, vol. 602, no. 1 I, pp. 123–134, 2004.

[86] C. Rodriguez, G. B. Taylor, R. T. Zavala, Y. M. Pihlström, and A. B. Peck, "H i observations of the supermassive binary black hole system in 0402+379," *Astrophysical Journal*, vol. 697, no. 1, pp. 37–44, 2009.

[87] S. Burke-Spolaor, "A radio Census of binary supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 4, pp. 2113–2122, 2011.

[88] M. J. Valtonen, H. J. Lehto, K. Nilsson et al., "A massive binary black-hole system in OJ 287 and a test of general relativity," *Nature*, vol. 452, no. 7189, pp. 851–853, 2008.

[89] C. Villforth, K. Nilsson, J. Heidt et al., "Variability and stability in blazar jets on time-scales of years: optical polarization monitoring of OJ 287 in 2005-2009," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 3, pp. 2087–2111, 2010.

[90] S. Komossa, H. Zhou, and H. Lu, "A recoiling supermassive black hole in the quasar SDSS J092712.65+294344.0?" *Astrophysical Journal*, vol. 678, no. 2, pp. L81–L84, 2008.

[91] T. Bogdanović, M. Eracleous, and S. Sigurdsson, "SDSS J092712.65+294344.0: recoiling black hole or a subparsec binary candidate?" *Astrophysical Journal*, vol. 697, no. 1, pp. 288–292, 2009.

[92] M. Dotti, C. Montuori, R. Decarli, M. Volonteri, M. Colpi, and F. Haardt, "SDSSJ092712.65+294344.0: a candidate massive black hole binary," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 398, no. 1, pp. L73–L77, 2009.

[93] T. A. Boroson and T. R. Lauer, "A candidate sub-parsec supermassive binary black hole system," *Nature*, vol. 458, no. 7234, pp. 53–55, 2009.

[94] G. A. Shields, E. W. Bonning, and S. Salviander, "Comment on the black hole recoil candidate quasar SDSS j092712.65+294344.0," *The Astrophysical Journal*, vol. 696, pp. 1367–1373, 2009.

[95] R. Decarli, M. Dotti, C. Montuori, T. Liimets, and A. Ederoclite, "The peculiar optical spectrum of 4c+22.25: imprint of a massive black hole binary?" *The Astrophysical Journal Letters*, vol. 720, no. 1, pp. L93–L96, 2010.

[96] R. S. Barrows, C. H. S. Lacy, D. Kennefick, J. Kennefick, and M. S. Seigar, "Unusual double-peaked emission in the SDSS quasar J093201.60 + 031858.7," *New Astronomy*, vol. 16, no. 2, pp. 122–127, 2011.

[97] P. Tsalmantza, R. Decarli, M. Dotti, and D. W. Hogg, "A systematic search for massive black hole binaries in the Sloan Digital Sky Survey spectroscopic sample," *Astrophysical Journal*, vol. 738, no. 1, article 20, 2011.

[98] P. Tsalmantza and D. W. Hogg, Submitted, 2011.

[99] M. Eracleous, T. A. Boroson, and J. P. Halpern, "A large systematic search for recoiling and close supermassive binary black holes," *Astrophysical Journal Supplement*. In press.

[100] D. P. Schneider, G. T. Richards, P. B. Hall et al., "The sloan digital sky survey quasar catalog. V. Seventh data release," *Astronomical Journal*, vol. 139, no. 6, pp. 2360–2373, 2010.

[101] E. W. Bonning, G. A. Shields, and S. Savliander, "Recoiling black holes in quasars," *The Astrophysical Journal Letters*, vol. 666, no. 1, pp. L13–L16, 2007.

[102] T. Bogdanović, C. S. Reynolds, and M. C. Miller, "Alignment of the spins of supermassive black holes prior to coalescence," *The Astrophysical Journal Letters*, vol. 661, no. 2, pp. L147–L150, 2007.

[103] M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruszkowski, and F. Haardt, "Dual black holes in merger remnants—II. Spin evolution and gravitational recoil," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 682–690, 2010.

[104] M. Volonteri, K. Gültekin, and M. Dotti, "Gravitational recoil: effects on massive black hole occupation fraction over cosmic time," *Monthly Notices of the Royal Astronomical Society*, vol. 404, no. 4, pp. 2143–2150, 2010.

[105] M. Kesden, U. Sperhake, and E. Berti, "Relativistic suppression of black hole recoils," *Astrophysical Journal*, vol. 715, no. 2, pp. 1006–1011, 2010.

[106] M. Eracleous and J. P. Halpern, "Double-peaked emission lines in active galactic nuclei," *Astrophysical Journal, Supplement Series*, vol. 90, no. 1, pp. 1–30, 1994.

[107] T. M. Heckman, J. H. Krolik, S. M. Moran et al., "SDSSJ092712.65+294344.0: NGC 1275 AT z = 0.7?" *The Astrophysical Journal*, vol. 695, no. 1, article 363, 2009.

[108] C. Montuori, M. Dotti, M. Colpi, R. Decarli, and F. Haardt, "Search for sub-parsec massive binary black holes through line diagnosis," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 1, pp. 26–32, 2011.

[109] Y. Shen and A. Loeb, "Identifying supermassive black hole binaries with broad emission line diagnosis," *Astrophysical Journal Letters*, vol. 725, no. 1, pp. 249–260, 2010.

[110] P. Artymowicz and S. H. Lubow, "Mass flow through gaps in circumbinary disks," *Astrophysical Journal*, vol. 467, no. 2, pp. L77–L80, 1996.

[111] K. Hayasaki, S. Mineshige, and L. C. Ho, "A supermassive binary black hole with triple disks," *Astrophysical Journal*, vol. 682, no. 2, pp. 1134–1140, 2008.

[112] A. Sesana, C. Roedig, M. T. Reynolds, and M. Dotti, "Mul timessenger astronomy with pulsar timing and X-ray observations of massive black hole binaries," *Monthly Notices of the Royal Astronomical Society: Letters*, 2011.

[113] T. Tanaka, K. Menou, and Z. Haiman, Submitted to MNRAS (arXiv:11072937) , 2011.

[114] A. I. Macfadyen and M. Milosavljevic, "An eccentric circumbinary accretion disk and the detection of binary massive black holes," *Astrophysical Journal*, vol. 672, no. 1, pp. 83–93, 2008.

[115] M. Dotti, M. Colpi, and F. Haardt, "Laser Interferometer Space Antenna double black holes: dynamics in gaseous nuclear discs," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 1, pp. 103–112, 2006.

[116] J. D. Schnittman, "Electromagnetic counterparts to black hole mergers," *Classical and Quantum Gravity*, vol. 28, no. 9, article 094021, 2011.

[117] B. D. Farris, Y. T. Liu, and S. L. Shapiro, "Binary black hole mergers in gaseous disks: simulations in general relativity," *Physical Review D*, vol. 84, no. 2, article 024024, 2011.

[118] T. Bode, R. Haas, T. Bogdanović, P. Laguna, and D. Shoemaker, "Relativistic mergers of supermassive black holes and their electromagnetic signatures," *Astrophysical Journal*, vol. 715, no. 2, pp. 1117–1131, 2010.

[119] C. Palenzuela, L. Lehner, and S. L. Liebling, "Dual jets from binary black holes," *Science*, vol. 329, no. 5994, pp. 927–930, 2010.

[120] P. Moesta, D. Alic, L. Rezzolla, O. Zanotti, and C. Palenzuela, "On the detectability of dual jets from binary black holes," *General Relativity and Quantum Cosmology*. In press.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 364973, 8 pages
doi:10.1155/2012/364973

Review Article

Recoiling Black Holes: Electromagnetic Signatures, Candidates, and Astrophysical Implications

S. Komossa[1, 2, 3]

[1] *Fakultät für Physik, Technische Universität München, James Franck Straße 1/I, 85748 Garching, Germany*
[2] *Excellence Cluster Universe, TUM, Boltzmannstraße 2, 85748 Garching, Germany*
[3] *Max-Planck-Institut für Plasmaphysik, Boltzmannstraße 2, 85748 Garching, Germany*

Correspondence should be addressed to S. Komossa, stefanie.komossa@gmx.de

Received 11 November 2011; Accepted 5 January 2012

Academic Editor: Francesca Civano

Supermassive black holes (SMBHs) may not always reside right at the centers of their host galaxies. This is a prediction of numerical relativity simulations, which imply that the newly formed single SMBH, after binary coalescence in a galaxy merger, can receive kick velocities up to several 1000 km/s due to anisotropic emission of gravitational waves. Long-lived oscillations of the SMBHs in galaxy cores, and in rare cases even SMBH ejections from their host galaxies, are the consequence. Observationally, accreting recoiling SMBHs would appear as quasars spatially and/or kinematically offset from their host galaxies. The presence of the "kicks" has a wide range of astrophysical implications which only now are beginning to be explored, including consequences for black hole and galaxy assembly at the epoch of structure formation, black hole feeding, and unified models of active galactic nuclei (AGN). Here, we review the observational signatures of recoiling SMBHs and the properties of the first candidates which have emerged, including follow-up studies of the candidate recoiling SMBH of SDSSJ092712.65+294344.0.

1. Introduction

Interaction and merging of galaxies occurs frequently throughout the history of the universe. If both galaxies do harbor SMBHs, binaries will inevitably form [1]. Galaxy mergers are believed to be the sites of major black hole growth, and an active search for SMBH pairs and binaries of wide and small separations is currently ongoing (see [2] for a review of electromagnetic signatures). When the two SMBHs ultimately coalesce, they are a source of strong gravitational waves. These are emitted anisotropically during coalescence and carry away linear momentum (e.g., [3]). As a result, the newly formed single SMBH recoils. Configurations of coalescing black holes can lead to kick velocities up to several thousand km/s (e.g., [4–16]; review by [17]). In the initial computations, kick velocity was highest for maximally spinning equal-mass black hole binaries with antialigned spins in the orbital plane ("superkicks"). More recently, based on a new recoil formula, Lousto and Zlochower [18] have estimated that recoil velocities up to 5000 km/s can be reached in configurations with spins partially aligned with the orbital angular momentum. In unbound encounters (not likely to occur in astrophysical environments), the kick velocity can exceed 15 000 km/s [19, 20].

After the kick, the recoiling SMBH will oscillate about the core of its host galaxy [21, 22] or will even escape, if its kick velocity exceeds the escape velocity of its host. In a "typical," gas-poor galaxy, a black hole kick velocity of 500 km/s will result in an initial amplitude of ~200 pc, and an oscillation timescale of order 10^7 yrs (Figure 1 of [23]). The kicks, including those large enough to remove SMBHs from their host galaxies, have potentially far-reaching astrophysical consequences, including for SMBH and galaxy assembly and AGN statistics. Upon recoil, the most tightly bound gas will remain bound to the recoiling black hole, and therefore high-velocity kicks imply the existence of interstellar and intergalactic quasars (e.g., [21–30]). Identifying recoiling SMBHs through observations is of great interest. Several key electromagnetic signatures of kicks have been predicted in the last few years, and first candidate recoiling SMBHs have emerged. This chapter is structured as follows. In Section 2, an overview of the predicted electromagnetic signatures of

recoiling SMBHs is given. In Section 3 the event frequency is discussed, while Sections 4 and 5 provide a review of the published candidate recoiling SMBHs. Section 6 explores consequences of recoil for unified models of AGN. Section 7 concludes with some astrophysical consequences and important future studies.

2. Electromagnetic Signatures of Recoiling SMBHs

2.1. Broad Emission-Line Shifts. After the kick, matter remains bound to the recoiling SMBH within a region whose radius r_k is given by

$$r_k = \frac{GM_{BH}}{v_k^2} \approx 0.4\left(\frac{M_{BH}}{10^8\,M_\odot}\right)\left(\frac{v_k}{10^3\,\mathrm{km\,s^{-1}}}\right)^{-2}\mathrm{pc}, \quad (1)$$

where v_k is the kick velocity [31]. This region is on the order of the size of the broad line region (BLR) of AGN [32]. The accretion disk and BLR will therefore typically remain bound to the SMBH while the bulk of the host galaxy's narrow-line region (NLR) will remain behind. The accreting recoiling SMBH will therefore appear as an off-nuclear "quasar" as long as its accretion supply lasts. However, spatial offsets are challenging to detect even with the *Hubble Space Telescope* (HST) except in the nearby universe. The kinematic Doppler shifts of the broad emission lines are, in principle, easy to measure out to high redshifts. Spectroscopically, recoiling SMBHs will appear as AGN which have their broad emission lines kinematically shifted by up to ∼5000 km/s with respect to their NLRs.

Bonning et al. [33] suggested several criteria, how to identify a recoiling SMBH spectroscopically. Apart from (1) the kinematic shift of the BLR, a candidate recoiling SMBH should (2) show symmetric broad line profiles, it should (3) lack an ionization stratification of its narrow emission lines, and it should (4) not show any shift between broad MgII and the broad Balmer lines. (In practice, individual recoil candidates may show some (temporary) deviations from this scheme, or exhibit extra features. For instance, just after recoil, the BLR emission profiles would likely be asymmetric. Feedback trails from partially bound gas and disk winds would produce emission-line signatures at various kinematic shifts between zero and the recoil velocity. Once the SMBH has travelled beyond the extent of the classical NLR of a few kpc extent, low-density "halo" gas would dominate the optical narrow-line spectrum, with emission-line ratios characteristically different from the classical NLR.) One object, the quasar SDSSJ092712.65+294344.0, fulfills all of these four criteria and is therefore an excellent candidate for a recoiling SMBH [34]. It will be further discussed in Section 4, together with several other candidate recoiling BHs. More candidates may hide in large samples of peculiar broad-line emitters recently identified in the Sloan Digital Sky Survey (SDSS; [35]).

2.2. Flaring Accretion Disks. In gas-rich mergers, an accretion disk is likely present, even though the inner part of the disk may only re-form after binary coalescence [27, 36–38]. UV, soft X-ray, and IR flares could result from shocks in the accretion disk surrounding the SMBH just after recoil, or when the inner disk reforms (e.g., [38–46]). These flares may last $\approx 10^4$ yrs and may be detectable in current and future sky surveys.

2.3. Tidal Disruption Flares from Stars Around Recoiling SMBHs. Even in the absence of an accretion disk, ejected SMBHs will always carry a retinue of bound stars. Observable effects related to these stars are therefore perhaps the most universal signature of recoil. As the SMBH moves through the galaxy, the bound, and unbound, stars are subject to tidal disruption, leading to powerful X-ray flares of quasar-like luminosity [47, 48], which would appear off-nuclear or even intergalactic. Komossa and Merritt [49] computed disruption rates for the bound, and the unbound, stellar populations under recoil conditions. In the resonant relaxation regime, they showed that the rates are of order $10^{-6}\,\mathrm{yr}^{-1}$ for a typical postmerger galaxy (Figure 2 of [49]); smaller than, but comparable to, rates for nonrecoiling SMBHs. At an early phase of recoil, the tidal disruption rate can be much higher, when the SMBH experiences a full loss cone, and travels through the clumpy core environment of a recent merger [49]. The flare rate may temporarily reach values as high [50] as during the peak of the premerger binary phase [51].

Another signature related to the stars bound to the recoiling SMBH is episodic X-ray emission from accretion due to stellar mass loss. Mass loss provides a reservoir of gas, and therefore also *optical emission lines from gas at the recoil velocity* even in the initial absence of a gaseous accretion disk. Other consequences include the presence of intergalactic planetary nebulae and supernovae, after the ejected SMBH has left its host galaxy [49].

All these signals would generically be associated with recoiling SMBHs, whether or not the galaxy merger is gas-rich or dry, and whether or not an accretion disk is present initially, and they would continue episodically for a time of ∼10 Gyr [49].

2.4. Hypercompact Stellar Systems. While the "tidal recoil flares" are very luminous and can be detected out to very large distances, the compact system of bound stars itself will be detectable in the nearby universe, and would resemble a globular cluster in total luminosity, but with a much greater velocity dispersion due to the large binding mass M_{BH} [49]. Merritt et al. [52] worked out the properties of these "hypercompact stellar systems" (HCSSs), and related the structural properties (mass, size, and density profile) of HCSSs to the properties of their host galaxies and to the amplitude of the kick. Since the kick velocity is encoded in the velocity dispersion of the bound stars, future detection of large samples of HCSSs would therefore allow us to determine empirically the kick distribution, and therefore the merger history of galaxies in clusters. Nearby clusters of galaxies are best suited to search for and identify HCSSs, and ∼100 of them should be detectable within 2 Mpc of the center of the Virgo cluster [52]. Depending on the merger history of our Milky Way [53], and the merger history of black holes in its globular clusters [54], 100 s of low-mass HCSSs and rogue black holes

may reside in the halo of our Milky Way, and a search for them is underway [55].

2.5. Other Observable Effects of Recoil. During the long-lived "Phase II" recoil oscillations [22], when the SMBH oscillation amplitude is on the torus scale, the SMBH might efficiently accrete from the dense molecular gas at *each* turning point, causing repeated flares of radiation [23]. Such flares would locally destroy the dust, while photoionization of the dense surrounding gas would produce a strong emission-line response. Such a signal would not only help in identifying kicks but also could be used as a new probe of the properties of the torus itself.

Other signatures of recoiling SMBHs include effects on the morphology and dynamics of the gaseous disk of the host galaxy [56], their imprints on the hot gas in early-type galaxies [57], accretion from the ISM [58], the possibility of star formation in the wake of the SMBH trajectory [59], and their influence on the jet structures in radio galaxies [30].

3. The Frequency of Recoiling SMBHs in Astrophysical Environments

Several factors affect the distribution of SMBH kick velocities and their observability; the system parameters of the SMBH binary on the one hand (mass ratio, spin magnitudes and spin directions), and the astrophysical environment on the other hand.

The frequency of high-velocity kicks depends on the distribution of mass ratios and spins of the binary SMBHs. In case of random distributions of spin directions, as expected in gas-poor galaxy mergers, the kick formula (e.g., [4, 9, 13]) has been used to predict the kick fraction in dependence of recoil velocity [4, 9, 23, 60]. In this case, kicks with velocities larger than 500 km/s are relatively common (Figure 1 of [23]). Spin precession further has the consequence that large kicks are deboosted if the angle Θ between the spin of the more massive BH and the orbital angular momentum is initially small, while large kicks are boosted, if Θ is initially large [61].

The other key factor is the astrophysical environment, which determines the spin magnitude (set by the mechanism of SMBH mass growth via random accretion, ordered accretion, or BH-BH merging; [62]) and the timescale of spin alignment with the orbital angular momentum (e.g., [63–65]) in gas-rich galaxy mergers. The latter depends on the rapidity of binary coalescence, the amount of gas accretion before versus after coalescence, the accretion rate, the disk properties (e.g., the viscosity law across the disk), and the mass of the SMBH. While the most massive black holes are more resistant to alignment, the process is generally relatively fast in gas rich environments (timescales of 10^5–10^9 yrs or less) [66, 67].

While initial results from numerical relativity have indicated that kick velocities are low in this case, the whole parameter space is still being explored, and Lousto and Zlochower [18] have recently shown that kick velocities up to 5000 km/s can be reached in configurations with spins partially aligned with the orbital angular momentum. As a consequence, the likelihood of observing high-velocity recoils in gas-rich environments is significantly higher than in some previous estimates (their Figure 3).

Given the large number of uncertain parameters in estimating the frequency of recoiling SMBHs, identifying them through observations is also important. Ultimately, observations will independently provide the distribution of kick velocities and the properties of the recoiling SMBHs. First candidates have emerged in recent years, and more are likely to come soon, given the number of operating or planned very large spectroscopic and time-domain surveys, like SDSS, LAMOST, LSST, and future X-ray surveys.

4. Candidate Recoiling SMBHs Identified by Kinematic Signatures

4.1. SDSSJ092712.65+294344.0 and X-Ray Followups. The quasar SDSSJ092712.65+294344.0 (SDSSJ0927+2943 hereafter) at $z = 0.7$ shows all the characteristic optical signatures of a recoiling SMBH [34], which were predicted earlier [33]. Its broad emission lines are shifted by 2650 km s^{-1} with respect to its narrow emission lines, the broad lines are symmetric, the broad MgII line shows the same shift as the broad Balmer lines, and the narrow emission lines lack an ionization stratification as expected if the accreting SMBH is no longer at the center of the system [34]. (SDSSJ0927+2943 also shows a second system of narrow emission lines with unusual properties when compared with other known quasars, including exceptionally broad Neon emission lines. The origin of these lines is still being explored; the lower-ionization lines are too narrow to have originally been bound to the recoiling SMBH (except in case of projection effects), and their low degree of ionization is not straightforward to understand [34]. A possible reservoir of narrow-line gas at the kick velocity is stellar mass loss, as a consequence of stellar evolution of the stars bound to the recoiling SMBH [49].) Its unique properties make SDSSJ0927+2943 an excellent candidate for a recoiling SMBH.

Two alternative models have been considered in order to explain some (but not all) of the unusual properties of this system; a chance projection, within 1 arcsec, of one or *two intrinsically peculiar* AGN in a very massive cluster of galaxies [34, 68, 69], and a close premerger binary SMBH [70, 71]. However, a rich and massive cluster has not been detected in NIR and X-ray imaging follow-up observations [72], (Komossa et al. 2012, in prep.). Neither was the predicted orbital motion of an SMBH binary detected in spectroscopic followups [68]; see also [73]. This leaves us with the recoil scenario for SDSSJ0927+2943. This scenario is also consistent with the recent measurement of an *offset* between the QSO and the host galaxy as traced by [OIII] emission [73].

We have obtained an imaging observation of SDSSJ0927+2943 with the *Chandra* X-ray observatory, in order to measure more precisely its X-ray luminosity (than was possible with a serendipitous off-axis ROSAT observation; [34]), and to study the properties of the field around SDSSJ-0927+2943, including the search for a possible massive cluster of galaxies. We detect point-like X-ray emission from the quasar coincident with the optical position of SDSSJ-0927+2943. A second X-ray source is present at a distance of

~17 arcsec from SDSSJ0927+2943. This second source coincides with the object SDSSJ092713.8+294336 and contributed approximately 70% to the ROSAT X-ray emission from the region of SDSSJ0927+2943. Luminous extended X-ray emission from a *rich* cluster, in the form predicted by Heckman et al. [69], is not present. The full results of the X-ray analysis will be presented by (Komossa et al. 2012, in prep.).

4.2. E1821+643. The well-known luminous quasar E1821+643 ($z = 0.297$) shows highly asymmetric broad Balmer lines which appear different in direct and in polarized light, and are strongly shifted with respect to the narrow lines. Based on their spectropolarimetric observations, Robinson et al. [74] favor a scenario where one component of the BLR is bound to a recoiling black hole, which is moving at 2100 km/s relative to its host galaxy. A second broad-line system is shifted by only 470 km/s, and its nature is currently unclear. If still related to recoil, in form of a marginally bound or unbound component of the BLR, the system is young, and Robinson et al. then estimate an age of $\sim 10^4$ years.

4.3. SDSSJ105041.35+345631.3. Shields et al. [68, 75] selected the quasar SDSSJ105041.35+345631.3 at $z = 0.272$ from the SDSS because of its large kinematic shift of the BLR of, in this case, 3500 km/s relative to the narrow emission lines. A projection effect is considered unlikely, as is a binary SMBH because of the lack of detectable orbital motion. While Shields et al. do not rule out an extreme case of a recoiling SMBH, they conclude that several aspects of the optical spectrum are best understood if this galaxy is an extreme case of a "double-peaked emitter."

5. Candidate Recoiling SMBHs Identified by Spatial Offsets

5.1. CID-42. The galaxy CID-42 (COSMOSJ1000+0206) at redshift $z = 0.359$ was discovered in the COSMOS survey [76], and caught attention due to its unusual morphology with two apparent optical "nuclei" [77] at a projected separation of 2 kpc, and an extended tidal tail. Initially suspected to be a binary AGN [77], it was then reinterpreted as a candidate recoiling SMBH, or alternatively, an SMBH ejection following 3-body interaction in a triple SMBH system, by Civano et al. [78]. An HST image analysis has shown that the northwestern core is slightly extended though compact, and consistent with being the nucleus of the galaxy, while only the southeastern bright source is point-like [78]. The optical spectrum of CID 42 shows a kinematic shift of 1200 km/s between the BLR and the major narrow-line component, and extra faint narrow Hβ emission at the same redshift as the broad lines [78], and perhaps further faint narrow-line emission shifted by ~150 km/s [77]. As such, the spectrum shares similarities with that of the recoil candidate SDSSJ092712.65+294344.0 [34, 68]. Another remarkable feature, not yet well understood, is the presence of a strong redshifted broad iron line with a P-Cygni profile, variable over four years, of high column density and highly ionized [78]. Follow-up optical and X-ray observations are currently underway (Civano et al. 2012, in prep.). Their new high-resolution Chandra data show the presence of only one X-ray emitting object which coincides with the position of the southeastern optical source, supporting the recoil scenario rather than the binary SMBH (Civano et al. 2012, in prep.).

5.2. M87. M87 is a nearby massive galaxy with a prominent radio jet. The photocenter of the host galaxy is offset by 7 pc from the nuclear point source (i.e., presumably the location of the SMBH) [79]. The displacement is in the direction of the counterjet. Among several scenarios (acceleration by a jet, presence of massive perturbers, binary orbital motion) considered, Batcheldor et al. [79] favor GW recoil as the most plausible. The observed offset can then be explained either by a moderate kick 1 Myr ago, or residual small-amplitude oscillations of a large recoil which happened <1 Gyr ago.

5.3. CXOJ122518.6+144545. Jonker et al. [80] reported the detection of an unusual off-nuclear X-ray source, at a projected separation of 3 kpc from the core of the galaxy SDSSJ122518.86+144547.7 at $z = 0.045$. CXOJ122518.6+144545 is X-ray luminous and has a bright optical counterpart and properties unlike those of other off-nuclear X-ray sources which were found in large numbers with *Chandra*. The authors offer three explanations of CXOJ122518.6+144545: a supernova of type IIn, an ultraluminous X-ray source with an unusually bright optical counterpart, or a recoiling SMBH. Bellovary et al. [81] further discuss the possibility of a wandering SMBH in the galaxy halo, produced by stripping of a satellite which merged with the primary galaxy.

5.4. ESO 1327–2041. The nearby galaxy ESO 1327–2041 ($z = 0.018$) shows a complex morphology indicative of a recent merger. *HST* imaging has revealed the presence of a compact source embedded in an extended "stellar stream" ([82]; their Figure 1), at a redshift similar to the core of ESO 1327–2041, and at a projected separation of 15 kpc. Keeney et al. discuss several possible interpretations of this compact object, and propose that it is the actual nucleus of the galaxy, ejected as a consequence of either tidal interaction between two galaxies or gravitational wave recoil following a past merger.

6. Implications of Recoil Oscillations for Unified Models of AGN

There are potentially far-reaching consequences of SMBH recoil for unified models of AGN. Spatial oscillations of the SMBHs about the cores of their host galaxies imply that the SMBHs spend a significant fraction of time *off-nucleus*, at scales beyond that of the molecular obscuring torus. An intrinsically *obscured* quasar of *type 2* with its BLR hidden by the torus will therefore appear as *unabsorbed, type 1* quasar during the recoil oscillations, when moving beyond the torus scale.

Assuming reasonable distributions of recoil velocities, Komossa and Merritt [23] have computed the off-core timescale of (intrinsically type 2) quasars (These calculations are based on models of Gualandris and Merritt [22], which did not include gas. Recent simulations of recoil oscillations in

a gaseous disk show, that oscillation timescales can either increase or decrease [83] with respect to the gas-free case.). It was shown that roughly 50% of all major mergers result in a SMBH being displaced beyond the torus for a time of $10^{7.5}$ yr or more. This is an interesting number, because it is comparable to quasar activity time scales. Since *major* mergers (i.e., quasars) are most strongly affected by gravitational wave recoil, our results imply a deficiency of luminous type 2 *quasars* in comparison to low-luminosity *Seyfert* 2 galaxies, as indeed observed (e.g., [84]). These may therefore naturally explain the long-standing puzzle, why few absorbed type 2 quasars exist at high luminosities; it would be these which are affected by the recoil oscillations, therefore appearing as type 1 rather than type 2 for a significant fraction of their lifetime [23].

Recoil oscillations further imply the presence of a fraction of quasars which lack a hot dust component, because the dusty torus is only illuminated from a distance. Such "hot-dust-poor" quasars have indeed been observed (e.g., [85]).

Recoil oscillations also have a number of other observable consequences related to AGN. For instance, they will affect the X-ray background and its modeling since a fraction of sources will be unobscured at any given time. In particular, small amplitude oscillations of the order the torus size will affect the ratio of Compton-thin to Compton-thick sources and could lead to measurable variability in the absorption and extinction of AGN spectra once the recoiling SMBH passes the individual clouds making up the torus [23].

7. Astrophysical Implications and Future Observations

The kicks and superkicks predicted by recent numerical relativity simulations of coalescing SMBHs have stimulated an active new field of research. Electromagnetic signatures of recoiling SMBHs are being predicted, several candidates have emerged in large data bases, and astrophysical implications of the kicks are still being explored. The fact that SMBHs will not always reside at the very cores of their host galaxies, or may even be ejected completely, has many potential implications for the topics discussed in this book; for galaxy and SMBH assembly and galaxy-SMBH (co)evolution, core structures in early type galaxies, the scatter in the host galaxy-SMBH scaling relations, the statistics of obscured quasars, and the redshift-dependence of gravitational wave signals (e.g., [26, 86–92]).

It is therefore important to identify more candidate recoiling SMBHs through observations. Promising future searches would include (1) emission-line signatures in large spectroscopic data bases such as SDSS or LAMOST; (2) recoil flares from accretion disks and stellar tidal disruptions in large-scale surveys like Pan-STARRS and LSST and in the X-ray regime; (3) the characteristic, large stellar velocity dispersions of HCSSs in spectroscopic followups of ongoing imaging surveys of nearby clusters of galaxies.

Detecting recoiling SMBHs in large numbers will open up a new window on measuring galaxy merger histories and kick amplitude distributions, and testing predictions of numerical relativity.

References

[1] M. C. Begelman, R. D. Blandford, and M. J. Rees, "Massive black hole binaries in active galactic nuclei," *Nature*, vol. 287, no. 5780, pp. 307–309, 1980.

[2] S. Komossa, "Observational evidence for binary black holes and active double nuclei," *Memorie della Società Astronomica Italiana*, vol. 77, p. 733, 2006.

[3] J.D. Bekenstein, "Gravitational-radiation recoil and runaway black holes," *The Astrophysical Journal*, vol. 183, pp. 657–664, 1973.

[4] M. Campanelli, C. Lousto, Y. Zlochower, and D. Merritt, "Large merger recoils and spin flips from generic black hole binaries," *The Astrophysical Journal*, vol. 659, no. 1, pp. L5–L8, 2007.

[5] M. Campanelli, C. O. Lousto, H. Nakano, and Y. Zlochower, "Comparison of numerical and post-Newtonian waveforms for generic precessing black-hole binaries," *Physical Review D*, vol. 79, no. 8, Article ID 084010, 2009.

[6] J. A. González, M. Hannam, U. Sperhake, B. Brügmann, and S. Husa, "Supermassive recoil velocities for binary black-hole mergers with antialigned spins," *Physical Review Letters*, vol. 98, no. 23, Article ID 231101, 2007.

[7] J. A. González, U. Sperhake, and B. Brügmann, "Black-hole binary simulations: the mass ratio 10 : 1," *Physical Review D*, vol. 79, no. 12, Article ID 124006, 2009.

[8] F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner, "Binary black holes: spin dynamics and gravitational recoil," *Physical Review D*, vol. 76, no. 8, Article ID 084032, 2007.

[9] J. G. Baker, W. D. Boggs, J. Centrella et al., "Modeling kicks from the merger of generic black hole binaries," *The Astrophysical Journal*, vol. 682, no. 1, pp. L29–L32, 2008.

[10] B. Brügmann, J. A. González, M. Hannam, S. Husa, and U. Sperhake, "Exploring black hole superkicks," *Physical Review D*, vol. 77, no. 12, Article ID 124047, 2008.

[11] S. Dain, C. O. Lousto, and Y. Zlochower, "Extra-large remnant recoil velocities and spins from near-extremal-Bowen- York-spin black-hole binaries," *Physical Review D*, vol. 78, no. 2, Article ID 024039, 2008.

[12] S. H. Miller and R. A. Matzner, "Multipole analysis of kicks in collision of spinning binary black holes," *General Relativity and Gravitation*, vol. 41, no. 3, pp. 525–539, 2009.

[13] C. O. Lousto and Y. Zlochower, "Modeling gravitational recoil from precessing highly spinning unequal-mass black-hole binaries," *Physical Review D*, vol. 79, no. 6, Article ID 064018, 2009.

[14] A. Le Tiec, L. Blanchet, and C. M. Will, "The gravitational-wave recoil from the ringdown phase of coalescing black hole binaries," *Classical and Quantum Gravity*, vol. 27, no. 1, Article ID 012001, 2010.

[15] C. O. Lousto, M. Campanelli, Y. Zlochower, and H. Nakano, "Remnant masses, spins and recoils from the merger of generic black hole binaries," *Classical and Quantum Gravity*, vol. 27, no. 11, Article ID 114006, 2010.

[16] C. O. Lousto and Y. Zlochower, "Modeling maximum astrophysical gravitational recoil velocities," *Physical Review D*, vol. 83, no. 2, 2011.

[17] J. Centrella, J. G. Baker, B. J. Kelly, and J. R. van Meter, "Black-hole binaries, gravitational waves, and numerical relativity," *Reviews of Modern Physics*, vol. 82, no. 4, pp. 3069–3119, 2010.

[18] C. Lousto and Y. Zlochower, "Hangup kicks: still larger recoils by partial spin/orbit alignment of black-hole binaries," *Physical Review Letters*, vol. 107, Article ID 231102, 2011.

[19] J. Healy, F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner, "Superkicks in hyperbolic encounters of binary black holes," *Physical Review Letters*, vol. 102, no. 4, Article ID 041101, 2009.

[20] U. Sperhake, E. Berti, V. Cardoso, F. Pretorius, and N. Yunes, "Superkicks in ultrarelativistic encounters of spinning black holes," *Physical Review D*, vol. 83, no. 2, Article ID 024037, 2011.

[21] P. Madau and E. Quataert, "The effect of gravitational-wave recoil on the demography of massive black holes," *The Astrophysical Journal*, vol. 606, no. 1, pp. L17–L20, 2004.

[22] A. Gualandris and D. Merritt, "Ejection of supermassive black holes from galaxy cores," *The Astrophysical Journal*, vol. 678, no. 2, pp. 780–797, 2008.

[23] S. Komossa and D. Merritt, "Gravitational wave recoil oscillations of black holes: implications for unified models of active galactic nuclei," *The Astrophysical Journal*, vol. 689, no. 2, pp. L89–L92, 2008.

[24] P. Madau, M. J. Rees, M. Volonteri, F. Haardt, and S. P. Oh, "Early reionization by miniquasars," *The Astrophysical Journal*, vol. 604, no. 2, pp. 484–494, 2004.

[25] D. Merritt, M. Milosavljević, M. Favata, S. A. Hughes, and D. E. Holz, "Consequences of gravitational radiation recoil," *The Astrophysical Journal*, vol. 607, no. 1, pp. L9–L12, 2004.

[26] N. I. Libeskind, S. Cole, C. S. Frenk, and J. C. Helly, "The effect of gravitational recoil on black holes forming in a hierarchical universe," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 3, pp. 1381–1391, 2006.

[27] A. Loeb, "Observable signatures of a black hole ejected by gravitational-radiation recoil in a Galaxy merger," *Physical Review Letters*, vol. 99, no. 4, Article ID 041103, 2007.

[28] L. Blecha and A. Loeb, "Effects of gravitational-wave recoil on the dynamics and growth of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 4, pp. 1311–1325, 2008.

[29] M. Volonteri and P. Madau, "Off-nuclear AGNs as a signature of recoiling massive black holes," *The Astrophysical Journal*, vol. 687, no. 2, Article ID 10.1086/593353, p. L57, 2008.

[30] F. K. Liu, D. Wang, and X. Chen, "Recoiling supermassive black holes in spin-flip radio galaxies," *The Astrophysical Journal*, http://arxiv.org/abs/1112.1081.

[31] D. Merritt, T. Storchi-Bergmann, A. Robinson, D. Batcheldor, D. Axon, and R. C. Fernandes, "The nature of the HE0450-2958 system," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 4, pp. 1746–1750, 2006.

[32] B. M. Peterson, "The masses of black holes in active galactic nuclei," in *Proceedings of the Central Engine of Active Galactic Nuclei*, vol. 373 of *ASP Conference Series*, p. 3, Jioatong University, October 2006.

[33] E. W. Bonning, G. A. Shields, and S. Savliander, "Recoiling black holes in quasars," *The Astrophysical Journal*, vol. 666, no. 1, pp. L13–L16, 2007.

[34] S. Komossa, H. Zhou, and H. Lu, "A recoiling supermassive black hole in the quasar SDSS J092712.65+294344.0?" *The Astrophysical Journal*, vol. 678, no. 2, pp. L81–L84, 2008.

[35] M. Eracleous, T. A. Boroson, J. P. Halpern, and J. Liu, "A large systematic search for recoiling and close supermassive binary black holes," http://arxiv.org/abs/1106.2952.

[36] F. K. Liu, X. B. Wu, and S. L. Cao, "Double-double radio galaxies: remnants of merged supermassive binary black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 340, no. 2, pp. 411–416, 2003.

[37] M. Milosavljević and E. S. Phinney, "The afterglow of massive black hole coalescence," *The Astrophysical Journal*, vol. 622, no. 2, pp. L93–L96, 2005.

[38] M. Ponce, J.A. Faber, and J.C. Lombardi, "Accretion disks around kicked black holes: post-kick dynamics," *The Astrophysical Journal, http://arxiv.org/abs/1107.1711*, vol. 745, p. 71, 2012.

[39] Z. Lippai, Z. Frei, and Z. Haiman, "Prompt shocks in the gas disk around a recoiling supermassive black hole binary," *The Astrophysical Journal Letters*, vol. 676, no. 1, pp. L5–L8, 2008.

[40] G. A. Shields and E. W. Bonning, "Powerful flares from recoiling black holes in quasars," *The Astrophysical Journal*, vol. 682, no. 2, pp. 758–766, 2008.

[41] J. D. Schnittman and J. H. Krolik, "The infrared afterglow of supermassive black hole mergers," *The Astrophysical Journal*, vol. 684, no. 2, pp. 835–844, 2008.

[42] M. Megevand, M. Anderson, J. Frank et al., "Perturbed disks get shocked: Binary black hole merger effects on accretion disks," *Physical Review D*, vol. 80, no. 2, Article ID 024012, 2009.

[43] E. M. Rossi, G. Lodato, P. J. Armitage, J. E. Pringle, and A. R. King, "Black hole mergers: the first light," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 3, pp. 2021–2035, 2010.

[44] L. R. Corrales, Z. Haiman, and A. MacFadyen, "Hydrodynamical response of a circumbinary gas disc to black hole recoil and mass loss," *Monthly Notices of the Royal Astronomical Society*, vol. 404, no. 2, pp. 947–962, 2010.

[45] T. Tanaka, Z. Haiman, and K. Menou, "Witnessing the birth of a quasar," *Astronomical Journal*, vol. 140, no. 2, pp. 642–651, 2010.

[46] O. Zanotti, C. Roedig, L. Rezzolla, and L. del Zanna, "General relativistic radiation hydrodynamics of accretion flows—I. Bondi-Hoyle accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 417, no. 4, pp. 2899–2915, 2011.

[47] S. Komossa and N. Bade, "The giant X-ray outbursts in NGC 5905 and IC 3599: follow-up observations and outburst scenarios," *Astronomy and Astrophysics*, vol. 343, no. 3, pp. 775–787, 1999.

[48] J. S. Bloom, D. Giannios, B. D. Metzger et al., "A possible relativistic jetted outburst from a massive black hole fed by a tidally disrupted star," *Science*, vol. 333, no. 6039, pp. 203–206, 2011.

[49] S. Komossa and D. Merritt, "Tidal disruption flares from recoiling supermassive black holes," *The Astrophysical Journal*, vol. 683, no. 1, pp. L21–L24, 2008.

[50] N. Stone and A. Loeb, "Prompt tidal disruption of stars as an electromagnetic signature of supermassive black hole coalescence," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 1, pp. 75–80, 2011.

[51] X. Chen, P. Madau, A. Sesana, and F. K. Liu, "Enhanced tidal disruption rates from massive black hole binaries," *The Astrophysical Journal*, vol. 697, no. 2, pp. L149–L152, 2009.

[52] D. Merritt, J. D. Schnittman, and S. Komossa, "Hypercompact stellar systems around recoiling supermassive black holes," *The Astrophysical Journal*, vol. 699, no. 2, pp. 1690–1710, 2009.

[53] R. M. O'Leary and A. Loeb, "Star clusters around recoiled black holes in the Milky Way halo," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 2, pp. 781–786, 2009.

[54] K. Holley-Bockelmann, K. Gültekin, D. Shoemaker, and N. Yunes, "Gravitational wave recoil and the retention of intermediate-mass black holes," *The Astrophysical Journal*, vol. 686, no. 2, pp. 829–837, 2008.

[55] R. M. O'Leary and A. Loeb, "Recoiled star clusters in the Milky Way halo: N-body simulations and a candidate search through SDSS," http://arxiv.org/abs/1102.3695.

[56] D. A. Kornreich and R. V. E. Lovelace, "Dynamics of kicked and accelerated massive black holes in galaxies," *The Astrophysical Journal*, vol. 681, no. 1, pp. 104–112, 2008.

[57] B. Devecchi, E. Rasia, M. Dotti, M. Volonteri, and M. Colpi, "Imprints of recoiling massive black holes on the hot gas of early-type galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 2, pp. 633–640, 2009.

[58] Y. Fujita, "Long-term evolution of and X-ray emission from a recoiling supermassive black hole in a disk galaxy," *The Astrophysical Journal*, vol. 691, no. 2, pp. 1050–1057, 2009.

[59] R. De La Fuente Marcos and C. De La Fuente Marcos, "The invisible hand: star formation triggered by runaway black holes," *The Astrophysical Journal*, vol. 677, no. 1, pp. L47–L50, 2008.

[60] J. D. Schnittman and A. Buonanno, "The distribution of recoil velocities from merging black holes," *The Astrophysical Journal*, vol. 662, no. 2, pp. L63–L66, 2007.

[61] M. Kesden, U. Sperhake, and E. Berti, "Relativistic suppression of black hole recoils," *The Astrophysical Journal*, vol. 715, no. 2, pp. 1006–1011, 2010.

[62] M. Volonteri, M. Sikora, and J. P. Lasota, "Black hole spin and galactic morphology," *The Astrophysical Journal*, vol. 667, no. 2, pp. 704–713, 2007.

[63] P. A. G. Scheuer and R. Feiler, "The realignment of a black hole misaligned with its accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 282, no. 1, pp. 291–294, 1996.

[64] P. Natarajan and P. J. Armitage, "Warped discs and the directional stability of jets in active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 309, no. 4, pp. 961–968, 1999.

[65] T. Bogdanović, C. S. Reynolds, and M. C. Miller, "Alignment of the spins of supermassive black holes prior to coalescence," *The Astrophysical Journal*, vol. 661, no. 2, pp. L147–L150, 2007.

[66] A. Perego, M. Dotti, M. Colpi, and M. Volonteri, "Mass and spin co-evolution during the alignment of a black hole in a warped accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 4, pp. 2249–2263, 2009.

[67] M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruszkowski, and F. Haardt, "Dual black holes in merger remnants-II. Spin evolution and gravitational recoil," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 682–690, 2010.

[68] G. A. Shields, E. W. Bonning, and S. Salviander, "Comment on the black hole recoil candidate quasar SDSS J092-712.65+294344.0," *The Astrophysical Journal*, vol. 696, no. 2, pp. 1367–1373, 2009.

[69] T. M. Heckman, J. H. Krolik, S. M. Moran, J. Schnittman, and S. Gezari, "SDSSJ092712.65+294344.0: NGC 1275 at $z = 0.7$?" *The Astrophysical Journal*, vol. 695, no. 1, article 363, 2009.

[70] M. Dotti, C. Montuori, R. Decarli, M. Volonteri, M. Colpi, and F. Haardt, "SDSSJ092712.65+294344.0: a candidate massive black hole binary," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 1, pp. L73–L77, 2009.

[71] T. Bogdanović, M. Eracleous, and S. Sigurdsson, "SDSS J092-712.65+294344.0: recoiling black hole or a subparsec binary candidate?" *The Astrophysical Journal*, vol. 697, no. 1, pp. L288–L292, 2009.

[72] R. Decarli, M. T. Reynolds, and M. Dotti, "A photometric study of the field around the candidate recoiling/binary black hole SDSS J092712.65+294344.0," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 1, pp. 458–466, 2009.

[73] M. Vivek, R. Srianand, P. Noterdaeme, V. Mohan, and V. C. Kuriakose, "SDSS J092712.64+294344.0: recoiling blackhole or merging galaxies?" *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 1, pp. L6–L9, 2009.

[74] A. Robinson, S. Young, D. J. Axon, P. Kharb, and J. E. Smith, "Spectropolarimetric evidence for a kicked supermassive black hole in the quasar E1821+643," *The Astrophysical Journal*, vol. 717, no. 2, pp. L122–L126, 2010.

[75] G. A. Shields, D. J. Rosario, K. L. Smith et al., "The quasar SDSS J105041.35+345631.3: black hole recoil or extreme double-peaked emitter?" *The Astrophysical Journal*, vol. 707, no. 2, pp. 936–941, 2009.

[76] M. Elvis, in *Proceedings of the BAAS*, vol. 41, p. 708, 2009.

[77] J. M. Comerford, R. L. Griffith, B. F. Gerke et al., "1.75 h -1 kpc separation dual active galactic nuclei at z = 0.36 in the cosmos field," *The Astrophysical Journal*, vol. 702, no. 1, pp. L82–L86, 2009.

[78] F. Civano, M. Elvis, G. Lanzuisi et al., "A runaway black hole in cosmos: gravitational wave or slingshot recoil?" *The Astrophysical Journal*, vol. 717, no. 1, pp. 209–222, 2010.

[79] D. Batcheldor, A. Robinson, D. J. Axon, E. S. Perlman, and D. Merritt, "A displaced supermassive black hole in M87," *The Astrophysical Journal*, vol. 717, pp. L6–L10, 2010.

[80] P. G. Jonker, M. A. P. Torres, A. C. Fabian, M. Heida, G. Miniutti, and D. Pooley, "A bright off-nuclear X-ray source: a type IIn supernova, a bright ULX or a recoiling supermassive black hole in CXO J122518.6+144545," *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 1, pp. 645–650, 2010.

[81] J. M. Bellovary, F. Governato, T. R. Quinn, J. Wadsley, S. Shen, and M. Volonteri, "Wandering black holes in bright disk galaxy halos," *The Astrophysical Journal*, vol. 721, no. 2, pp. L148–L152, 2010.

[82] B. A. Keeney, J. T. Stocke, C. W. Danforth, and C. L. Carilli, "The quasar/galaxy pair PKS 1327-206/ESO 1327-2041: absorption associated with a recent galaxy merger," *Astronomical Journal*, vol. 141, no. 2, article 66, 2011.

[83] D. Sijacki, V. Springel, and M. G. Haehnelt, "Gravitational recoils of supermassive black holes in hydrodynamical simulations of gas-rich galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 4, pp. 3656–3670, 2011.

[84] G. Hasinger, "Absorption properties and evolution of active galactic nuclei," *Astronomy and Astrophysics*, vol. 490, no. 3, pp. 905–922, 2008.

[85] H. Hao, M. Elvis, F. Civano, and A. Lawrence, "Hot-dust-poor quasars in mid-infrared and optically selected samples," *The Astrophysical Journal*, vol. 733, no. 2, article 108, 2011.

[86] Z. Haiman, "Constraints from gravitational recoil on the growth of supermassive black holes at high redshift," *The Astrophysical Journal*, vol. 613, no. 1, pp. 36–40, 2004.

[87] M. Boylan-Kolchin, C. P. Ma, and E. Quataert, "Core formation in galactic nuclei due to recoiling black holes," *The Astrophysical Journal*, vol. 613, no. 1, pp. L37–L40, 2004.

[88] J. D. Schnittman and A. Buonanno, "The distribution of recoil velocities from merging black holes," *The Astrophysical Journal*, vol. 662, no. 2, pp. L63–L66, 2007.

[89] A. Sesana, "Extreme recoils: impact on the detection of gravitational waves from massive black hole binaries," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 1, pp. L6–L10, 2007.

[90] T. Tanaka and Z. Haiman, "The assembly of supermassive black holes at high redshifts," *The Astrophysical Journal*, vol. 696, no. 2, pp. 1798–1822, 2009.

[91] M. Volonteri, K. Gültekin, and M. Dotti, "Gravitational recoil: effects on massive black hole occupation fraction over cosmic time," *Monthly Notices of the Royal Astronomical Society*, vol. 404, no. 4, pp. 2143–2150, 2010.

[92] M. Micic, K. Holley-Bockelmann, and S. Sigurdsson, "Growing massive black holes in a Local Group environment: the central supermassive, slowly sinking and ejected populations," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 2, pp. 1127–1144, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 805402, 16 pages
doi:10.1155/2012/805402

Review Article

A Practical Guide to the Massive Black Hole Cosmic History

A. Sesana

Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institut, Am Mülenber 1, 14476 Golm, Germany

Correspondence should be addressed to A. Sesana, alberto.sesana@aei.mpg.de

Received 1 September 2011; Accepted 26 October 2011

Academic Editor: Francesco Shankar

I review our current understanding of massive black hole (MBH) formation and evolution along the cosmic history. After a brief introductory overview of the relevance of MBHs in the hierarchical structure formation paradigm, I discuss the main viable channels for seed BH formation at high redshift and for their subsequent mass growth and spin evolution. The emerging hierarchical picture, where MBHs grow through merger triggered accretion episodes, acquiring their mass while shining as quasars, is overall robust, but too simplistic to explain the diversity observed in MBH phenomenology. I briefly discuss which future observations will help to shed light on the MBH cosmic history in the near future, paying particular attention to the upcoming gravitational wave window.

1. Introduction

There is nowadays a general consensus that massive black holes (MBHs) are fundamental pieces of the puzzle of the galaxy evolution process in the Universe. 10^8–$10^9\,M_\odot$ MBHs have long been recognized to be the central engines of powerful quasars, whose extreme luminosity ($>10^{45}\,\mathrm{erg\,s^{-1}}$) emitted from a tiny volume ($<10^{-3}$ pc, as inferred by variability studies) can be explained only through radiatively efficient accretion of matter onto supermassive compact objects. Observations of quasars out to redshift $\lesssim 7$ [1–3] imply that some of these objects were already in place when the Universe was less than a billion year old, although the bulk of quasar population in the Universe comes later, at $z \approx 2$ [4]. In the last twenty years, evidence for the existence of MBHs has also been found in nearby quiescent galaxies. The most compelling evidence involves our Milky Way, where accurate dynamical measurements of stellar orbits have revealed a $4 \times 10^6\,M_\odot$ MBH [5, 6]. Thanks to reverberation mapping [7] as well as stellar and gas kinematics measurements, the masses of $\sim$100 MBHs in nearby galaxies have been measured (see, e.g., [8] and references therein), providing compelling evidence that MBHs are ubiquitous in the center of massive nearby galaxies [9]. Their masses have been found to correlate with several properties of the hosting galaxy, in particular, with the bulge mass, luminosity, and velocity dispersion [8, 10–16], and probably with the mass of the dark matter halo ([17], although such relation is debated, see [18]), indicating an intimate connection linking the MBHs to their hosts. Such intimate connection is supported by the strong correlation between the redshift evolution of the global star formation rate and the luminosity density of optically selected quasar [19, 20]. Moreover, the dustiest starburst galaxies (high redshift submillimeter galaxies and ultraluminous infrared galaxies, ULIRGs) show sign of a recent merger and are often associated with quasar activity [21–23].

Putting the pieces together, a well-defined picture has emerged, in which the dormant MBHs that populate galaxies today are the relics of the luminous quasars that shone in the past. In the widely accepted ΛCDM cosmology, structure formation proceeds in a hierarchical fashion [24], in which massive galaxies are the result of several merging and accretion events involving smaller building blocks. In this framework, the MBHs we see in todays galaxies are expected to be the natural end-product of a complex evolutionary path, in which black holes seeded in protogalaxies at high redshift grow through cosmic history via a sequence of MBH-MBH mergers and accretion episodes [25–27]. In this general picture, galaxy mergers trigger copious infall of cold gas in the nucleus of the merger remnant [28], resulting in vigorous star formation and accretion on the central MBH. The energy output of the accreting MBH acts as a feedback on the surrounding environment either by removing [29] or heating [30] the cold gas, regulating the star formation process,

and self-regulating its own accretion [31], possibly resulting in the observed MBH-host relations [32, 33]. Hierarchical models for MBH evolution, associating quasar activity to gas-fueled accretion following galaxy mergers, have been successful in reproducing several properties of the observed Universe, such as the present day mass density of nuclear MBHs and the optical and X-ray luminosity functions of quasars [25, 26, 30, 34–37]. Though generally successful, most of the key ingredients of the picture are poorly understood. It is still unclear how and when the first seed BHs at high redshift form, and which accretion channel may possibly lead to $>10^9\,M_\odot$ MBHs at $z \approx 7$. The MBH-environment interplay establishing the observed MBH-host relation is poorly known, and beyond the resolution of state-of-the-art simulations. Moreover, especially at the low mass end, most of the MBH accretion seems to occur in isolated systems, indicating that the merger-accretion paradigm is not the whole story.

A complete understanding of the formation and evolution of MBHs is one of the most exciting goals of contemporary astrophysics and cosmology. Here, we review the key ideas composing this fascinating puzzle; the first seed formation and their subsequent mass growth, the galaxy merging process leading to the formation of MBH binaries, the possible paths for the evolution of MBH spins. We caution that given the vastness of the covered topics, this review is necessarily incomplete. As a matter of fact, each of the following sections would be worthy several reviews in itself. We try to give a general overview, pointing to the relevant literature for in-depth examination by the reader, with the goal of designing a broad background where other specific topics tackled in this volume can be framed.

The review is organized as follows. In Section 2, we inspect the main routes for seed BH formation at high redshift in the frame of ΛCDM cosmology. The subsequent accretion and merging history is considered in Section 3, where we describe different possible MBH mass growth paths. Section 4 is devoted to spin evolution and to its relation to the MBH merging history and host galaxy properties. We discuss which future observations will help to shed light on the MBH cosmic history in the near future, paying particular attention to the upcoming gravitational wave window, in Section 5, and we briefly summarize in Section 6.

2. The Infancy of the Giants: Seed Black Hole Formation

The first obvious relevant question is where and when the seeds of the MBHs powering quasars at redshifts as high as 7, as well as lurking in the center of inactive galaxies today form. In this section, we describe possible paths of seed BH formation in the currently favored general framework of ΛCDM cosmology. In the standard ΛCDM picture, the mass content of the Universe is largely dominated by cold dark matter (DM), with baryons contributing to a 10% level only. Starting from a Gaussian density fluctuation field in a quasi-homogeneous Universe, DM perturbations grow in time, to the point they decouple from the Hubble flow and collapse and virialize forming self-gravitating DM halos (see [38, 39] for a comprehensive treatment). This is a bottom up path, in the sense that small halos collapse first and get bigger and bigger by accreting from ambient DM and merging with other halos, in a process known as hierarchical clustering. Press and Schechter [40] provided a statistical description of the halo growth in terms of σ overdensities of a Gaussian density field. Figure 1, adapted from [41] by Barkana and Loeb, illustrates this concept. At any given time (redshift) 1σ fluctuations of the density field correspond to the characteristic mass of a collapsing halo at that redshift. More massive halos correspond to rarer peaks (i.e., 2-3σ) of the density field; such halos are much rarer, but they constitute, at each redshift, the densest environments where the formation of the first collapsed objects took place [42, 43].

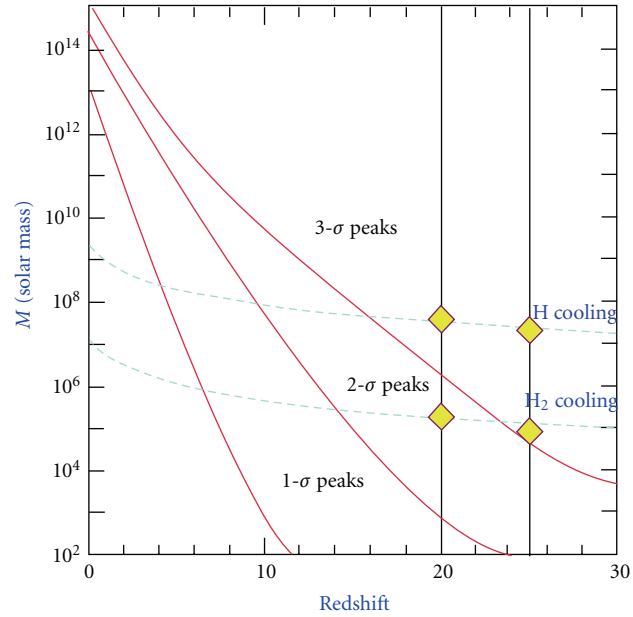


Figure 1: First DM halos allowing for efficient gas cooling. Red lines show the mass of different σ peaks of the halo mass function as a function of redshift; dashed blue lines show the minimum mass required for efficient H_2 (lower) and H (upper) cooling. Diamonds at $z = 20$ and $z = 25$ mark halo masses needed for efficient H and H_2 cooling at those redshifts. Adapted by Volonteri from [41].

In the early halo assembly process, baryons follow the DM halo potential well, without plying any relevant dynamical role. The first halos suitable to seed BH formation are those exceeding the Jeans limit, $M_J \approx 10^4\,M_\odot[(1+z)/10]^{3/2}$, that is, the mass needed for the self-gravity of the halo to overcome thermal gas pressure, allowing the gas to contract together with the halo and to shock heat at the halo virial temperature. Note that M_J is an increasing function of z, whereas the characteristic halo mass increases with decreasing z. Baryons start to efficiently virialize in rare DM halos of $\sim 10^5\,M_\odot$ at $z \approx 50$. The formation of collapsed objects (stars, BHs, etc.) within this halos depends on the ability of the gas to cool efficiently. The primordial zero metallicity gas, is mostly composed of atomic and molecular hydrogen [44], whose efficient cooling is possible at temperatures higher than 10^4 K and 300 K, respectively.

At any given z, the characteristic halo virial temperature of a DM halo is given by $T_v \approx 2 \times 10^4\,\mathrm{K}\, M_{h,8}^{2/3}(1+z)/10$, where $M_{h,8}$ is the halo mass in units of $10^8\,\mathrm{M}_\odot$ [41]. Efficient cooling is therefore possible in rare, high σ fluctuations of $\sim 10^6\,\mathrm{M}_\odot$ at $z \approx 25$ for molecular hydrogen, or of $\sim 10^8\,\mathrm{M}_\odot$ at $z \approx 15$ for atomic hydrogen [45], as indicated by the intersection of the blue-dotted and the upper red-solid curves in Figure 1. The precise outcome of such cooling process depends on largely unknown details of the involved micro- and macrophysics, as well as on the evolution of the environment; in the following, we spell out the three main seed BH formation mechanisms that have been proposed, the resulting seed mass functions are plotted in Figure 2.

2.1. Population III Star Remnants. After recombination, a small amount of hydrogen combines in H_2 molecules [47]. In absence of any primordial soft UV background, H_2 is not photodissociated and acts as an effective coolant in $\sim 10^6\,\mathrm{M}_\odot$ halos at $z \approx 25$ with $T_v \gtrsim 300\,\mathrm{K}$ [45]. In absence of metals, H_2 driven contraction is subsonic, preventing the cloud to fragment (differently to present day metal enriched star forming clouds). Several simulations of the primordial collapse of these metal-free clouds suggest that the first generation of stars (the so called PopIII stars) was massive $m_* > 100\,\mathrm{M}_\odot$ ([48–51], but see discussion below). The final fate of such massive low metallicity stars has been extensively studied by Heger and collaborators [52]; if $m_* > 260\,\mathrm{M}_\odot$, after only $\sim 2\,\mathrm{Myr}$, the star directly collapses into a BH of half its initial mass, leaving behind a few hundred solar mass remnant that can possibly play the role of a seed BH [53].

This very natural scenario presents, however, a lot of uncertainties, mostly related to the mass of the final stars. Early simulations just showed a $\sim 10^3\,\mathrm{M}_\odot$ gas overdensity surrounding a $0.01\,\mathrm{M}_\odot$ optically thick core. The mass is then accreted on the stellar core, but the accretion rate and the final mass of the stars are likely affected by feedback effects [54], possibly resulting in lighter stars. Moreover, recent higher-resolution simulations found a significant degree of fragmentation of the collapsing clump [55, 56]. There are indications that PopIII stars usually form in binary reach clusters with a much less top heavy mass function as previously thought. Typical stellar masses cover a wide 0.1–$50\,\mathrm{M}_\odot$ range [57, 58], well below the $260\,\mathrm{M}_\odot$ threshold required for efficient seed BH formation. In principle, a compact cluster of massive stars can still undergo runaway collision, forming a massive single star [59], eventually collapsing in a seed BH, a scenario we will discuss later in Section 2.3. In any case, these recent results are seriously challenging the viability of PopIII remnants as seed BH candidates.

2.2. Direct Collapse. Another route leading to a seed BH involves the formation and subsequent direct collapse of a very massive star or disk of gas at high redshift. Several variants of this process have been studied in the literature [60–67]; the advantage with respect to the PopIII scenario is the formation of fairly massive seeds (usually $\sim 10^5\,\mathrm{M}_\odot$), which can more easily grow to $10^9\,\mathrm{M}_\odot$ to power the observed $z > 6$ quasars.

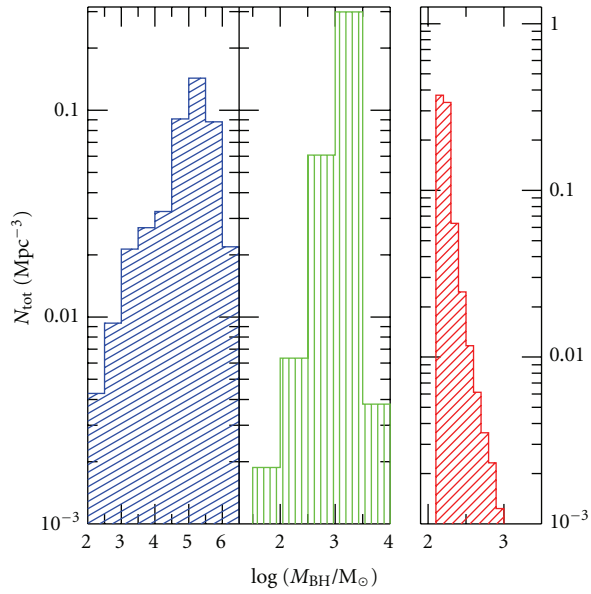


Figure 2: Mass function of seed BHs for the three different formation scenarios discussed in the text: direct collapse (left panel), runaway mergers in high redshift star clusters (central panel), and PopIII remnant (right panel). From [46].

It has been suggested that direct collapse is efficient only in metal-free halos with $T_v \gtrsim 10^4\,\mathrm{K}(M_h > 10^8\,\mathrm{M}_\odot)$, where H_2 cooling is suppressed. Under such condition, the gas cloud collapses isothermally at $\sim 8000\,\mathrm{K}$ set by atomic line cooling. The work in [63] carried numerical SPH simulations of this scenario, finding a central gas condensation of $>10^6\,\mathrm{M}_\odot$ in a radius of $<1\,\mathrm{pc}$, without any signal of fragmentation. However, the work in [68] showed that H_2 suppression is rather inefficient in such massive halos at this high z, and it is likely that H_2 also acted as efficient coolant. Moreover, the gas cloud has a significant angular momentum related to its halo spin [69]. For typical halo parameters, the gas forms a rotationally supported disk on 10 pc scales [46], and additional angular momentum transport is required to reach the MBH formation conditions.

Dynamical instabilities, either global or local, provide a viable solution to the angular momentum transport problem. Centrifugally supported clouds become bar unstable if the level of rotation is higher than a given threshold. Bars are efficient in transporting angular momentum outwards creating global inflows of matter. If the gas can cool efficiently, the instability grows on smaller scales, creating a cascade process, called "bars within bars" instability [70]. Numerical simulations of the process show that it allows accumulation of $>10^6\,\mathrm{M}_\odot$ of gas at subparsec scales, and it works also in presence of star formation. An alternative way to transport angular momentum is via local instabilities [66]. By acquiring mass from the ambient gas, the disk eventually feels its self-gravity, becoming marginally stable. Disk stability can be described via the Toomre parameter $Q = c_s k/\pi G\Sigma$, where c_s is the sound speed, k the epicyclic frequency, and Σ the disk surface density. A marginally stable

disk has $Q \sim 1$. In this situation, rather than fragmenting, the disk develops spiral structures that efficiently transport angular momentum outwards, causing mass inflows in the center [71]. Such mass inflows have a double self-regulating effect: the disk surface density gets lower, and the epicyclic frequency of the disk increases, preventing the Q parameter from becoming too small resulting in disk fragmentation. The work in [66] showed that also in this way $\sim 10^6\,M_\odot$ of gas can be accumulated in the center.

The accumulated gas will eventually form a massive seed BH. If the mass accumulation timescale is fast enough, a supermassive star (SMS, [72]) will form in the center. Fully relativistic simulations of isolated rotating SMSs showed that they eventually collapse in a Kerr MBH with negligible mass loss [73–75]. However, the mass is usually accreted onto the SMS core gradually, and the isolated approximation may be too simplistic. Begelman and collaborators [65] studied in details the SMS formation and evolution. They found that, as accretion proceeds, the central optically thick core collapses into a few $M_\odot$ BH, accreting efficiently at the Eddington limit of the whole external envelope (called quasistar). The allowed accretion rate therefore greatly exceeds the Eddington rate for the BH only, resulting in its rapid growth. The resulting BH mass is in the $>10^4$–$10^5\,M_\odot$ range. However, recent calculations [67] taking into account for strong wind-driven mass loss of the quasistar show that the final mass of the seed BH can be considerably smaller compared to the quasistar mass. In this case, central gas accumulations $>10^7\,M_\odot$ are required in order to leave a final seed with mass $>10^4\,M_\odot$, setting a minimal halo mass for efficient seed BH formation of $M_h > 10^9\,M_\odot$. In such scenario, seed BHs would be rarer and would form at slightly lower redshift ($z \approx 12$).

2.3. Runaway Stellar Dynamics. BHs in the 10^2–$10^4\,M_\odot$ range can be the endproduct of runaway collisions in dense stellar clusters [59, 76–79]. The runaway process has been studied in detail in the context of intermediate MBH formation in globular and massive star clusters. Dense stellar clusters have negative heat capacity, implying that they tend to get denser in the center (a phenomenon known as core collapse [80, 81]) while the less bound stars get unbound. Core collapse usually occurs on a relaxation timescale $t_{\rm relax}$ but is much faster if the mass spectrum of the involved stars covers a broad range. In this case, the most massive stars segregate to the center in a timescale $t_{cc} \ll t_{\rm relax}$ (t_{cc} stands for core collapse time). If $t_{cc} < 3$ Myr (the lifetime of the most massive stars), segregated stars undergo runaway collisions forming a very massive star (VMS, [82]). This usually happens in clusters more massive than few $\times 10^5\,M_\odot$ with half mass radius <1 pc. Devecchi and Volonteri [83] showed that runaway collapse is a viable route for seed BH formation in mildly metal polluted ($Z < 10^{-3} Z_\odot$), efficient hydrogen cooling ($T_v > 10^4$ K) halos at high redshift. Their scenario is similar to the direct collapse mechanism discussed above, with the difference that the mass accumulated in the central parsec overcomes the density threshold (function of the metallicity, see [84]) at which fragmentation and star formation occur efficiently. The most massive stars in the resulting compact cluster undergo runaway collisions forming a VMS of several thousand solar masses, leaving behind a seed BH remnant of $\sim$1000–2000 $M_\odot$. As already mentioned, a similar process may operate at zero metallicity even during the first episode of star formation: if PopIII stars form in clusters, following fragmentation of the collapsing cloud, runaway collisions can in principle still lead to the formation of a $\sim 1000\,M_\odot$ seed BH. Davies and collaborators [85] recently proposed a variant of this scenario, in which the requirement $t_{cc} < 3$ Myr is relaxed. If a dense star cluster is formed, subsequent massive gas inflows following frequent halo mergers at high redshift cause significant contraction of the cluster core, which undergoes a global collapse. Stellar remnants merge together leaving behind a large seed BH of $\sim 10^5\,M_\odot$, comparable to direct collapse endproducts.

3. The Making of the Giants: Massive Black Hole Accretion History

Once they are formed, seed BHs will inevitably interact with their environment. Nearby elliptical galaxies host MBHs of $>10^9\,M_\odot$, and observations of distant quasars imply that such monsters (at least few of them) were already in place at redshift $\lesssim 7$. Therefore, seed BHs need to accrete an enormous amount of mass, and they need to do it fast. We can identify three principal mass growth mechanisms:

(i) merger with other MBHs;

(ii) episodic accretion of compact objects, disrupted stars, or gas clouds;

(iii) prolonged accretion of large supplies of gas via accretion disks.

Soltan [86] first noticed that the optical luminosity function of quasars directly implies a large population of nuclear MBHs lurking in quiescent galaxies today. Infact, to an observed luminosity L corresponds a mass accretion rate $\dot{M} = f_{\rm bol}(1-\epsilon)L/(\epsilon c^2)$, where $f_{\rm bol}$ is a bolometric correction to the luminosity and ϵ is a mass-to-energy conversion efficiency. Subsequent mass measurements of the local nuclear MBHs suggested that the MBH mass density in the local Universe is consistent with the accreted mass inferred by integrating the quasar luminosity function at all redshifts assuming $\epsilon \approx 0.1$ [20, 87–90]. Therefore, the "quasar mode," in which large amount of gas are accreted in single coherent episodes via accretion disks shining close to the Eddington luminosity, appears to be principal path of MBH cosmic growth. (We recall that the Eddington luminosity is the maximum admitted luminosity for which the radiation pressure exerted by the photons emitted in the accretion process is smaller than the gravitational binding energy of the accreting material. If the contrary is true, radiation pressure blows away the reservoir of gas, suppressing the accretion process. For standard radiatively efficient accretion flows [91], the Eddington luminosity corresponds to an accretion rate of $\dot{M}_{\rm Edd} \approx 2.5 M_8\,M_\odot\,{\rm yr}^{-1}$, where M_8 is the MBH mass normalized to $10^8\,M_\odot$.) We notice, however, that there are caveats to this argument. If $z \approx 2$ quasars have $\epsilon \approx 0.2$ (which may be the case if they have, on average, a substantial

spin), then the integrated MBH mass density implied by the quasar luminosity function would be much lower, leaving room to a significant contribution from obscured accretion. In fact, hard X-ray source counts suggest that most of the MBH growth occurs in highly obscured objects ([92, 93], see Treister and Urry contribution to this volume [94]). A significant "obscured" growth is still consistent with the Soltan argument (which is usually applied to the quasar optical luminosity function) if the typical ϵ is high. We notice, moreover, that the local MBH mass density has in itself a factor of two uncertainty, readily allowing equal amount of obscured and unobscured accretion even assuming $\epsilon \approx 0.1$.

The Soltan argument is consistent with the general picture in which galaxy mergers trigger gas inflow in the nuclear region, feeding quasar activity [25, 27, 95]. A $10^9\,M_\odot$ MBH is thought to have undergone several of such mergers in its lifetime, acquiring most of its mass through relatively short ($\sim 10^7$ yrs) Eddington limited accretion episodes. Although this general picture has proved successful in reproducing the properties of the observed quasar population (see, e.g., [26, 36]), the detailed accretion history of MBHs is far more complex. Firstly, it is hard to grow $10^9\,M_\odot$ MBHs at redshift ≈ 7, when the Universe was ~ 0.8 Gyr old; it is therefore likely that early accretion at high redshift proceeded at super-Eddington rates (see next section). Secondly, although the merger-driven paradigm seems to apply to bright quasars, the same is not true for fainter, less massive, accreting MBHs. Most of those are hosted in spiral galaxies that likely did not undergo any major merger in their lifetime. There is a wide variety of accretion modes, likely related to the environmental conditions surrounding the MBHs, and all of them have to be explored in order to get a comprehensive picture of the MBH accretion history.

3.1. The Highest Redshift Quasars. One of the main challenges of hierarchical formation models is to address the presence of $>10^9\,M_\odot$ MBHs at $z \lesssim 7$, when the Universe was less than 0.8 Gyr old. Note that these systems have a space density of $1\,\mathrm{Gpc}^{-3}$ comoving volume [2], corresponding to 5σ fluctuations of the original density field [96]. This means that they are not representative of the typical MBH cosmic evolution; they are extremely rare and grew in the richest overdensities of the early Universe. To grow to $>10^9\,M_\odot$ at $z \lesssim 7$, a putative seed BH would require almost continuous Eddington limited accretion over its whole lifetime. The mass accretion depends both on environmental factors (the supply of available gas) and the detailed accretion process (e.g., the effects of accretion feedback). The environment does not seem to be a problem for these objects. Numerical simulations show that the nuclear regions of extreme high-z overdensities are constantly fed by filamentary cold gas flows [97, 98]. If efficiently accreted, the MBH can grow to $>10^9\,M_\odot$ without even experiencing a merger event with another protogalaxy hosting another MBH. Therefore, the merger triggered quasar accretion scheme seems to break down in such extremely dense environments.

Whether the mass supplied through the filaments can be efficiently accreted is a delicate issue. Eddington limited accretion for $<10^9$ yrs can produce an MBH of $>10^9\,M_\odot$ only if $\epsilon \lesssim 0.1$, that is, if the MBH spin remains small [99, 100]. Infact, writing the mass growth as $M(t) = M(0)\exp(((1-\epsilon)/\epsilon)(t/t_{\mathrm{Edd}}))$, where $t_{\mathrm{Edd}} = 0.45$ Gyr, is easy to see that it needs ≈ 0.5 Gyr to grow a $10^9\,M_\odot$ MBH out of a PopIII seed if $\epsilon = 0.06$ (accretion onto a Schwarschild BH), but it takes ≈ 5 Gyr if $\epsilon = 0.4$ (prograde accretion onto an highly spinning BH). However, continuous coherent accretion through a standard accretion disk [91] is expected to efficiently spin up the MBH [101, 102], increasing ϵ to >0.3, hence considerably slowing down the MBH growth. To solve this problem, King and collaborators [103] proposed that mass is accreted in a series of small incoherent packets (chaotic accretion). In this case, depending on the angular momentum of the accreted material, the MBH is spun up or down, performing a random walk in spin magnitude that keeps it close to zero [104]. Although such solution is appealing (and probably relevant in different contexts, see Section 3.3), a continuous supply of gas powering a quasar for such long time ($\sim 10^9$ yr) is likely to be coherent, settling in a well-defined accretion disk, implying efficient spinup.

It has been suggested that MBH growth at high redshifts proceeds through radiatively inefficient accretion flows. If the mass supply rate is much larger than the Eddington rate ($\dot{m} = \dot{M}/\dot{M}_{\mathrm{Edd}} \gg 1$, which can be easily sustained by cold inflows at high z), photons are trapped in the accretion flows because the time it takes them to diffuse out of the flow is longer than the accretion time [105, 106]. Consequently, the emitted luminosity is suppressed to the point that $\dot{m}$ can greatly exceed 1 (supercritical accretion), still resulting in a sub-Eddington luminosity. Note that early supercritical accretion is still consistent with the Soltan [86] argument: MBHs can in principle grow this way to 10^7–$10^8\,M_\odot$, complete the last couple of e-folding in mass via radiatively efficient accretion, and still satisfy the quasar luminosity function-mass conversion constraints. However, simulations of radiative inefficient accretion flows [107–111] showed that most of the infalling matter is driven away by wind-like outflows, and the actual accretion rate can exceed the Eddington limit by a factor of ≈ 10 only. Supercritical accretion is also allowed in quasispherical Bondi-like accretion flows [105]. Volonteri and Rees [112] discussed how a short phase of supercritical accretion can easily produce $>10^9\,M_\odot$ MBH at $z > 6$ out of PopIII remnant seeds, however, how and when this happens in practice is an open question.

3.2. The Standard Paradigm of MBH Evolution. Most of the MBH mass density of the Universe is built up during the peak of quasar activity at $z \approx 2$. At those low redshifts, MBHs had all the time to grow in mass through a sequence of merger-triggered, Eddington-limited accretion events, without invoking continuous high redshift cold gas inflows or supercritical accretion. As already discussed, standard hierarchical MBH formation models have been proved successful in producing the population of MBHs powering medium redshift quasars, and their luminosity function.

One of the main features of the hierarchical assembly is the formation of a large number of MBH binaries along the cosmic history. A detailed description of their dynamical formation and evolution can be found in Dotti et al.

contribution to this volume [113]. In few words, following galaxy mergers, MBHs sink to the center of the remnant owe to dynamical friction against the DM/stellar/gaseous background [114, 115]. At subparsec scales, the mass enclosed within the relative orbits of the two MBHs becomes less then their own mass; the MBHs feel each other pull forming a bound binary. The binary fate is defined by its interaction with the stellar and gaseous environment. In gas-rich nuclei, torques exerted by a circumbinary disk can effectively extract the MBH binary energy and angular momentum, possibly resulting in a fast ($\sim 10^7$ yr) shrink to milliparsec separations [116, 117], where gravitational wave (GW) emission efficiently drives its coalescence [118]. Also in gas-poor environments, recent simulations [119–122] showed that three body interactions with ambient stars efficiently fed (by rotation and triaxiality of the stellar distribution) into the binary loss cone can bring the system to final coalescence in $\sim 10^8$ years. Along with the binary formation, during the merger, the cold gas content of the interacting system is highly destabilized, triggering inflows in the nuclear region [28, 123] that provide a large reservoir of fuel for the active phase of the MBHs. Concurrently, the dense nuclear cold gas triggers an efficient star formation episode [27], which competes with the MBH gas fueling depleting the reservoir of available gas. MBH feedback onto the surrounding environment regulates the MBH growth, shaping the MBH-bulge relations observed in the local Universe. This general scenario naturally reproduces the observed "downsizing" [124–126], which was initially thought to be in contrast with the hierarchical formation process. The downsizing is the phenomenon whereby luminous activity appears to occur in progressively lower mass objects as the redshift decreases; that is, the most massive MBHs seem to accumulate their mass before the lighter ones [127]. As noted by Malbon and collaborators [36], downsizing is a natural outcome of hierarchical structure formation. Even without invoking complex feedback mechanisms, the most massive black holes are those that grew in the densest environments, where dynamical evolution occurs on shorter timescales. Such MBHs are likely to experience several early mergers that led to an early exhaustion of their cold gas reservoir (because of accretion, star formation and feedback), suppressing their luminous activity at lower redshift. Conversely, less massive objects are hosted in less dense environments and experience a smoother evolution: at low redshift is still plenty of cold gas for them to shine as luminous quasars, creating the "downsizing" effect.

There is plenty of compelling evidence supporting, to some level, this scenario. The observed star formation history and the quasar activity in the Universe mimic each other peaking at $z \approx 2$ [19, 20]; many luminous quasars are found in "disturbed" galaxies, where the presence of tails or clumps in the matter distribution indicates a recent merger with another object (see, e.g., [128, 129]); dual quasars have been found in interacting galaxy pairs with projected separation <50 kpc (NGC 6240 [130] being the prototypical case); most ULIRGs (i.e., highly star forming galaxies) are associated to merging systems and show active galactic nuclei (AGN) in their center [21]. Nevertheless, the overall picture suffers of many uncertainties. Firstly, galaxy mergers generally trigger AGN activity, but the contrary is not true; most of the AGN (especially in the local universe) occurs in "unperturbed," isolated galaxies (see next Section). Secondly, the relation between the AGN and star formation triggering timescale and the MBH binary evolution timescale is poorly constrained. The presence of dual quasars indicates that at least in some cases accretion is triggered well before merger, however, that might not be the general case. Finally, evidence for bound MBH binaries remains circumstantial [131–135], and most of the proposed candidates have different explanations (see Dotti et al. contribution to this volume for a detailed discussion [113]).

3.3. Secular Evolution at the Small End of the MBH Mass Function. There is growing observational evidence that the merger-driven accretion paradigm is not the whole story. In the clustering scenario, elliptical galaxies form as a consequence of galaxy mergers [31], and we can infer that most of the residing MBH mass has been acquired by merger-triggered accretion episodes (consistent with the cosmic population of bright quasars). However, present day spirals (which dominate the population of low mass galaxies [136]) have likely experienced a much quieter cosmic evolution, possibly without undergoing any major merger event. Still, many of them show significant nuclear activity, shining as Seyfert galaxies [137]. Seyfert galaxies are in general characterized by lumpy structures rich of cold gas [138], and most of them show dense nuclear stellar clusters with densities of $\sim 10^6\,M_\odot/pc^3$ (an environment similar to the Milky Way center [139]). In such a rich environment, the main MBH growth channel is thought to be the accretion of small packets of material reaching the galactic nucleus because of dynamical relaxation and secular evolution processes.

In dense stellar environments, tidal disruptions of main sequence stars scattered in the MBH loss cone by two body relaxation are quite common. During a disruption event, half of the debris is accreted by the MBH, powering a luminous episode which can last for years, with a characteristic power law decline $L \propto t^{-5/3}$ [140]. Theoretical studies predict rates in the range 10^{-5}–few $\times\, 10^{-4}\,yr^{-1}$ in relaxed dense nuclei hosting a 10^6–$10^7\,M_\odot$ MBH [141]. Moreover, such rates can be highly increased for a short time (~Myr) by the presence of a secondary inspiralling MBH as a consequence of a minor merger [142–144] or a star cluster accretion event [145]. To date, a total of ~15 tidal disruption flare candidates has been identified; most of them in X-ray and UV [146–150], and two in optical [151]. The soft X emission is generally found to be consistent with thermal emission from a $T \sim 10^5$ K accretion disk, in line with expectations for a 10^6–$10^7\,M_\odot$ MBH [152, 153]. The inferred rates are broadly consistent with theoretical predictions, implying that this may indeed be a major mass growth channel for MBHs with $M <$ few $10^6\ M_\odot$ [154]. The presence of segregated nuclear clusters of compact objects [155, 156] is also expected to result in a significant number of "dark" accretion events (extreme mass ratio inspirals, EMRIs). Even though the event rates are likely too small to significantly contribute to the MBH mass buildup, such events may offer a unique way to probe MBHs in quiescent nuclei through GW observations (see Section 5).

The principal growth channel of MBHs in spirals, however, is likely to be the random accretion of molecular clouds [157]. Observations suggest that typical accretion episodes in Seyferts last 10^4–10^5 yr, whereas the total active lifetime (based on the fraction of disk galaxies that are Seyferts) is in the range 10^8–10^9 yr [137]. Disk galaxies show a rich structure of nuclear molecular clouds down to parsec scales (see, as a reference, the MW [158, 159]), with masses in the range 10^3–$10^5\,M_\odot$. Relaxation processes will likely drive some of these clouds toward the nucleus, fueling the dormant MBH. This feeding channel is likely unimportant in giant ellipticals, where the nuclear densities are much lower, the bulge less lumpy, and the content of cold gas very small [160] as a consequence of AGN feedback and stellar wind loss related to the last accretion episode (see, e.g., [161]), resulting in a pronounced lack of molecular clouds.

4. Spin Evolution

Astrophysical black holes are fully described by two quantities only: mass and spin. The latter can be expressed by the dimensionless parameter $\hat{a} = J/J_{\max} = cJ/GM_{\rm BH}^2$. By definition, $0 \le \hat{a} \le 1$. At present, we have several effective ways to measure MBH masses, and we can construct MBH mass functions and study their evolution with redshift; conversely, spin measurements are way more problematic, and only a handful of estimates are available at the time of writing. The principal measurement technique is through fitting the skewed relativistic Kα fluorescent line [162], whose shape is highly spin-dependent. There is just a handful of objects with reliable spin measurements. Notable examples are Fairall 9, $\hat{a} = 0.60 \pm 0.07$ [163], and $\hat{a} = 0.67^{+0.11}_{-0.10}$ [164]; 1H 0707-495, $\hat{a} > 0.93$ [165]; MRK509, $\hat{a} = 0.78^{+0.04}_{-0.03}$ [165]; MRK 79, $\hat{a} = 0.7 \pm 0.1$ [166]. However, several free parameters enter in the fitting procedure, creating a severe degeneracy problem. The measurement is particularly sensitive to the modeling of the soft X excess of the continuum, which blends with the low-energy tail of the line (which is crucial for the spin estimation). Different fits to the Kα line in MCG-6-30-15 gave $\hat{a} > 0.98$ [167], $\hat{a} = 0.86 \pm 0.01$ [165], and $\hat{a} = 0.49^{+0.20}_{-0.12}$ [164]. Even more problematic is the case of NGC 3783, for which different groups found $\hat{a} > 0.88$ [168] and $\hat{a} < 0.32$ [164].

Measuring and understanding spins is crucial to assess the MBH cosmic evolution. Firstly, spins affect the accretion-luminosity conversion efficiency; highly spinning BHs can convert up to ~40% of the accreted matter in radiation, growing much slowly. Secondly, the "spin paradigm" for MBH jets [169] assumes that radio jets observed in AGNs are launched by highly spinning BHs [170]. Lastly, spins dramatically affect the gravitational recoil suffered by the remnant MBH after a binary merger. It has infact been shown that highly spinning BHs can experience kicks up to 5000 km s^{-1} depending on their progenitor spin magnitude and orientation [171–173]. These "superkicks" are sufficient to eject the remnant from the deepest potential well of the most massive galaxy clusters, with potentially important implications for the hierarchical MBH formation paradigm, and for the occupation fraction of MBHs in galaxies.

4.1. Black Hole Spin Evolution Channels. MBH spins can significantly evolve via two major channels: accretion of matter and coalescences with other MBHs. Different evolution paths lead to different spin distributions, as it has been extensively discussed by Berti and Volonteri [174].

4.1.1. Coalescences. MBH binary coalescence is an unavoidable key element in the hierarchical formation scenario. The binary system is characterized by an orbital angular momentum and two spin vectors, whose sum (plus the angular momentum radiated in GWs) has to be conserved during the coalescence. Under this basic assumption, Hughes and Blandford [175] found that, in general, coalescences of MBHs with random spin directions result in a broad remnant spin distribution; in particular, highly spinning MBHs tend to spin down. With the advent of numerical relativity [176–178], it is now possible to construct accurate mappings of initial-to-final BH spins following purely GW-driven coalescences [179]. In general, unless the two initial MBH spins and the binary orbital angular momentum are aligned, coalescences are unlikely to result in highly spinning MBHs.

4.1.2. Accretion. The accretion imprint on the MBH spin crucially depends on the characteristics of the accretion flow. Firstly, let us consider an MBH accreting from an extended planar accretion disk with $M_{\rm disk} \gg M_{\rm BH}$ and $L_{\rm disk} \gg S_{\rm BH}$, where $L_{\rm disk}$ is the disk angular momentum and $S_{\rm BH}$ is the MBH spin. The natural consequence of such accretion mode is MBH spinup. Given an initial mass M_0, coherent accretion of a $\sqrt{6}M_0$ mass of gas will turn a nonrotating MBH into maximally spinning [102]. If $L_{\rm disk}$ and $S_{\rm BH}$ are misaligned, the Bardeen-Petterson effect [180] will act to align the MBH spin to the disk angular momentum in a very short timescale ($t_{\rm align} \ll t_{\rm acc}$ [181, 182]), leaving the picture unchanged. However, things can be significantly different if $L_{\rm disk} \lesssim S_{\rm BH}$. Defining θ to be the angle between $L_{\rm disk}$ and $S_{\rm BH}$, King and collaborators [103] found that if $\cos\theta > -L_{\rm disk}/2S_{\rm BH}$, the accreted material counteraligns with the MBH. For $L_{\rm disk} \ll S_{\rm BH}$, there is equal chance of prograde and retrograde accretion; the latter, however, transfers more angular momentum per unit mass, because the counterrotating innermost last stable orbit is larger. If, therefore, MBH growth occurs through accretion of incoherent packets of material with $\delta m_{\rm acc} \ll M_{\rm BH}$ and random angular momentum orientation, its final spin will be generally close to zero (see Figure 3).

4.2. Connecting MBH Spins to Their Accretion History. Following our discussion on the MBH mass growth in Section 3, the cosmic MBH spin distribution might therefore be multimodal, and strongly correlated with the merging and accretion history (and therefore morphology) of the host galaxies [104]. In particular, MBHs residing in giant ellipticals that experienced several gas-rich mergers are expected to be highly spinning. During such mergers infact, the two sinking MBH spins align quickly [183, 184] to the nuclear disk angular momentum, and both coherent accretion and the final coalescence contribute to the remnant MBH spinup.

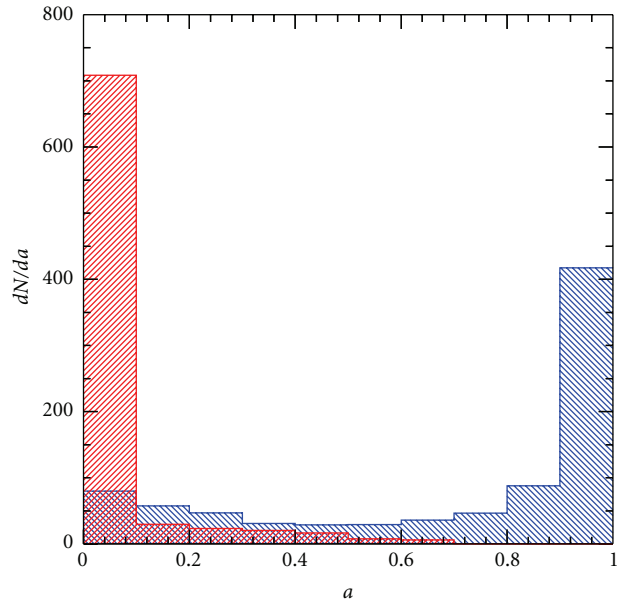


Figure 3: Spin distribution of MBH coalescence remnants for two different MBH accretion models. Red histogram: accretion proceeds in a chaotic fashion according to [103], resulting in small spins; blue histogram: accretion is coherent, resulting in efficient spinup. The seed BH spins at formation were set to zero.

Alignment has also important consequences for the gravitation recoil; since binaries with aligned spins experience the weakest kicks, their coalescence remnants will be generally unable to escape the galaxy core. For dry mergers, however, no spin alignment is expected, and the MBH remnants will generally have a broader spin distribution, experiencing higher kicks that can in principle evacuate them from their host galaxies. Many disk galaxies, conversely, may have not experienced any major merger event at low redshift. In this case, if the central MBH spin was initially high, random (chaotic) accretion of molecular clouds or tidally disrupted stars would bring it to low (possibly close to zero) values [104]. However, we should mention that, in many cases, studies of the kinematic in the nuclear region of spirals have revealed disk-like structures in the gas distribution (see, e.g., [185, 186]), leaving the possibility of substantial coherent accretion open. Although the question of accretion in spirals is not a neat as in ellipticals, it is likely that the average spins of the MBHs hosted by the formers is lower. This general picture predicts high spins in bright quasars (and present giant ellipticals). If this is the case, and we assume the "spin paradigm" as the source of radio jets, then elliptical galaxies have to be generally radio louder than spirals, in agreement with the finding of Sikora and collaborators [187]. This scenario is supported also by Capetti and Balmaverde [188] that found that radio bimodality correlates with bimodality of stellar brightness profiles in galactic nuclei. The inner regions of radio-loud galaxies display shallow cores, whereas radio quiet galaxies, including nearby low-luminosity Seyferts, have instead power-law brightness profiles (cusps) and preferentially reside in S0 and spiral galaxies. However, in the modeling of radio loud/radio quiet sources, retrograde accretion is possibly a key ingredient [189, 190] (see [191] and reference therein for a captivating scenario connecting radio loudness to retrograde accretion). On the other hand, at the time of writing, the paucity (and uncertainty) of reliable spin measurements does not offer a robust guideline for theoretical modeling.

5. Probing Massive Black Holes in the Era of Gravitational Wave Observations

The MBH evolution picture emerging from the previous sections is rich and diverse, and only loosely bound by observations. Different MBH seeding models and accretion recipes are consistent with current observational constraints about the present day MBH mass density and mass function, the quasar optical and X-ray luminosity functions, the count of faint X-ray sources in deep fields, and the unresolved X-ray background [192, 193]. On the other hand, the MBH spin distribution is basically unconstrained by observations; just a handful of spin measurements are available, and conjectured spin distributions rely on theoretical models (e.g., the "spin paradigm" for jet production can be used to infer that MBHs in bright quasars are close to maximally spinning [194]).

Several future observations can help placing better constrains to current MBH evolution models. For example, the work in [195] argues that the imprint of the seed BH population should be particularly evident at the small end of the MBH mass function. We therefore need to target the nuclei of dwarf galaxies, to unveil their central MBHs. Depending on the seed formation mechanism, such MBHs might systematically lie above or below the $M-\sigma$ relation observed for masses $>10^6\,M_\odot$. Moreover, the MBH occupation fraction in such small galaxies may offer valuable insights on the seed BH formation efficiency [196] and on the effectiveness of gravitational recoils, ultimately resulting in useful information on the MBH spins. Counting high redshift faint objects in future deep fields with the JWST will also help constraining the evolution of the MBH population at high redshift [197], while long exposures with the X-ray telescope Athena (http://www.mpe.mpg.de/athena/home.php?lang=en) will likely increase the number of measured spins.

The next two decades, will witness the dawn of GW astronomy. While signals coming from compact stars and binaries fall in the observational domain of operating and planned ground-based interferometers (such as LIGO, VIRGO, and the proposed Einstein Telescope (ET)), MBH binaries are expected to be among the primary actors on the upcoming low-frequency stage. The 10^{-4}–10^{-1} Hz window is going to be probed by the spaceborne New Gravitational Observatory (NGO), which is under consideration by ESA as the possible next L1 space mission; whereas the nHz (10^{-9}–10^{-7} Hz) range will be covered by precise timing of an ensemble of millisecond pulsars (forming a so called pulsar timing array, PTA [198–202]). (Following the termination of the ESA-NASA partnership at the basis of the Laser Interferometer Space Antenna (LISA) project, a new design study is being submitted to ESA, for a smaller mission called NGO.

More information can be found at https://lisa-light.aei.mpg.de/bin/view) While NGO will detect 10^3–$10^7\,M_\odot$ coalescing MBH binaries, mostly at $z > 2$, PTAs will be sensitive to the inspiral phase of genuine supermassive systems ($M > 10^8\,M_\odot$) at lower redshift ($z < 1$), providing a complementary probe of the MBH population in the Universe, as shown in Figure 4.

NGO will provide a unique survey of coalescing MBH binaries up to $z \approx 15$ [203–210], offering the unique possibility of disentangling different seed formation mechanisms [211–213]. (All the relevant literature was written considering LISA as reference mission concept; numbers and rates are similar for NGO.) Masses and spins of coalescing MBHs will be measured with unprecedented precision (see, e.g., [214–217]), placing strong constrains on the MBH accretion history. Moreover, NGO will also give the opportunity of detecting several (up to a hundred, but the rates are highly uncertain [218, 219]) extreme mass ratio inspirals (EMRIs), that is, inspirals of stellar mass compact objects into 10^5–$10^6\,M_\odot$ MBHs, out to $z \lesssim 0.7$. This is particularly important, not only for the unique opportunity to test GR in the strong field regime [220], but also because it will allow to unveil the low end of the mass function [221] and the spin distribution (both MBH mass and spin will be measured with an accuracy of <1% from GW observations [222]) of dormant MBHs. Infact, all the MBH observations beyond the local Universe are inevitably biased toward active systems; conversely, EMRIs are likely to occur in dense relaxed nuclei (e.g., the Milky Way) with no preference for the active ones.

PTAs exploit the characteristic fingerprint left by passing GWs in the time of arrival of the radio pulses propagating from the pulsar to the receiver on Earth [223–227]. Given the low-frequency window (10^{-9}–10^{-7} Hz), PTAs will be sensitive to the collective signal coming from the inspiralling population of supermassive, low redshift binaries [228, 229], providing a strong test for the effectiveness of the galaxy merger process at low redshift, and helping placing constrains to the high mass end of the MBH mass function. The work in [230] found that several hundred sources will contribute to the signal at a 1 ns level, considered the ultimate goal for the Square Kilometre Array (SKA [201]). Interestingly, such systems are far from coalescence, and they can still retain much of their original eccentricity against GW circularization [121, 231, 232]. Eccentricity measurements of individually resolved sources may help in constraining the evolution of MBH binaries, testing our current models of their dynamical evolution in star/gas dominated environments. Moreover, the identification of putative electromagnetic counterparts will open new avenues in the era of multimessenger astronomy [233, 234].

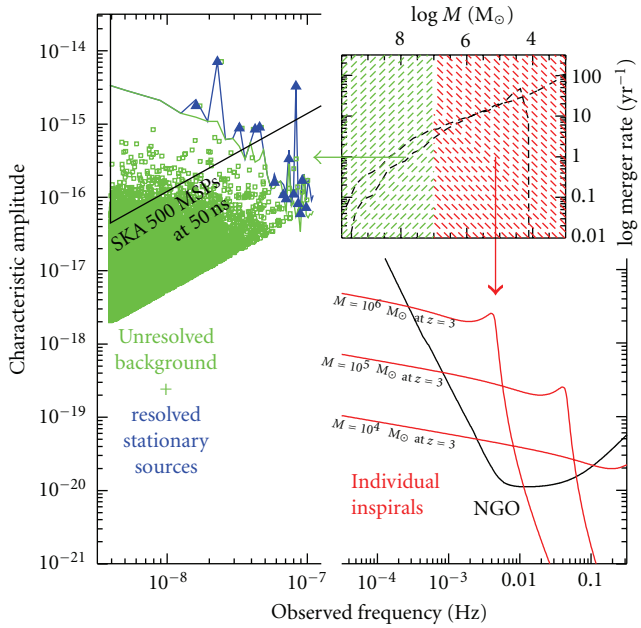


Figure 4: The GW landscape. A representative summary of GW observations of MBH binaries with PTAs (through the detection of a GW stochastic background) and NGO (via the direct detection the final stage of the coalescence of individual binary systems). In the main panel, we show the characteristic amplitude of GW signals as a function of frequency, compared to the sensitivity of SKA and NGO. On the left side, we show a Montecarlo realization of the GW signal coming from a cosmological population of MBH binaries: green dots represent the strain of each individual source, the overall signal is shown as a blue solid line, individually resolvable sources are shown as blue triangles, and unresolved background is shown as a green solid line. On the right side, red lines track the inspiral of selected MBH binaries at $z = 3$. The upper right panel shows the massive black hole binary coalescence rate as a function of the total mass M for two representative MBH population models. The green- and red-shaded areas highlight the portion of the mass function that will be probed by PTA and NGO, respectively, highlighting the complementarity of the two probes.

6. Summary

The cosmic evolution of MBHs is one of the most diverse and fascinating puzzles of the structure formation in the Universe. MBHs are ubiquitous in galactic nuclei today, and their tight correlations with their hosts witness an intimate link between their evolution and galaxy formation. In this review, we followed this link reconstructing the key elements of the puzzle. It is now widely accepted that the seeds of the MBHs powering quasars and lurking in the galactic centers today were formed in protogalactic DM halos of 10^6–$10^8\,M_\odot$ at high redshift, $z \approx 20$. The details of the formation process are largely unknown, and at least three distinctive mechanisms have been put forward: (i) light ($\sim 100\,M_\odot$) seed formation as PopIII star remnants; (ii) massive ($\sim 10^5\,M_\odot$) seed formation from direct collapse of protogalactic massive gas clouds; (iii) intermediate ($\sim 10^3\,M_\odot$) seed formation following runaway dynamics in protogalactic star clusters. The subsequent seed growth is largely determined by the environmental conditions in which they are formed. Seeds populating the deepest overdensities in the early Universe are likely to grow fast, possibly at a supercritical rate, fed by massive streams of cold gas, forming the rare massive objects powering the $z \approx 7$ quasars we see today. The MBHs powering the much more common $z \approx 2$ quasars may instead have experienced a less

violent fate, undergoing a series of merger events triggering Eddington-limited accretion episodes lasting for $\sim 10^7$ yr. Many MBH binaries inevitably form during such process, which are now hunted by several observational campaigns, and will be in the future primary targets of GW probes. At the small end of the mass function, MBHs populating spirals and dwarf galaxies evolved in a very quiet environment, without possibly even experiencing a single merger event. Their mass growth is thought to be driven by short accretion episodes following tidal disruption of stars, capture of molecular clouds, and inspirals of compact objects. Different growth modes are expected to result in different spins of the growing objects. Although coherent prolonged accretion via a radiatively efficient thin disk effectively spins up the accreting objects, both MBH binary mergers and accretion of incoherent packets of matter with randomly oriented angular momenta are expected to produce a spindown. It is therefore likely that the MBH spin distribution reflects their accretion history, being linked with the morphology of their host galaxies. Current measurements, however, are still too sparse and uncertain to test theoretical hypothesis. Future high redshift observations, together with the detection of smaller and smaller MBHs in local galaxies, will help in shedding further light on the cosmic evolution of these fascinating objects. In particular, forthcoming GW observations will make possible mass and spin measurements of coalescing MBH binaries out to $z \approx 15$ with unprecedented precision, offering a unique opportunity to study their growth along the cosmic history from the very beginning.

Acknowledgments

I am indebted to Marta Volonteri for suggesting a lot of literature relevant to this review, for making Figure 2 available, and for the modifications to Figure 1. I thank Rennan Barkana for the original version of Figure 1, and Massimo Dotti for his detailed and thoughtful comments to the paper.

References

[1] X. Fan, V. K. Narayanan, R. H. Lupton et al., "A survey of z > 5.8 quasars in the sloan digital sky survey. I. Discovery of three new quasars and the spatial density of luminous quasars at z ~ 6," *Astronomical Journal*, vol. 122, no. 6, pp. 2833–2849, 2001.

[2] X. Fan, M. A. Strauss, D. P. Schneider et al., "A survey of z > 5.7 quasars in the Sloan Digital Sky Survey. II—discovery of three additional quasars at z > 6," *Astronomical Journal*, vol. 125, no. 4, pp. 1649–1659, 2003.

[3] D. J. Mortlock, S. J. Warren, B. P. Venemans et al., "A luminous quasar at a redshift of z = 7.085," *Nature*, vol. 474, no. 7353, pp. 616–619, 2011.

[4] G. T. Richards, A. D. Myers, A. G. Gray et al., "Efficient photometric selection of quasars from the sloan digital sky survey. II. ~1,000,000 quasars from data release 6," *Astrophysical Journal, Supplement Series*, vol. 180, no. 1, pp. 67–83, 2009.

[5] A. M. Ghez, S. Salim, S. D. Hornstein et al., "Stellar orbits around the galactic center black hole," *Astrophysical Journal*, vol. 620, no. 2, pp. 744–757, 2005.

[6] S. Gillessen, F. Eisenhauer, T. K. Fritz et al., "The orbit of the star S2 around Sgr A∗ from very large telescope and Keck data," *Astrophysical Journal*, vol. 707, no. 2, pp. L114–L117, 2009.

[7] B. M. Peterson, L. Ferrarese, K. M. Gilbert et al., "Central masses and broad-line region sizes of active galactic nuclei. II. A homogeneous analysis of a large reverberation-mapping database," *Astrophysical Journal*, vol. 613, no. 2, pp. 682–699, 2004.

[8] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal*, vol. 698, no. 1, pp. 198–221, 2009.

[9] D. Richstone, E. A. Ajhar, and R. Bender, "Supermassive black holes and the evolution of galaxies," *Nature*, vol. 385, no. 6701, supplement, article A14, 1998.

[10] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[11] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.

[12] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal Letters*, vol. 539, no. 1, part 2, pp. L13–L16, 2000.

[13] S. Tremaine, K. Gebhardt, R. Bender et al., "The slope of the black hole mass versus velocity dispersion correlation," *Astrophysical Journal*, vol. 574, no. 2, pp. 740–753, 2002.

[14] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *Astrophysical Journal*, vol. 589, no. 1, pp. L21–L24, 2003.

[15] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *Astrophysical Journal Letters*, vol. 604, no. 2, pp. L89–L92, 2004.

[16] A. W. Graham, C. A. Onken, E. Athanassoula, and F. Combes, "An expanded Mbh-σ diagram, and a new calibration of active galactic nuclei masses," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2211–2228, 2011.

[17] L. Ferrarese, "Beyond the bulge: a fundamental relation between supermassive black holes and dark matter halos," *Astrophysical Journal*, vol. 578, no. 1, pp. 90–97, 2002.

[18] J. Kormendy, R. Bender, and M. E. Cornell, "Supermassive black holes do not correlate with galaxy disks or pseudobulges," *Nature*, vol. 469, no. 7330, pp. 374–376, 2011.

[19] B. J. Boyle and R. J. Terlevich, "The cosmological evolution of the QSO luminosity density and of the star formation rate," *Monthly Notices of the Royal Astronomical Society*, vol. 293, no. 2, pp. L49–L51, 1998.

[20] F. Shankar, D. H. Weinberg, and J. Miralda-Escudé, "Self-consistent models of the AGN and black hole populations: duty cycles, accretion rates, and the mean radiative efficiency," *Astrophysical Journal*, vol. 690, no. 1, pp. 20–41, 2009.

[21] D. B. Sanders and I. F. Mirabel, "Luminous infrared galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 34, no. 1, pp. 749–792, 1996.

[22] D. M. Alexander, F. E. Bauer, W. N. Brandt et al., "The Chandra Deep Field North survey. XIV. X-Ray-detected obscured AGNs and starburst galaxies in the bright submillimeter source population," *Astronomical Journal*, vol. 125, no. 2, pp. 383–397, 2003.

[23] A. M. Swinbank, I. Smail, S. C. Chapman, A. W. Blain, R. J. Ivision, and W. C. Keel, "The rest-frame optical spectra of

Scuba galaxies," *Astrophysical Journal*, vol. 617, no. 1, pp. 64–80, 2004.

[24] S. D. M. White and M. J. Rees, "Core condensation in heavy halos—a two-stage theory for galaxy formation and clustering," *Monthly Notices of the Royal Astronomical Society*, vol. 183, pp. 341–358, 1978.

[25] G. Kauffmann and M. Haehnelt, "A unified model for the evolution of galaxies and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 311, no. 3, pp. 576–588, 2000.

[26] M. Volonteri, F. Haardt, and P. Madau, "The assembly and merging history of supermassive black holes in hierarchical models of galaxy formation," *Astrophysical Journal*, vol. 582, no. 2, pp. 559–573, 2003.

[27] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-Ray background, supermassive black holes, and galaxy spheroids," *Astrophysical Journal, Supplement Series*, vol. 163, no. 1, pp. 1–49, 2006.

[28] J. C. Minos and L. Hernquist, "Gasdynamics and starbursts in major mergers," *Astrophysical Journal*, vol. 464, no. 2, pp. 641–663, 1996.

[29] N. Menci, A. Cavaliere, A. Fontana, E. Giallongo, F. Poli, and V. Vittorini, "Early hierarchical formation of massive galaxies triggered by interactions," *Astrophysical Journal*, vol. 604, no. 1, pp. 12–17, 2004.

[30] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[31] T. Di Matteo, V. Springel, and L. Ilernquist, "Energy input from quasars regulates the growth and activity of black holes and their host galaxies," *Nature*, vol. 433, no. 7026, pp. 604–607, 2005.

[32] J. Silk and M. J. Rees, "Quasars and galaxy formation," *Astronomy and Astrophysics*, vol. 331, no. 2, pp. L1–L4, 1998.

[33] A. King, "Black holes, galaxy formation, and the MBH-σ relation," *Astrophysical Journal Letters*, vol. 596, no. 1, pp. L27–L29, 2003.

[34] Z. Haiman and K. Menou, "On the cosmological evolution of the luminosity function and the accretion rate of quasars," *Astrophysical Journal*, vol. 531, no. 1, pp. 42–51, 2000.

[35] J. S. B. Wyithe and A. Loeb, "Self-regulated growth of supermassive black holes in galaxies as the origin of the optical and X-Ray luminosity functions of quasars," *Astrophysical Journal*, vol. 595, no. 2, pp. 614–623, 2003.

[36] R. K. Malbon, C. M. Baugh, C. S. Frenk, and C. G. Lacey, "Black hole growth in hierarchical galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 4, pp. 1394–1414, 2007.

[37] T. Di Matteo, J. Colberg, V. Springel, L. Hernquist, and D. Sijacki, "Direct cosmological simulations of the growth of black holes and galaxies," *Astrophysical Journal*, vol. 676, no. 1, pp. 33–53, 2008.

[38] T. Padmanabhan, *Structure Formation in the Universe*, Cambridge University Press, New York, NY, USA, 1993.

[39] P. Coles and F. Lucchin, *Cosmology, Cosmological Models, Large Scale Structure, Origin, Evolution*, John Wiley & Sons, Chichester, UK, 1995.

[40] W. H. Press and P. Schechter, "Formation of galaxies and clusters of galaxies by self-similar gravitational condensation," *Astrophysical Journal*, vol. 17, pp. 425–438, 1974.

[41] R. Barkana and A. Loeb, "In the beginning: the first sources of light and the reionization of the universe," *Physics Report*, vol. 349, no. 2, pp. 125–238, 2001.

[42] Z. Haiman, A. A. Thoul, and A. Loeb, "Cosmological formation of low-mass objects," *Astrophysical Journal*, vol. 464, no. 2, pp. 523–528, 1996.

[43] T. Abel, G. L. Bryan, and M. L. Norman, "The formation of the first star in the universe," *Science*, vol. 295, no. 5552, pp. 93–98, 2002.

[44] P. R. Shapiro and H. Kang, "Hydrogen molecules and the radiative cooling of pregalactic shocks," *Astrophysical Journal*, vol. 318, pp. 32–65, 1987.

[45] M. Tegmark, J. Silk, M. J. Rees, A. Blanchard, T. Abel, and F. Palla, "How small were the first cosmological objects?" *Astrophysical Journal*, vol. 474, no. 1, pp. 1–12, 1997.

[46] M. Volonteri, "Formation of supermassive black holes," *Astronomy and Astrophysics Review*, vol. 18, no. 3, pp. 279–315, 2010.

[47] D. Galli and F. Palla, "The chemistry of the early universe," *Astronomy and Astrophysics*, vol. 335, no. 2, pp. 403–420, 1998.

[48] V. Bromm, P. S. Coppi, and R. B. Larson, "The formation of the first stars. I. The primordial star-forming cloud," *Astrophysical Journal*, vol. 564, no. 1, pp. 23–51, 2002.

[49] T. Abel, G. L. Bryan, and M. L. Norman, "The formation and fragmentation of primordial molecular clouds," *Astrophysical Journal*, vol. 540, no. 1, pp. 39–44, 2000.

[50] N. Yoshida, K. Omukai, L. Hernquist, and T. Abel, "Formation of primordial stars in a ACDM universe," *Astrophysical Journal*, vol. 652, no. 1, pp. 6–25, 2006.

[51] L. Gao, N. Yoshida, T. Abel, C. S. Frenk, A. Jenkins, and V. Springel, "The first generation of stars in the Λ cold dark matter cosmology," *Monthly Notices of the Royal Astronomical Society*, vol. 378, no. 2, pp. 449–468, 2007.

[52] A. Heger, C. L. Fryer, S. E. Woosley, N. Langer, and D. D. H. Hartmann, "How massive single stars end their life," *Astrophysical Journal*, vol. 591, no. 1, pp. 288–300, 2003.

[53] P. Madau and M. J. Rees, "Massive black holes as population III remnants," *Astrophysical Journal*, vol. 551, no. 1, pp. L27–L30, 2001.

[54] C. F. McKee and J. C. Tan, "The formation of the first stars. II. Radiative feedback processes and implications for the initial mass function," *Astrophysical Journal*, vol. 681, no. 2, pp. 771–797, 2008.

[55] M. J. Turk, T. Abel, and B. O'Shea, "The formation of population III binaries from cosmological initial conditions," *Science*, vol. 325, no. 5940, pp. 601–605, 2009.

[56] A. Stacy, T. H. Greif, and V. Bromm, "The first stars: formation of binaries and small multiple systems," *Monthly Notices of the Royal Astronomical Society*, vol. 403, no. 1, pp. 45–60, 2010.

[57] P. C. Clark, S. C.O. Glover, R. S. Klessen, and V. Bromm, "Gravitational fragmentation in turbulent primordial gas and the initial mass function of Population III stars," *Astrophysical Journal Letters*, vol. 727, no. 2, article 110, 2011.

[58] T. H. Greif, V. Springel, S. D. M. White et al., "Simulations on a moving mesh: the clustered formation of population III protostars," *Astrophysical Journal*, vol. 737, no. 2, article 75, 2011.

[59] S. F. Portegies Zwart, H. Baumgardt, P. Hut, J. Makino, and S. L. W. McMillan, "Formation of massive black holes through runaway collisions in dense young star clusters," *Nature*, vol. 428, no. 6984, pp. 724–726, 2004.

[60] M. G. Haehnelt and M. J. Rees, "The formation of nuclei in newly formed galaxies and the evolution of the quasar population," *Royal Astronomical Society*, vol. 263, no. 1, pp. 168–178, 1993.

[61] A. Loeb and F. A. Rasio, "Collapse of primordial gas clouds and the formation of quasar black holes," *Astrophysical Journal*, vol. 432, no. 1, pp. 52–61, 1994.

[62] D. J. Eisenstein and A. Loeb, "Origin of quasar progenitors from the collapse of low-spin cosmological perturbations," *Astrophysical Journal*, vol. 443, no. 1, pp. 11–17, 1995.

[63] V. Bromm and A. Loeb, "Formation of the first supermassive black holes," *Astrophysical Journal*, vol. 596, no. 1, pp. 34–46, 2003.

[64] S. M. Koushiappas, J. S. Bullock, and A. Dekel, "Massive black hole seeds from low angular momentum material," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 1, pp. 292–304, 2004.

[65] M. C. Begelman, M. Volonteri, and M. J. Rees, "Formation of supermassive black holes by direct collapse in pre-galactic haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 1, pp. 289–298, 2006.

[66] G. Lodato and P. Natarajan, "Supermassive black hole formation during the assembly of pre-galactic discs," *Monthly Notices of the Royal Astronomical Society*, vol. 371, no. 4, pp. 1813–1823, 2006.

[67] C. Dotan, E. M. Rossi, and N. J. Shaviv, "A lower limit on the halo mass to form supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 417, no. 4, pp. 3035–3046, 2011.

[68] M. Dijkstra, Z. Haiman, A. Mesinger, and J. S. B. Wyithe, "Fluctuations in the high-redshift Lyman-Werner background: close halo pairs as the origin of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 391, no. 4, pp. 1961–1972, 2008.

[69] J. S. Bullock, A. Dekel, T. S. Kolatt et al., "A universal angular momentum profile for galactic halos," *Astrophysical Journal*, vol. 555, no. 1, pp. 240–257, 2001.

[70] I. Shlosman, J. Frank, and M. C. Begelman, "Bars within bars: a mechanism for fuelling active galactic nuclei," *Nature*, vol. 338, no. 6210, pp. 45–47, 1989.

[71] J. H. Wise, M. J. Turk, and T. Abel, "Resolving the formation of protogalaxies. II. Central gravitational collapse," *Astrophysical Journal*, vol. 682, no. 2, pp. 745–757, 2008.

[72] F. Hoyle and W. A. Fowler, "On the nature of strong radio sources," *Monthly Notices of the Royal Astronomical Society*, vol. 125, article 169, 1963.

[73] M. Saijo, T. W. Baumgarte, S. L. Shapiro, and M. Shibata, "Collapse of a rotating supermassive star to a supermassive black hole: post-Newtonian simulations," *Astrophysical Journal*, vol. 569, no. 1, pp. 349–361, 2002.

[74] M. Shibata and S. L. Shapiro, "Collapse of a rotating supermassive star to a supermassive black hole: fully relattvistic simulations," *Astrophysical Journal*, vol. 572, no. 1, pp. L39–L43, 2002.

[75] P. J. Montero, H. T. Janka, and E. Mueller, "Relativistic collapse and explosion of rotating supermassive stars with thermonuclear effects," http://arxiv.org/abs/1108.3090.

[76] M. C. Begelman and M. J. Rees, "The fate of dense stellar systems," *Monthly Notices of the Royal Astronomical Society*, vol. 185, pp. 847–860, 1978.

[77] T. Ebisuzaki, "Missing link found? The "runaway" path to supermassive black holes," *Astrophysical Journal*, vol. 562, no. 1, pp. L19–L22, 2001.

[78] S. F. Portegies Zwart and S. L. W. McMillan, "The runaway growth of intermediate-mass black holes in dense star clusters," *Astrophysical Journal*, vol. 576, no. 2, pp. 899–907, 2002.

[79] M. A. Gürkan, M. Freitag, and F. A. Rasio, "Formation of massive black holes in dense star clusters. I. Mass segregation and core collapse," *Astrophysical Journal*, vol. 604, no. 2, pp. 632–652, 2004.

[80] L. Spitzer Jr., "Equipartition and the formation of compact nuclei in spherical stellar systems," *Astrophysical Journal*, vol. 158, article L139, 1969.

[81] L. Spitzer Jr., *Dynamical Evolution of Globular Clusters*, Princeton University Press, Princeton, NJ, USA, 1987.

[82] S. F. Portegies Zwart, J. Makino, S. L. W. McMillan, and P. Hut, "Star cluster ecology: III. Runaway collisions in young compact star clusters," *Astronomy and Astrophysics*, vol. 348, no. 1, pp. 117–126, 1999.

[83] B. Devecchi and M. Volonteri, "Formation of the first nuclear clusters and massive black holes at high redshift," *Astrophysical Journal*, vol. 694, no. 1, pp. 302–313, 2009.

[84] F. Santoro and J. M. Shuli, "Critical metallicity and fine-structure emission of primordial gas enriched by the first stars," *Astrophysical Journal*, vol. 643, no. 1, pp. 26–37, 2006.

[85] M. B. Davies, M. C. Miller, and J. M. Bellovary, "SMBH formation via gas accretion in nuclear stellar clusters," *The Astrophysical Journal*, vol. 740, no. 2, article L42, 2011.

[86] A. Soltan, "Masses of quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 200, pp. 115–122, 1982.

[87] Q. Yu and S. Tremaine, "Observational constraints on growth of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 4, pp. 965–976, 2002.

[88] M. Elvis, G. Risaliti, and G. Zamorani, "Most supermassive black holes must be rapidly rotating," *Astrophysical Journal*, vol. 565, no. 2, pp. L75–L77, 2002.

[89] A. Marconi, G. Risaliti, R. Gilli, L. K. Hunt, R. Maiolino, and M. Salvati, "Local supermassive black holes, relics of active galactic nuclei and the X-Ray background," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 1, pp. 169–185, 2004.

[90] F. Shankar, P. Salucci, G. L. Granato, G. De Zotti, and L. Danese, "Supermassive black hole demography: the match between the local and accreted mass functions," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 4, pp. 1020–1030, 2004.

[91] N. I. Shakura and R. A. Sunyaev, "Black holes in binary systems. Observational appearance," *Astronomy & Astrophysics*, vol. 24, pp. 337–355, 1973.

[92] F. E. Bauer, D. M. Alexander, W. N. Brandt et al., "The fall of active galactic nuclei and the rise of star-forming galaxies: a close look at the Chandra Deep Field X-Ray number counts," *Astronomical Journal*, vol. 128, no. 5, pp. 2048–2065, 2004.

[93] F. Fiore, A. Grazian, P. Santini et al., "Unveiling obscured accretion in the Chandra Deep Field-South," *Astrophysical Journal*, vol. 672, no. 1, pp. 94–101, 2008.

[94] E. Treister and C. M. Urry, "The cosmic history of black hole growth from deep multiwavelength surveys," http://arxiv.org/abs/1112.0320.

[95] S. L. Ellison, D. R. Patton, J. T. Mendel, and J. M. Scudder, "Galaxy pairs in the Sloan Digital Sky Survey—IV. Interactions trigger active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 3, pp. 2043–2053, 2011.

[96] Z. Haiman and A. Loeb, "What is the highest plausible redshift of luminous quasars?" *Astrophysical Journal*, vol. 552, no. 2, pp. 459–463, 2001.

[97] A. Dekel and Y. Birnboim, "Galaxy bimodality due to cold flows and shock heating," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 1, pp. 2–20, 2006.

[98] A. Dekel, Y. Birnboim, G. Engel et al., "Cold streams in early massive hot haloes as the main mode of galaxy formation," *Nature*, vol. 457, no. 7228, pp. 451–454, 2009.

[99] J. Yoo and J. Miralda-Escudé, "Formation of the black holes in the highest redshift quasars," *Astrophysical Journal*, vol. 614, no. 1, pp. L25–L28, 2004.

[100] Y. Li, L. Hernquist, B. Robertson et al., "Formation of z ~ 6 quasars from hierarchical galaxy mergers," *Astrophysical Journal*, vol. 665, no. 1, pp. 187–208, 2007.

[101] J. M. Bardeen, "Kerr metric black holes," *Nature*, vol. 226, no. 5240, pp. 64–65, 1970.

[102] K. S. Thorne, "Disk-accretion onto a black hole. II. Evolution of the hole," *Astrophysical Journal*, vol. 191, pp. 507–520, 1974.

[103] A. R. King, S. H. Lubow, G. I. Ogilvie, and J. E. Pringle, "Aligning spinning black holes and accretion discs," *Monthly Notices of the Royal Astronomical Society*, vol. 363, no. 1, pp. 49–56, 2005.

[104] M. Volonteri, M. Sikora, and J. P. Lasota, "Black hole spin and galactic morphology," *Astrophysical Journal*, vol. 667, no. 2, pp. 704–713, 2007.

[105] M. C. Begelman, "Black holes in radiation-dominated gas—an analogue of the Bondi accretion problem," *Monthly Notices of the Royal Astronomical Society*, vol. 184, pp. 53–67, 1978.

[106] M. C. Begelman and D. L. Meier, "Thick accretion disks—self-similar, supercritical models," *Astrophysical Journal*, vol. 253, pp. 873–896, 1982.

[107] J. M. Stone, J. E. Pringle, and M. C. Begelman, "Hydrodynamical non-radiative accretion flows in two dimensions," *Monthly Notices of the Royal Astronomical Society*, vol. 310, no. 4, pp. 1002–1016, 1999.

[108] J. F. Hawley and S. A. Balbus, "The dynamical structure of nonradiative black hole accretion flows," *Astrophysical Journal*, vol. 573, no. 2, pp. 738–748, 2002.

[109] I. V. Igumenshchev, R. Narayan, and M. A. Abramowicz, "Three-dimensional magnetohydrodynamic simulations of radiatively inefficient accretion flows," *Astrophysical Journal*, vol. 592, no. 2, pp. 1042–1059, 2003.

[110] D. Proga and M. C. Begelman, "Accretion of low angular momentum material onto black holes: two-dimensional magnetohydrodynamic case," *Astrophysical Journal*, vol. 592, no. 2, pp. 767–781, 2003.

[111] K. Ohsuga, M. Mori, T. Nakamoto, and S. Mineshige, "Supercritical accretion flows around black holes: two-dimensional, radiation pressure-dominated disks with photon trapping," *Astrophysical Journal*, vol. 628, no. 1, pp. 368–381, 2005.

[112] M. Volonteri and M. J. Rees, "Rapid growth of high-redshift black holes," *Astrophysical Journal*, vol. 633, no. 2, pp. 624–629, 2005.

[113] M. Dotti, A. Sesana, and R. Decarli, "Massive black hole binaries: dynamical evolution and observational signatures," http://arxiv.org/abs/1111.0664.

[114] M. Colpi, L. Mayer, and F. Governato, "Dynamical friction and the evolution of satellites in virialized halos: the theory of linear response," *Astrophysical Journal*, vol. 525, no. 2, pp. 720–733, 1999.

[115] E. C. Ostriker, "Dynamical friction in a gaseous medium," *Astrophysical Journal*, vol. 513, no. 1, pp. 252–258, 1999.

[116] A. Escala, R. B. Larson, P. S. Coppi, and D. Mardones, "The role of gas in the merging of massive black holes in galactic nuclei. II. Black hole merging in a nuclear gas disk," *Astrophysical Journal*, vol. 630, no. 1, pp. 152–166, 2005.

[117] M. Dotti, M. Colpi, F. Haardt, and L. Mayer, "Supermassive black hole binaries in gaseous and stellar circumnuclear discs: orbital dynamics and gas accretion," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 3, pp. 956–962, 2007.

[118] P. C. Peters and J. Mathews, "Gravitational radiation from point masses in a keplerian orbit," *Physical Review*, vol. 131, no. 1, pp. 435–440, 1963.

[119] P. Berczik, D. Merritt, R. Spurzem, and H. P. Bischof, "Efficient merger of binary supermassive black holes in nonaxisymmetric galaxies," *Astrophysical Journal*, vol. 642, no. 1, pp. L21–L24, 2006.

[120] F. M. Khan, A. Just, and D. Merritt, "Efficient merger of binary supermassive black holes in merging galaxies," *Astrophysical Journal Letters*, vol. 732, no. 2, article 89, 2011.

[121] M. Preto, I. Berentzen, P. Berczik, and R. Spurzem, "Fast coalescence of massive black hole binaries from mergers of galactic nuclei: implications for low-frequency gravitational-wave astrophysics," *Astrophysical Journal Letters*, vol. 732, no. 2, part 2, article L26, 2011.

[122] A. Gualandris and D. Merritt, "Long-term evolution of massive black hole binaries. IV. Mergers of galaxies with collisionally relaxed nuclei," *The Astrophysical Journal*, vol. 744, no. 1, p. 74, 2011, http://arxiv.org/abs/1107.4095.

[123] J. E. Barnes and L. Hernquist, "Transformations of galaxies. II. gasdynamics in merging disk galaxies," *Astrophysical Journal*, vol. 471, no. 1, pp. 115–142, 1996.

[124] L. L. Cowie, A. Songaila, E. M. Hu, and J. G. Cohen, "New insight on galaxy formation and evolution from keck spectroscopy of the Hawaii deep fields," *Astronomical Journal*, vol. 112, no. 3, pp. 839–864, 1996.

[125] L. L. Cowie, A. J. Barger, M. W. Bautz, W. N. Brandt, and G. P. Garmire, "The redshift evolution of the 2-8 keV X-Ray luminosity function," *Astrophysical Journal*, vol. 584, no. 2, pp. L57–L60, 2003.

[126] Y. Ueda, M. Akiyama, K. Ohta, and T. Miyaji, "Cosmological evolution of the hard X-Ray active galactic nucleus luminosity function and the origin of the hard X-Ray background," *Astrophysical Journal*, vol. 598, no. 2, pp. 886–908, 2003.

[127] A. Merloni, "The anti-hierarchical growth of supermassive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 353, no. 4, pp. 1035–1047, 2004.

[128] P. F. Hopkins, K. Bunpy, L. Hernquist, and R. S. Ellis, "Observational evidence for the coevolution of galaxy mergers, quasars, and the blue/red galaxy transition," *Astrophysical Journal*, vol. 659, no. 2, pp. 976–996, 2007.

[129] T. Urrutia, M. Lacy, and R. H. Becker, "Evidence for quasar activity triggered by galaxy mergers in HST observations of dust-reddened quasars," *Astrophysical Journal*, vol. 674, no. 1, pp. 80–96, 2008.

[130] S. Komossa, V. Burwitz, G. Hasinger, P. Predehl, J. S. Kaastra, and Y. Ikebe, "Discovery of a binary active galactic nucleus in the ultraluminous infrared galaxy NGC 6240 using Chandra," *Astrophysical Journal*, vol. 582, no. 1, pp. L15–L19, 2003.

[131] M. J. Valtonen, H. J. Lehto, K. Nilsson et al., "A massive binary black-hole system in OJ 287 and a test of general relativity," *Nature*, vol. 452, no. 7189, pp. 851–853, 2008.

[132] S. Komossa, H. Zhou, and H. Lu, "A recoiling supermassive black hole in the quasar SDSS J092712.65+294344.0?" *Astrophysical Journal Letters*, vol. 678, no. 2, part 2, pp. L81–L84, 2008.

[133] T. A. Boroson and T. R. Lauer, "A candidate sub-parsec supermassive binary black hole system," *Nature*, vol. 458, no. 7234, pp. 53–55, 2009.

[134] P. Tsalmantza, R. Decarli, M. Dotti, and D. W. Hogg, "A systematic search for massive black hole binaries in the Sloan Digital Sky Survey spectroscopic sample," *Astrophysical Journal*, vol. 738, no. 1, article 20, 2011.

[135] M. Eracleous, T. A. Boroson, J. P. Halpern, and J. Liu, "A large systematic search for recoiling and close supermassive binary black holes," http://arxiv.org/abs/1106.2952.

[136] E. F. Bell, D. H. McIntosh, N. Katz, and M. D. Weinberg, "The optical and near-infrared properties of galaxies. I. Luminosity and stellar MASS functions," *Astrophysical Journal, Supplement Series*, vol. 149, no. 2, pp. 289–312, 2003.

[137] L. C. Ho, A. V. Filippenko, and W. L. W. Sargent, "The influence of bars on nuclear activity," *Astrophysical Journal*, vol. 487, no. 2, pp. 591–602, 1997.

[138] R. W. Pogge and P. Martini, "Hubble space telescope imaging of the circumnuclear environments of the CfA seyfert galaxies: nuclear spirals and fueling," *Astrophysical Journal*, vol. 569, no. 2, pp. 624–640, 2002.

[139] R. Schödel, A. Eckart, T. Alexander et al., "The structure of the nuclear stellar cluster of the Milky Way," *Astronomy and Astrophysics*, vol. 469, no. 1, pp. 125–146, 2007.

[140] M. J. Rees, "Tidal disruption of stars by black holes of 106-108 solar masses in nearby galaxies," *Nature*, vol. 333, no. 6173, pp. 523–528, 1988.

[141] J. Magorrian and S. Tremaine, "Rates of tidal disruption of stars by massive central black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 309, no. 2, pp. 447–460, 1999.

[142] X. Chen, P. Madau, A. Sesana, and F. K. Liu, "Enhanced tidal disruption rates from massive black hole binaries," *Astrophysical Journal*, vol. 697, no. 2, pp. L149–L152, 2009.

[143] X. Chen, A. Sesana, P. Madau, and F. K. Liu, "Tidal stellar disruptions by massive black hole pairs. II. Decaying binaries," *Astrophysical Journal Letters*, vol. 729, no. 1, article 13, 2011.

[144] C. Wegg and J. Nate Bode, "Multiple tidal disruptions as an indicator of binary supermassive black hole systems," *Astrophysical Journal Letters*, vol. 738, no. 1, article L8, 2011.

[145] S. F. Portegies Zwart, J. Making, S. L. W. McMillan, and P. Hut, "The lives and deaths of star clusters near the Galactic center," *Astrophysical Journal*, vol. 565, no. 1, pp. 265–279, 2002.

[146] J. P. Halpern, S. Gezari, and S. Komossa, "Follow-up Chandra observations of three candidate tidal disruption events," *Astrophysical Journal*, vol. 604, no. 2, pp. 572–578, 2004.

[147] S. Komossa, J. Halpern, N. Schartel, G. Hasinger, M. Santos-Lleo, and P. Predehl, "A huge drop in the X-Ray luminosity of the nonactive galaxy RX J1242.6-1119A, and the first postflare spectrum: testing the tidal disruption scenario," *Astrophysical Journal*, vol. 603, no. 1, pp. L17–L20, 2004.

[148] S. Gezari, D. C. Martin, B. Milliard et al., "Ultraviolet detection of the tidal disruption of a star by a supermassive black hole," *Astrophysical Journal*, vol. 653, no. 1, pp. L25–L28, 2006.

[149] P. Esquej, R. D. Saxton, S. Komossa et al., "Evolution of tidal disruption candidates discovered by XMM-Newton," *Astronomy and Astrophysics*, vol. 489, no. 2, pp. 543–554, 2008.

[150] S. Gezari, H. Tim, and C. S. Bradley, "Luminous thermal flares from quiescent supermassive black holes," *Astrophysical Journal*, vol. 698, no. 2, pp. 1367–1379, 2009.

[151] S. van Velzen, G. R. Farrar, S. Gezari et al., "Optical discovery of probable stellar tidal disruption flares," *The Astrophysical Journal*, vol. 741, no. 2, p. 73, 2011.

[152] L. E. Strubbe and E. Quataert, "Optical flares from the tidal disruption of stars by massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 4, pp. 2070–2084, 2009.

[153] G. Lodato and E. M. Rossi, "Multiband light curves of tidal disruption events," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 1, pp. 359–367, 2011.

[154] M. Milosavljević, D. Merritt, and L. C. Ho, "Contribution of stellar tidal disruptions to the X-Ray luminosity function of active galaxies," *Astrophysical Journal*, vol. 652, no. 1, pp. 120–125, 2006.

[155] C. Hopman and T. Alexander, "The effect of mass segregation on gravitational wave sources near massive black holes," *Astrophysical Journal*, vol. 645, no. 2, pp. L133–L136, 2006.

[156] M. Preto and P. Amaro-Seoane, "On strong mass segregation around a massive black hole: implications for lower-frequency gravitational-wave astrophysics," *Astrophysical Journal*, vol. 708, no. 1, pp. L42–L46, 2010.

[157] P. F. Hopkins and L. Hernquist, "Fueling low-level AGN activity through stochastic accretion of cold gas," *Astrophysical Journal, Supplement Series*, vol. 166, no. 1, pp. 1–36, 2006.

[158] A. Miyazaki and M. Tsuboi, "Dense molecular clouds in the Galactic center region. II. Statistical properties of the Galactic center molecular clouds," *Astrophysical Journal*, vol. 536, no. 1, pp. 357–367, 2000.

[159] T. Oka, T. Hasegawa, F. Sato, M. Tsuboi, A. Miyazaki, and M. Sugimoto, "Statistical properties of molecular clouds in the galactic center," *Astrophysical Journal*, vol. 562, no. 1, pp. 348–362, 2001.

[160] L. J. Sage, G. A. Welch, and L. M. Young, "The cool ISM in elliptical galaxies. I. A survey of molecular gas," *Astrophysical Journal*, vol. 657, no. 1, pp. 232–240, 2007.

[161] L. Ciotti and J. P. Ostriker, "Radiative feedback from massive black holes in elliptical galaxies: agn flaring and central starburst fueled by recycled gas," *Astrophysical Journal*, vol. 665, no. 2, pp. 1038–1056, 2007.

[162] A. Laor, "Line profiles from a disk around a rotating black hole," *Astrophysical Journal Letters*, vol. 376, no. 1, pp. 90–94, 1991.

[163] S. Schmoll, J. M. Miller, M. Volonteri et al., "Constraining the spin of the black hole in fairall 9 with Suzaku," *Astrophysical Journal*, vol. 703, no. 2, pp. 2171–2176, 2009.

[164] A. R. Patrick, J. N. Reeves, A. P. Lobban, D. Porquet, and A. G. Markowitz, "Iron line profiles in Suzaku spectra of bare Seyfert galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 411, pp. 2353–2370, 2011.

[165] I. de la Calle Pérez, A. L. Longinotti, M. Guainazzi et al., "FERO: finding extreme relativistic objects : I. Statistics of relativistic Fe K lines in radio-quiet Type 1 AGN," *Astronomy and Astrophysics*, vol. 524, article A50, 2010.

[166] L. C. Gallo, G. Miniutti, J. M. Miller et al., "Multi-epoch X-Ray observations of the Seyfert 1.2 galaxy Mrk79: bulk motion of the illuminating X-Ray source," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 1, pp. 607–619, 2011.

[167] L. W. Brenneman and C. S. Reynolds, "Constraining black hole spin via X-Ray spectroscopy," *Astrophysical Journal*, vol. 652, no. 2, pp. 1028–1043, 2006.

[168] L. W. Brenneman, C. S. Reynolds, M. A. Nowak et al., "The spin of the supermassive black hole in NGC 3783," *Astrophysical Journal*, vol. 736, no. 2, article 103, 2011.

[169] R. D. Blandford, H. Netzer, L. Woltjer, T. J. L. Courvoisier, and M. Mayor, *Saas-Fee Advanced Course*, Lecture Notes. Swiss Society for Astrophysics and Astronomy, XII, Springer, Heidelberg, Germany, 1990.

[170] R. D. Blandford and R. L. Znajek, "Electromagnetic extraction of energy from Kerr black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 179, pp. 433–456, 1977.

[171] M. Campanelli, C. O. Lousto, Y. Zlochower, B. Krishnan, and D. Merritt, "Spin flips and precession in black-hole-binary mergers," *Physical Review D*, vol. 75, no. 6, Article ID 064030, 17 pages, 2007.

[172] W. Tichy and P. Marronetti, "Binary black hole mergers: large kicks for generic spin orientations," *Physical Review D*, vol. 76, no. 6, Article ID 061502, 5 pages, 2007.

[173] C. O. Lousto and Y. Zlochower, "Hangup kicks: still larger recoils by partial spin/orbit alignment of black-hole binaries," *Physical Review Letters*, vol. 107, no. 23, Article ID 231102, 2011.

[174] E. Berti and M. Volonteri, "Cosmological black hole spin evolution by mergers and accretion," *Astrophysical Journal*, vol. 684, no. 2, pp. 822–828, 2008.

[175] S. A. Hughes and R. D. Blandford, "Black hole mass and spin coevolution by mergers," *Astrophysical Journal*, vol. 585, no. 2, pp. L101–L104, 2003.

[176] F. Pretorius, "Evolution of binary black-hole spacetimes," *Physical Review Letters*, vol. 95, no. 12, Article ID 121101, 4 pages, 2005.

[177] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. Van Meter, "Gravitational-wave extraction from an inspiraling configuration of merging black holes," *Physical Review Letters*, vol. 96, no. 11, Article ID 111102, 4 pages, 2006.

[178] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, "Accurate evolutions of orbiting black-hole binaries without excision," *Physical Review Letters*, vol. 96, no. 11, Article ID 111101, 4 pages, 2006.

[179] L. Rezzolla, E. Barausse, E. N. Dorband et al., "Final spin from the coalescence of two black holes," *Physical Review D*, vol. 78, no. 4, Article ID 044002, 6 pages, 2008.

[180] J. M. Bardeen and J. A. Petterson, "The lense-thirring effect and accretion disks around kerr black holes," *Astrophysical Journal Letters*, vol. 195, article L65, 1975.

[181] P. A. G. Scheuer and R. Feiler, "The realignment of a black hole misaligned with its accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 282, no. 1, pp. 291–294, 1996.

[182] P. Natarajan and J. E. Pringle, "The alignment of disk and black hole spins in active galactic nuclei," *Astrophysical Journal*, vol. 506, no. 2, pp. L97–L100, 1998.

[183] A. Perego, M. Dotti, M. Colpi, and M. Volonteri, "Mass and spin co-evolution during the alignment of a black hole in a warped accretion disc," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 4, pp. 2249–2263, 2009.

[184] M. Dotti, M. Volonteri, A. Perego, M. Colpi, M. Ruszkowski, and F. Haardt, "Dual black holes in merger remnants—II. Spin evolution and gravitational recoil," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 682–690, 2010.

[185] R. I. Davies, L. J. Tacconi, and R. Genzel, "The nuclear gasdynamics and star formation of NGC 7469," *Astrophysical Journal*, vol. 602, no. 1, pp. 148–161, 2004.

[186] R. I. Davies, L. J. Tacconi, and R. Genzel, "The nuclear gas dynamics and star formation of Markarian 231," *Astrophysical Journal*, vol. 613, no. 2, pp. 781–793, 2004.

[187] M. Sikora, Ł. Stawarz, and J. P. Lasota, "Radio loudness of active galactic nuclei: observational facts and theoretical implications," *Astrophysical Journal*, vol. 658, no. 2, pp. 815–828, 2007.

[188] A. Capetti and B. Balmaverde, "The host galaxy/AGN connection in nearby early-type galaxies a new view of the origin of the radio-quiet/radio-loud dichotomy?" *Astronomy and Astrophysics*, vol. 453, no. 1, pp. 27–37, 2006.

[189] D. Garofalo, "The spin dependence of the blandford-znajek effect," *Astrophysical Journal*, vol. 699, no. 1, pp. 400–408, 2009.

[190] D. Garofalo, "Signatures of black hole spin in galaxy evolution," *Astrophysical Journal Letters*, vol. 699, no. 1, part 2, pp. L52–L54, 2009.

[191] M. Dotti, M. Colpi, L. Maraschi, A. Perego, and M. Volonteri, "A path to radio-loudness through gas-poor galaxy mergers and the role of retrograde accretion," *Astronomical Society of the Pacific*, vol. 427, pp. 19–26, 2010.

[192] M. Volonteri, R. Salvaterra, and F. Haardt, "Constraints on the accretion history of massive black holes from faint X-Ray counts," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 1, pp. 121–127, 2006.

[193] R. Salvaterra, F. Haardt, and M. Volonteri, "Unresolved X-Ray background: clues on galactic nuclear activity at $z > 6$," *Monthly Notices of the Royal Astronomical Society*, vol. 374, no. 2, pp. 761–768, 2007.

[194] R. S. Nemmen, R. G. Bower, A. Babul, and T. Storchi-Bergmann, "Models for jet power in elliptical galaxies: a case for rapidly spinning black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 4, pp. 1652–1662, 2007.

[195] M. Volonteri and P. Natarajan, "Journey to the MBH-σ relation: the fate of low-mass black holes in the universe," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 4, pp. 1911–1918, 2009.

[196] M. Volonteri, G. Lodato, and P. Natarajan, "The evolution of massive black hole seeds," *Monthly Notices of the Royal Astronomical Society*, vol. 383, no. 3, pp. 1079–1088, 2008.

[197] M. Volonteri and M. C. Begelman, "Quasi-stars and the cosmic evolution of massive black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 409, no. 3, pp. 1022–1032, 2010.

[198] R. N. Manchester, "The parkes pulsar timing array project," in *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, C. Bassa, Z. Wang, A. Cumming, and V. M. Kaspi, Eds., vol. 983 of *American Institute of Physics Conference Series*, February 2008.

[199] G. H. Janssen, B. W. Stappers, M. Kramer, M. Purver, A. Jessner, and I. Cognard, "European pulsar timing array," in *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, C. Bassa, Z. Wang, A. Cumming, and V. M. Kaspi, Eds., vol. 983 of *American Institute of Physics Conference Series*, February 2008.

[200] F. Jenet, L. S. Finn, J. Lazio et al., "The North American nanohertz observatory for gravitational waves," *Astrophysics, NANOGrav White Paper*, 2009.

[201] J. Lazio, "The square kilometre array," in *Panoramic Radio Astronomy: Wide-field 1-2 GHz Research on Galaxy Evolution The Square Kilometre Array (PRA '09)*, September 2009.

[202] G. Hobbs, A. Archibald, Z. Arzoumanian et al., "The International Pulsar Timing Array project: using pulsars as a gravitational wave detector," *Classical and Quantum Gravity*, vol. 27, no. 8, Article ID 084013, 2010.

[203] M. G. Haehnelt, "Low-frequency gravitational waves from supermassive black-holes," *Monthly Notices of the Royal Astronomical Society*, vol. 269, article 199, 1994.

[204] S. A. Hughes, "Untangling the merger history of massive black holes with LISA," *Monthly Notices of the Royal Astronomical Society*, vol. 331, no. 3, pp. 805–816, 2002.

[205] J. S. B. Wyithe and A. Loeb, "Low-frequency gravitational waves from massive black hole binaries: predictions for LISA and pulsar timing arrays," *Astrophysical Journal*, vol. 590, no. 2, pp. 691–706, 2003.

[206] A. Sesana, F. Haardt, P. Madau, and M. Volonteri, "Low-frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies," *Astrophysical Journal*, vol. 611, no. 2, pp. 623–632, 2004.

[207] M. Enoki, K. T. Inoue, M. Nagashima, and N. Sugiyama, "Gravitational waves from supermassive black hole coalescence in a hierarchical galaxy formation model," *Astrophysical Journal*, vol. 615, no. 1, pp. 19–28, 2004.

[208] K. J. Rhook and J. S. B. Wyithe, "Realistic event rates for detection of supermassive black hole coalescence by LISA," *Monthly Notices of the Royal Astronomical Society*, vol. 361, no. 4, pp. 1145–1152, 2005.

[209] A. Sesana, F. Haardt, P. Madau, and M. Volonteri, "The gravitational wave signal from massive black hole binaries and its contribution to the LISA data stream," *Astrophysical Journal*, vol. 623, no. 1, pp. 23–30, 2005.

[210] A. Sesana, M. Volonteri, and F. Haardt, "The imprint of massive black hole formation models on the LISA data stream," *Monthly Notices of the Royal Astronomical Society*, vol. 377, no. 4, pp. 1711–1716, 2007.

[211] A. Sesana, J. Gair, E. Berti, and M. Volonteri, "Reconstructing the massive black hole cosmic history through gravitational waves," *Physical Review D*, vol. 83, no. 4, Article ID 044036, 26 pages, 2011.

[212] J. R. Gair, A. Sesana, E. Berti, and M. Volonteri, "Constraining properties of the black hole population using LISA," *Classical and Quantum Gravity*, vol. 28, no. 9, 2011.

[213] J. E. Plowman, R. W. Hellings, and S. Tsuruta, "Constraining the black hole mass spectrum with gravitational wave observations—II. Direct comparison of detailed models," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 1, pp. 333–352, 2011.

[214] A. Vecchio, "LISA observations of rapidly spinning massive black hole binary systems," *Physical Review D*, vol. 70, no. 4, Article ID 42001, 17 pages, 2004.

[215] R. N. Lang and S. A. Hughes, "Measuring coalescing massive binary black holes with gravitational waves: the impact of spin-induced precession," *Physical Review D*, vol. 74, no. 12, Article ID 122001, 22 pages, 2006.

[216] R. N. Lang and S. A. Hughes, "Localizing coalescing massive black hole binaries with gravitational waves," *Astrophysical Journal*, vol. 677, no. 2, Article ID 122001, 22 pages, 2008.

[217] S. T. McWilliams, R. N. Lang, J. G. Baker, and J. I. Thorpe, "Sky localization of complete inspiral-merger-ringdown signals for nonspinning massive black hole binaries," *Physical Review D*, vol. 84, no. 6, Article ID 064003, 9 pages, 2011.

[218] J. R. Gair, L. Barack, T. Creighton et al., "Event rate estimates for LISA extreme mass ratio capture sources," *Classical and Quantum Gravity*, vol. 21, no. 20, pp. S1595–S1606, 2004.

[219] P. Amaro-Seoane, J. R. Gair, M. Freitag et al., "Intermediate and extreme mass-ratio inspirals—astrophysics, science applications and detection using LISA," *Classical and Quantum Gravity*, vol. 24, no. 17, R113, 2007.

[220] L. Barack and C. Cutler, "Using LISA extreme-mass-ratio inspiral sources to test off-Kerr deviations in the geometry of massive black holes," *Physical Review D*, vol. 75, no. 4, Article ID 042003, 11 pages, 2007.

[221] J. R. Gair, C. Tang, and M. Volonteri, "LISA extreme-mass-ratio inspiral events as probes of the black hole mass function," *Physical Review D*, vol. 81, no. 10, Article ID 104014, 19 pages, 2010.

[222] L. Barack and C. Cutler, "LISA capture sources: approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy," *Physical Review D*, vol. 69, no. 8, Article ID 082005, 24 pages, 2004.

[223] M. V. Sazhin, "Opportunities for detecting ultralong gravitational waves," *Soviet Astronomy*, vol. 22, pp. 36–38, 1978.

[224] S. Detweiler, "Pulsar timing measurements and the search for gravitational waves," *Astrophysical Journal*, vol. 234, pp. 1100–1104, 1979.

[225] B. Bertotti, B. J. Carr, and M. J. Rees, "Limits from the timing of pulsars on the cosmic gravitational wave background," *Monthly Notices of the Royal Astronomical Society*, vol. 203, pp. 945–954, 1983.

[226] R. W. Hellings and G. S. Downs, "Upper limits on the isotropic gravitational radiation background from pulsar timing analysis," *Astrophysical Journal*, vol. 265, pp. L39–L42, 1983.

[227] F. A. Jenet, G. B. Hobbs, K. J. Lee, and R. N. Manchester, "Detecting the stochastic gravitational wave background using pulsar timing," *Astrophysical Journal*, vol. 625, no. 2, pp. L123–L126, 2005.

[228] A. Sesana, A. Vecchio, and C. N. Colacino, "The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 1, pp. 192–209, 2008.

[229] A. Sesana, A. Vecchio, and M. Volonteri, "Gravitational waves from resolvable massive black hole binary systems and observations with Pulsar Timing Arrays," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 4, pp. 2255–2265, 2009.

[230] B. Kocsis and A. Sesana, "Gas-driven massive black hole binaries: signatures in the nHz gravitational wave background," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 3, pp. 1467–1479, 2011.

[231] A. Sesana, "Self consistent model for the evolution of eccentric massive black hole binaries in stellar environments: implications for gravitational wave observations," *Astrophysical Journal*, vol. 719, no. 1, pp. 851–864, 2010.

[232] C. Roedig, M. Dotti, A. Sesana, J. Cuadra, and M. Colpi, "Limiting eccentricity of subparsec massive black hole binaries surrounded by self-gravitating gas discs," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 4, pp. 3033–3041, 2011.

[233] A. Sesana, C. Roedig, M. T. Reynolds, and M. Dotti, "Multimessenger astronomy with pulsar timing and X-Ray observations of massive black hole binaries," *Monthly Notices of the Royal Astronomical Society*, vol. 420, no. 1, pp. 860–877, 2012.

[234] T. Tanaka, K. Menou, and Z. Haiman, "Electromagnetic counterparts of supermassive black hole binaries resolved by pulsar timing arrays," http://arxiv.org/abs/1107.2937.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 783451, 10 pages
doi:10.1155/2012/783451

Review Article

The Circumnuclear Environment of IRAS 20551-4250: A Case Study of AGN/Starburst Connection for *JWST*

E. Sani[1] and E. Nardini[2]

[1] *INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy*
[2] *High Energy Astrophysics Division, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

Correspondence should be addressed to E. Sani, sani@arcetri.astro.it

Received 15 November 2011; Accepted 20 December 2011

Academic Editor: Antonis Georgakakis

We present a general review of the current knowledge of IRAS 20551-4250 and its circumnuclear environment. This Ultraluminous Infrared Galaxy is one of the most puzzling sources of its class in the nearby Universe: the near-IR spectrum is typical of a galaxy experiencing a very intense starburst, but a highly obscured active nucleus is identified beyond $\sim 5\,\mu$m and possibly dominates the mid-IR energy output of the system. At longer wavelengths star formation is again the main driver of the global spectral shape and features. We interpret all the available IR diagnostics in the framework of simultaneous black hole growth and star formation and discuss the key properties that make this source an ideal laboratory for the forthcoming *James Webb Space Telescope*.

1. Introduction

Two main physical processes characterize the nuclear regions of active galaxies: intense star formation at rates of $\sim 10^2$–$10^3\,M_\odot\,yr^{-1}$ (starburst, SB) and accretion on to a supermassive black hole (active galactic nucleus, AGN). The issue of SB and AGN connection in both local and distant galaxies is critical for a proper understanding of galaxy formation and evolution, of star formation history and metal enrichment of the Universe, and of the origin of the extragalactic background at low and high energies. There is indeed increasing evidence of a strong link between the starburst and AGN mechanisms in active systems. The empirical correlation between the mass of black holes (BHs) located at the centre of nearby galaxies (both active and passive/quiescent) and the mass of their spheroids (see [1] and references therein) suggests that the formation of bulges and the growth of the central BHs are tightly connected. Also the presence of circumnuclear star formation in a substantial fraction of local AGN [2–5] hints at the relation between the two phenomena. The overall conclusion of these studies is that in 30–50% of the cases the accreting supermassive BHs are associated with young (i.e., of age less than a few $\times 100$ Myr) star-forming regions, with clear evidence of an enhanced star formation rate (reaching up to starburst intensities) in most AGN. However, this does not necessarily imply any causal connection between the two physical processes. It could be simply the natural consequence of massive gas fuelling into the nuclear regions, due to either interactions/mergers or secular evolution such as bar-driven inflows. Both star formation and nuclear accretion, in fact, are triggered and subsequently fed by this gas reservoir.

In the local Universe, the optimal targets to study the AGN/SB interplay are the so-called Ultraluminous Infrared Galaxies (ULIRGs; [6]). These sources are the result of major mergers, during which the redistribution of the gaseous component drives vigorous starburst events and obscured nuclear accretion. It is now well established that ULIRGs are usually powered by a combination of both processes, giving rise to their huge luminosities ($L_{bol} \sim L_{IR} > 10^{12}\,L_\odot$). However, since the primary radiation field is reprocessed by dust, the identification of the dominant energy supply is often unclear. The simultaneous presence of star formation and AGN signatures in the mid-IR makes this a really favourable band to disentangle the AGN and SB components and explore their environment. In particular, (i) the available spectra of *bona fide* starburst-dominated and, respectively, unobscured AGN-dominated sources are widely different, and show little

dispersion within the separate classes [7–10]. This allowed us to reproduce the AGN/SB contributions with fixed templates, especially over the the 3–8 μm spectral interval. (ii) For a given bolometric luminosity, the mid-IR AGN emission is higher than that of a starburst by a factor that rapidly declines with wavelength, ranging from ~100 at 3-4 μm [7] to ~25 at 5–8 μm [11]. Such a large difference is due to the key contribution of the hot dust layers directly exposed to the AGN radiation field. Together with the relatively low dust extinction at these wavelengths, this allows the detection of an AGN even when it is heavily obscured and/or its total luminosity is small compared with the SB counterpart. Based on the previous points, we successfully fitted the observed ULIRG spectra with a two-component analytical model, with only two free parameters: the relative AGN/SB contribution and the optical depth of the screen-like obscuration (if any) affecting the compact AGN component.

To understand whether the link between star formation and nuclear activity is a matter of *nature* (i.e., feedback processes) or *nurture* (i.e., host environments), here we investigate the circumnuclear structure of IRAS 20551-4250, an ideal laboratory thanks to its unique physical properties (in terms of both relative AGN/SB contribution and AGN obscuration), and to the fairly large multiwavelength dataset available. The paper is organized as follows: in Section 2 we review the present knowledge of the mid-IR properties of IRAS 20551-4250. The dust extinction law and gas column density are dealt with in Section 3. A possible general picture and the feasibility of future observations with *James Webb Space Telescope* (*JWST*) are discussed in Section 4. In Section 5 we summarize our findings and draw the conclusions. Throughout this work we adopt a standard cosmology ($H_0 = 70$ km/s/Mpc, $\Omega_m = 0.3$, $\Omega_\lambda = 0.7$).

2. IRAS 20551-4250: General Properties

IRAS 20551-4250 is a nearby ($z = 0.043$) ULIRG lying in the luminosity range of IR quasars, with $L_{\rm IR} = 4.3 \times 10^{45}$ erg/s. It mostly lacks of targeted studies; none the less in literature there are several related measurements among the statistical analyses of the local ULIRG population. IRAS 20551-4250 is a merging system in a fairly advanced state (Figure 1, left panel), characterized by a single nucleus with a prominent tidal tail and a slightly disturbed core, likely caused by a minor merger or strong secular evolution effects. From the high-resolution near-IR data Haan et al. [12] ascribe the large ratio of nuclear excess to bulge luminosity (see also Figure 1, right panel) to the possible presence of an AGN with BH mass $\sim 4.4 \times 10^8\,M_\odot$. The spectral classification changes significantly with the observed waveband. It is optically classified as an H II region [13], while in the mid-IR it resembles an SB galaxy [2]. However, diagnostic methods exclusively based on emission lines, as the ones mentioned previously, suffer from limited extensibility to faint sources and fail in identifying the heavily absorbed AGN detected in the hard X-rays. Indeed, the hard X-rays emission of IRAS 20551-4250 is clearly dominated by an obscured AGN, with luminosity $L_{2-10\,\rm keV} \sim 7.0 \times 10^{42}$ erg s^{-1} and column density $N_H \sim 8 \times 10^{23}$ cm^{-2} [14]. According to all these pieces of observational evidence, the relative AGN contribution to the bolometric luminosity is uncertain, but probably highly significant, while the circumnuclear environment is still poorly characterized. The first quantitative determination of the AGN contribution to the mid-IR emission of IRAS 20551-4250 was obtained by Farrah et al. [15] thanks to a series of effective diagnostics based on fine-structure lines. Their analysis of *Spitzer*/IRS high-resolution spectra suggests a moderate AGN contribution, even though a peculiar geometry and/or extreme optical depth are responsible for the lack of typical AGN tracers (e.g., [Ne V], [O IV]).

2.1. L- and M-Band Spectroscopy. Risaliti et al. [16] obtained *L*-band observations of ULIRGs with 8 m class telescopes (*VLT* and *Subaru*). The resulting high-quality spectra have revealed the great power of *L*-band diagnostics in characterizing AGN and SB components inside ULIRGs. The main results of these studies are summarized in the following. (1) A large (~110 nm) equivalent width (EW) of the 3.3 μm polycyclic aromatic hydrocarbon (PAH) emission feature is typical of SB-dominated sources, while the strong radiation field of an AGN, extending up to the X-ray domain, partially or completely destroys the PAH carriers. (2) A strong ($\tau_{3.4} > 0.2$) absorption feature at 3.4 μm due to aliphatic hydrocarbon grains is an indicator of an obscured AGN; indeed, such a deep absorption requires the presence of a bright, point-like source behind a screen of dusty gas. (3) A steep continuum ($\Gamma > 3$ when describing the flux density as a power law $f_\nu \propto \lambda^\Gamma$) hints at the presence of a highly obscured AGN. Again, a large value of Γ implies the strong dust reddening of a compact source.

The *L*-band spectrum of IRAS 20551-4250 shows somewhat puzzling properties [7]: a strong 3.3 μm emission feature (EW $\simeq$ 90 nm) suggests a dominant starburst contribution. On the other hand, the steep observed slope ($\Gamma \sim 5$) and the detection of the 3.4 μm absorption feature point to the presence of a significant AGN affecting the continuum emission. Sani and coauthors [17] added the *M*-band (4-5 μm) data to better determine the continuum trend and analyse the broad CO absorption band near 4.65 μm. By combining the *L*- and *M*-band data (as shown in Figure 2), we estimated a very large AGN contribution at 3.5 μm, exceeding ~90% once corrected for extinction (see [7] for the analytical details). The observed AGN component, however, is heavily obscured and shows extreme dust reddening. The large optical depth ($\tau_L > 5$, assuming the extinction law of Draine [18]) is necessary to reconcile the apparently contradictory observational results, that is, the high equivalent width of the 3.3 μm PAH feature and the steep, intense continuum. The presence of a dust and gas screen absorbing the AGN emission is also revealed by the deep absorption profiles due to aliphatic hydrocarbons ($\tau_{3.4} = 1.5$) and gaseous CO ($\tau_{4.6} = 2.2$). This step-wise correlation between continuum reddening and absorption features appears to be a general property of ULIRGs hosting an obscured AGN [11, 17]. Anyhow, this does not hold under a quantitative point of view: no tight correlation is found among the values of the optical depth, not even between the two absorption features themselves. This suggests a nonuniform dust composition

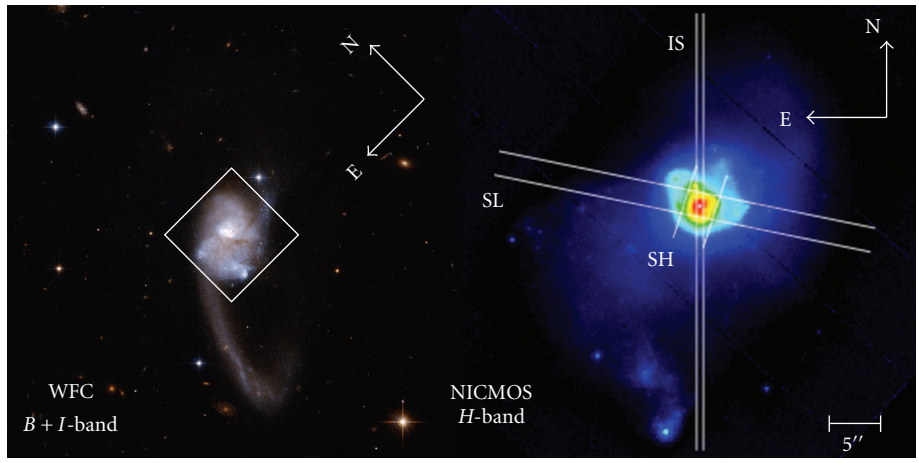


Figure 1: *HST* images of IRAS 20551-4250. Left: composite $B + I$ band image obtained with the WFC F435W and F814W filters (PID 10592, PI A. Evans). The white square identifies the nuclear (30″ × 30″) region. Right: NICMOS F160W image (PID 11235, PI J. Surace) of the nuclear region. The *Spitzer*/IRS short-wavelength/low-resolution (SL, 3.6″ width) and short-wavelength/high-resolution (SH, 4.7″ width) slits are shown together with the *VLT*/ISAAC slit (IS, 1″ width in the L-band). Given a spatial scale of 950 pc/″, the SL, SH, and IS slits cover regions of 3.4 kpc, 4.5 kpc, and 950 pc, respectively.

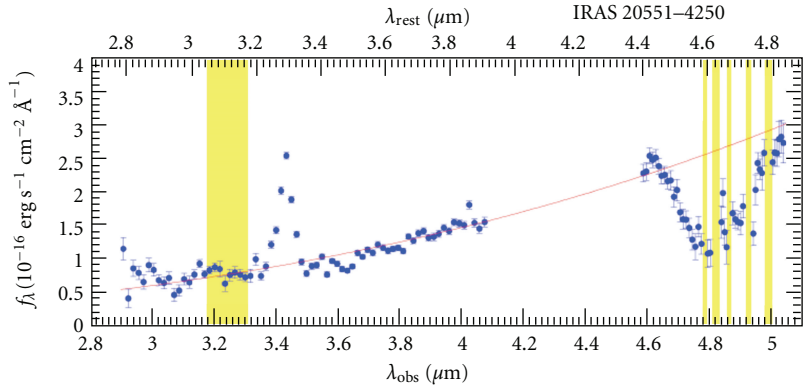


Figure 2: $L + M$ band of IRAS 20551-4250 [17]. The heavily reddened continuum (red curve) is present together with strong CO absorption in the M-band. The regions of bad atmospheric transmission are shaded in yellow. (Note the units on the vertical axis, $f_\lambda \propto \lambda^{-2} f_\nu$.)

among ULIRGs. The implications on the shape of the extinction law are discussed in the following section.

2.2. Spitzer/IRS Spectroscopy. In a series of papers [10, 11] we have shown that the high quality of *Spitzer*-IRS data allows a very effective *quantitative* determination of the AGN/SB components around 5–8 μm; this method is much more accurate than those possible in other bands in spite of the lower AGN over SB brightness ratio, which rapidly declines with wavelength. Summarizing, once applied to large, virtually complete samples of local ULIRGs, the 5–8 μm analysis yields the main results listed hereinafter. (1) The large variations in the observed spectral shape of ULIRGs can be successfully explained in terms of the relative AGN contribution and its degree of obscuration. (2) Although the larger fraction of ULIRG bolometric energy output is associated with the intense SB events, the AGN contribution is nonnegligible (~25–30%) and increases with both the total IR luminosity of the host galaxy and, possibly, with the merger stage [19]. (3) The apparent lack of continuum reddening and the simultaneous detection of deep absorption troughs in some of the most obscured sources (when a stepwise correlation is generally found, as mentioned earlier) suggests that the extinction of the AGN component in a ULIRG environment is not universal. Both a power-law and a quasigrey behaviour of the optical depth as a function of wavelength are necessary to account for the emission of different objects and seem to be involved among ULIRGs.

Consistently with the 3-4 μm analysis, also the 5–8 μm spectrum of IRAS 20551-4250 (in Figures 3 and 4) shows remarkable properties: the AGN continuum can be hardly determined due to strong absorption around 6 and 6.85 μm, respectively, attributed to a mixture of ices and hydrogenated amorphous carbons (HACs). The standard spectral decomposition yields again a very bright but strongly reddened AGN, with a mid-IR intrinsic contribution of ~90% and a 6 μm optical depth $\tau_6 = 1.2$ (following the same extinction law introduced before [18]). Although the starburst dominates the bolometric luminosity, the AGN contribution is significant (26 ± 3%). At longer wavelengths ($\lambda > 8\,\mu$m),

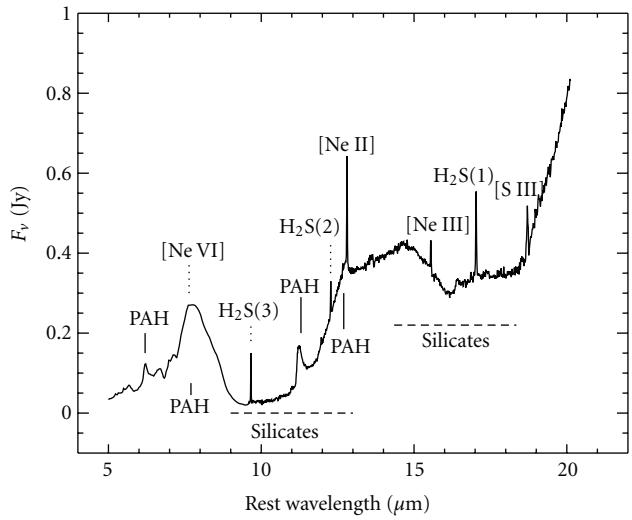


FIGURE 3: *Spitzer*/IRS 5–20 μm emission. We have already analysed the low-resolution data in a previous work [11], while the high-resolution spectrum (above ~10 μm) has been extracted from the same dataset following Schweitzer et al. [4]. The main features are labelled for ease of identification.

the huge silicate absorption troughs at 9.7 and 18 μm require the nuclear source to be deeply embedded in a smooth distribution of dust, both geometrically and optically thick. Ground-based imaging at 18 μm reveals a compact unresolved source (<120 pc) with high surface brightness and large Si optical depth ($\tau_{18} = 0.7$), in agreement with a buried AGN interpretation [20]. It is also worth noting that $\tau_{9.7}$ can be combined with the EW of the 6.2 μm PAH feature in a diagnostic diagram that provides not only a direct classification but also possible indications on the evolutionary path of a source, by probing the age of the SB and the geometrical structure of the dust [21]. The location of IRAS 20551-4250 in such a diagram is typical of an intermediate stage between a fully obscured AGN and an unobscured nuclear starburst.

As mentioned before, in IRAS 20551-4250 also fine-structure lines from highly ionized atoms are detected, as well as H_2 pure vibrational transitions (Figure 3). Our new measurements of the mid-IR line fluxes are listed in Table 1. Notably, the standard coronal lines produced by the hard AGN photons, such as [Ne V] (14.3 μm) and [O IV] (25.9 μm), are not detected (only upper limits are reported also in [15]); moreover, the [Ne III]/[Ne II] line ratio of ~0.2 is well consistent with an SB-dominated radiation field. As a result, taking into account only mid-IR emission lines would lead to a misclassification of IRAS 20551-4250 as a pure SB source. The lack of high-ionization lines and low [Ne III]/[Ne II] ratio can be actually reconciled with the presence of a deeply obscured AGN by allowing for a peculiar geometry of the gaseous/dusty absorber. Indeed, a large covering factor of the putative torus predicted by AGN unification models [22] can even prevent the formation of the narrow-line region and the production of high-ionization species. The geometrical properties of the absorber in a ULIRG are likely much more complicated, and a cocoon-like structure can be reasonably expected. Also the other standard diagnostic ratio [S III](λ18.71)/[S II](λ33.48) has an intermediate value among ULIRGs (~0.7) and tends to confirm the latter interpretation.

Four lines from pure rotational transitions of warm H_2 are clearly detected (see Figure 3, Table 1 and [15]): 0-0S(3) 9.67 μm, 0-0S(2) 12.28 μm, 0-0S(1) 17.04 μm and 0-0S(0) 28.22 μm. The upper levels of these transitions are populated via UV pumping, formation of H_2 in excited states or collisional excitation; therefore these lines directly probe the warm component of the molecular gas. A standard ortho-to-para ratio of 3 is found for gas with typical temperature $T \sim 300$ K. The heating mechanisms can be associated to either the SB (e.g., in photodissociation regions (PDRs), shocks/outflows in supernova remnants (SNRs)) or the AGN (due to the X-ray heating). From the line ratios and excitation temperatures measured among ULIRGs (including IRAS 20551-4250) Higdon et al. [23] ascribed the warm H_2 component to PDRs associated with massive SB. A more detailed investigation of the physical parameters of the H_2 gas is presented in Section 3.2.

3. The Circumnuclear Medium

The combined analysis of 3–8 μm data gives the immediate advantage to trace the coexisting AGN and SB environments. Indeed, after the review of all the mid-IR spectral properties, the presence of a heavily absorbed AGN combined with a vigorous SB in IRAS 20551-4250 is well established. None the less, a comprehensive interpretation of all the observables (AGN hot-dust emission, continuum reddening, absorption features, PAH strength) is not straightforward. The general picture is complicated by the different spatial extent of the nuclear region that has been explored in the works mentioned previously. In fact, there can be some aperture effects related to the slit widths, as the nuclear emission is quite diffuse and has a large surface brightness. The slit widths and orientations of the main instruments considered in this work are shown in the right panel of Figure 1. IRAS 20551-4250 presents a very small fraction (<10%) of extended emission in the 13.2 μm continuum, which can be mainly associated with the compact, unresolved hot/warm dust component in proximity of the AGN [24]. Conversely, the extra-nuclear emission is substantial for both the 7.7 μm PAH feature and the [Ne II] line at 12.8 μm (~40 and 25%, resp.; [25]), which are obviously related to the circumnuclear SB. Here, in order to further investigate the physical conditions responsible for reddening/absorption, we try (i) to fit simultaneously the *L*-band and 5–8 μm data, and (ii) to measure the column density of the circumnuclear gas for both the atomic and molecular components.

3.1. The Extinction Law. Figure 4 shows the observed spectrum of IRAS 20551-4250 between 3 and 8 μm once the ground-based *VLT* data are combined with the first part of the Short-Low *Spitzer*/IRS orders. We did not apply any cross-scaling factor, since it would be a very complex task and we are confident about the reliability of the absolute flux

Table 1: Properties of the detected mid-IR emission lines. We measured the fluxes with the IDL package SMART by means of a Gaussian fitting. We use the most recent version of the IRS pipelines; thus our estimates are more reliable than those previously published.

Mid-IR emission lines								
Line	$H_2S(3)$	$H_2S(2)$	[Ne II]	[Ne III]	$H_2S(1)$	[S III]	$H_2S(0)$	[S III]
λ_{rest} (μm)	9.662	12.275	12.814	15.555	17.030	18.713	28.219	33.481
E_{ion} (eV)	—	—	21.6	41.0	—	23.3	—	23.3
Flux (10^{21} W/cm^2)	5.36 ± 0.15	3.06 ± 0.17	13.0 ± 0.4	2.6 ± 0.3	6.9 ± 0.2	5.7 ± 0.6	3.8 ± 0.8	8.0 ± 0.3

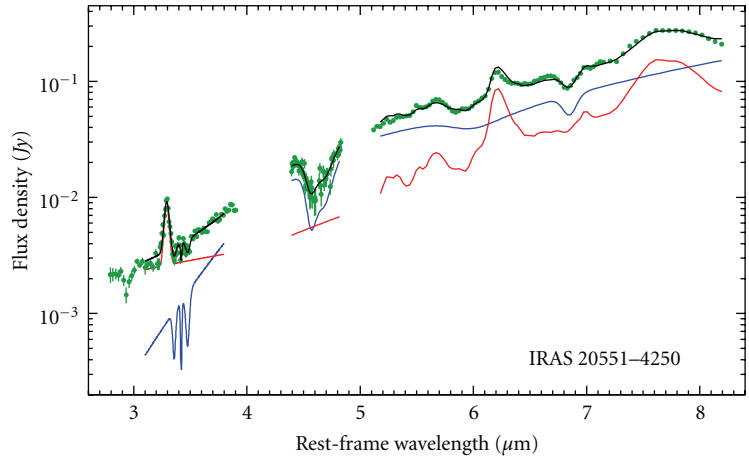


Figure 4: *VLT*/ISAAC *L*- and *M*-bands, plus *Spitzer*/IRS low-resolution spectra of IRAS 20551-4250. The data and best-fit models for each band are plotted in green and black, respectively. The different contributions to the observed spectra (blue for AGN component, red for the SB one) are also shown. Absorption features are fitted separately and then combined with the AGN template. Due to the strong reddening of the AGN continuum, the only AGN signature within the *L*-band is the shallow, structured 3.4 μm hydrocarbon absorption, highly diluted by the SB emission. The *M*-band continuum cannot be constrained due to the deep CO features, while the 6.0 μm water ice and 6.85 μm hydrocarbon absorption profiles are detected at 5–8 μm [9, 15].

calibrations, which are affected only by small relative errors (~10%; [7, 11, 17]). From a visual inspection of the three spectra, it is clear that the observed continuum slope, which is expected to be heavily shaped by the AGN contribution, cannot be reproduced with a single spectral index over the whole range under investigation. In our separate *L*-band and 5–8 μm studies, we have assumed an intrinsic slope of $\Gamma = 1.5$ for the AGN hot-dust continuum and then applied a power-law extinction of the form $\tau(\lambda) \propto \lambda^{-1.75}$ [18]. This screen-like absorption is possibly due to colder dust in the outer layers of the putative torus, or it might be associated with some star-formation region in the circumnuclear environment of the host galaxy. It is now evident that the latter assumptions do not allow us to reproduce simultaneously the AGN emission for the different datasets. In fact, by extending the best-fitting AGN model from the *L*-band to longer wavelengths, we largely overestimate the 8 μm observed flux. Of course, it is possible that the intrinsic AGN spectrum is more complex than the one adopted in our spectral decomposition. A more detailed analysis should allow for different dust components with individual temperature and emissivity, and also radiative transfer effects need to be taken into account. However, a broken power-law trend seems to describe with fairly good precision the observed spectral curvature. Interestingly, we can try to obtain some empirical (*a posteriori*) indication about the extinction suffered by the AGN hot-dust emission. Virtually all the available extinction curves in this wavelength range, in fact, are derived from lines of sight within our own Galaxy, while the composition of the interstellar medium (ISM) in active galaxies is expected to be very different, as proved, for example, by the dust-to-gas ratios estimated through a comparison between the mid-IR dust obscuration and the gas column density in the X-rays of these objects [26, 27].

We have therefore fitted all the three bands allowing for different slopes of the observed AGN continuum. The *M*-band is clearly poorly constrained and the value of Γ is frozen to give a smooth connection among the spectral intervals for both the AGN and SB templates. We have then computed the trend of the extinction law by making the easiest assumption about the intrinsic shape of the hot-dust emission, that is, the simple power-law dependence of the flux density from wavelength. Figure 5 shows the comparison between two possible extinction laws, corresponding to different values of the intrinsic Γ, and three standard Galactic curves. Although no conclusive indication can be drawn, the similarity is quite remarkable and suggests that the dust extinction law and the AGN intrinsic continuum are partially degenerate. This anyway does not affect the quantitative results of our analysis, as the AGN and SB 6 μm to bolometric

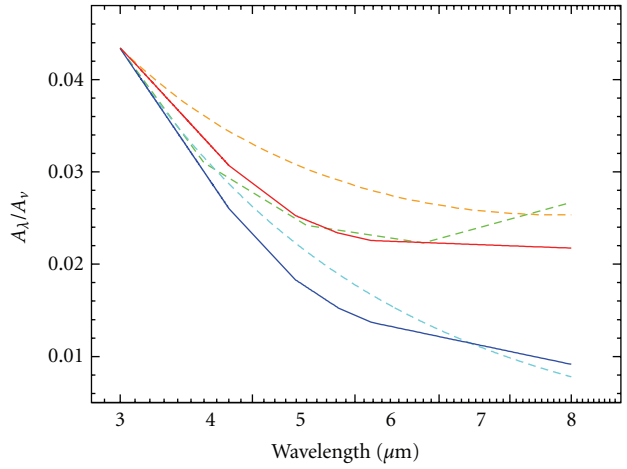


Figure 5: Dust extinction laws obtained by assuming a single power-law form ($f_\nu \propto \lambda^\Gamma$) to reproduce the intrinsic AGN hot-dust emission over the 3–8 μm band. The solid lines correspond to different choices of the spectral index: blue for $\Gamma = 2$, and red for $\Gamma = 3$. The dashed lines are standard Galactic extinction curves for comparison: cyan for [18], orange for Chiar and Tielens [28], and green for Nishiyama et al. [29, 30]. The relation with the extinction in visual magnitudes plotted on the vertical axis is based on the latter works, and all the curves are normalized in order to have the same value at 3 μm. While the exact extinction shape is not so important in the L-band, at 5–8 μm (and most likely beyond, with the presence of the silicate absorption feature) the difference in terms of optical depth in the different cases can be as large as a factor of 3. However, it seems quite hard to obtain such a large dust extinction around 6 μm to be consistent with the gas column density of nearly $\sim 10^{24}$ cm^{-2} measured in the X-rays.

corrections are averaged over large samples and this systematic effect is greatly reduced (see also the discussion on the AGN template and dust extinction in [11]).

3.2. Gas and Dust Content. To further constrain the absorbing/emitting medium in IRAS 20551-4250, we attempted at estimating the gas column density by means of a multiwavelength approach. We start by assuming a Galactic gas-to-dust ratio [31]:

$$\frac{N_H}{A_V} = 1.9 \times 10^{21}\ \text{mag}^{-1}\ \text{cm}^{-2} \quad (1)$$

with $A_L \sim 0.04\ A_V$ and $A_6 \sim 0.012\ A_V$. We then employ the following estimates: (a) the column density of the gas absorbing the X-ray radiation directly measured in the 2–10 keV energy range [14], (b) the L-band and 6 μm optical depth assessed through the continuum reddening in our decomposition method [11], and (c) the optical depth of the 3.4 μm hydrocarbon feature [7]. The corresponding visual extinction values are listed in Table 2.

From a comparison among these independent A_V predictions, we can draw four main considerations. (1) Independently from the adopted proxy, we infer a huge extinction in the visual band, which naturally explains the optical misclassification of IRAS 20551-4250. (2) As discussed in the previous section, a flatter extinction law over the 3–8 μm range with respect to a steep power-law trend [18] seems to be more appropriate to reproduce the observed AGN emission. Otherwise, the values of A_V derived from the 3 μm and 6 μm reddening differ by a factor of two. (3) By using the depth of the hydrocarbon feature to deabsorb the continuum, following the Galactic relation $A_L = (12 \pm 4)\tau_{3.4}$ [32], the resulting AGN intrinsic luminosity would exceed the source bolometric emission. The abundance of hydrocarbons dust grains is therefore higher in IRAS 20551-4250 than in the Galactic ISM. (4) The X-ray column density corresponds to an $A_V(X)$ at least a factor of two larger than that expected from our mid-IR modelling (τ_L, τ_6). Irrespectively of the actual dust extinction law, any reasonable value of the mid-IR optical depth implies a lower dust-to-gas ratio than in the Milky Way ISM. As a ULIRG is by definition a dust-rich system, this apparent inconsistency can be explained in two ways, which are in part complementary. (i) Due to orientation effects, our line of sight pierces through the regions of highest column density in the circumnuclear absorber. (ii) There is little coupling between the dust and gas components because the bulk of X-ray absorption occurs close to the central engine, in a region comparable in size with the dust sublimation radius.

Table 2: Extinction obtained by assuming a gas-to-dust ratio as in (1). Columns: (1) observational band used for the direct measurement with the relative reference. (2) and (3) L-band and 6 μm optical depth for the AGN continuum as obtained with our spectral decomposition model. (4) Optical depth of the aliphatic hydrocarbon feature once corrected for the SB dilution. (5) Gas column density measured from the X-ray spectrum. (6) Optical extinction predicted by Galactic relations (see text for details).

Extinction estimates					
Band	τ_L	τ_6	$\tau_{3.4}$	N_H cm^{-2}	Expected A_V mag
3 μm [9]	8				220
6 μm [15]		1.2			110
3.4 μm [9]			1.5		450
2–10 keV [11]				8×10^{23}	420

Another line of investigation into the physical properties of the circumnuclear medium relies on the rotation diagram of the warm molecular hydrogen, from which we can derive the temperature, column density, and mass of the gas. To this purpose, the observed fluxes listed in Table 1 are converted into column densities of the Jth state (N_j) assuming the LTE regime, an ortho-to-para ratio of 3, a point-like source, and no extinction [23] (see also [33]). While Higdon and collaborators [23] construct the rotation diagram for IRAS 20551-4250 using only the H_2 S(1) and S(3) transitions detected in the low-resolution mode, here we make use of high-resolution detections and add the S(0) and S(2) lines. In this way, the parameters derived from the linear fitting in Figure 6 are more reliable and accurate. Clearly a single temperature model applies to the S(1), S(2), and S(3) transitions, with the excitation temperature (T_{ex}) given by minus the reciprocal of the slope, while the total H_2 column density (N_{H_2}) depends on the fit normalization and the partition

functions of the populations. We thus obtain $T_{ex} = 347^{+5}_{-6}$ K, $N_{H_2} = (2.7 \times 10^{20})$ cm^{-2} and a corresponding H_2 mass of $M_{H_2} = 6.8 \times 10^8\,M_\odot$. (Uncertainties on T_{ex} and N_{H_2} are estimated by refitting the data while varying the S(1) and S(3) fluxes within their errors.) Our estimate gives a higher temperature (~8%) and correspondingly lower gas mass with respect to [23]. The inclusion of the S(0) line requires some caution, as it is detected with the IRS-LH slit, much larger (11.1″) than the SH one (4.7″) that samples the previous fluxes. For completeness, we plot in Figure 6 the observed S(0) value as a cross and the value corrected for the relative slit apertures SH/LH as a green point. Including also the corrected S(0) significantly steepens the linear regression and leads to a lower temperature $T_{ex} = 303$ K, hence doubling the column density and mass. As a matter of fact, a single-temperature component is not suitable to properly reproduce complex systems such as IRAS 20551-4250, and a multitemperature model should be adopted [23, 33]. Unfortunately the nondetection of higher level transitions (e.g., from S(4) to S(7)), or their blending with PAH features, prevents us from modelling a hot ($T \simeq 1000$ K) H_2 component. None the less, as an exercise, we can exclude the S(3) point and adopt the corrected S(0) in the linear regression. We now trace a colder H_2 component with $T_{ex} = 265$ K, characterized by a huge, likely unphysical (A mass for the molecular hydrogen larger than $\sim 10^9\,M_\odot$ would correspond to enormous star formation rates, with an IR radiation even greater than IRAS 20551-4250 bolometric luminosity.) gas mass ($M_{H_2} \sim 2 \times 10^9\,M_\odot$). We remind that ortho-$H_2$ exists only in states of odd rotational quantum number, while para-H_2 is represented only by states of even J; therefore the S(1)/S(3) line ratio is independent from the ortho-to-para ratio. The measured S(1)/S(3) is 1.29 ± 0.07, in agreement with the theoretical value of 1.23 computed for *no* extinction and $T_{ex} = 350$ K. From this, we conclude that the obscuring material along the line of sight producing the continuum reddening, deep features, and X-ray absorption lies in between the AGN and molecular H_2 clouds and is possibly associated with the SB region.

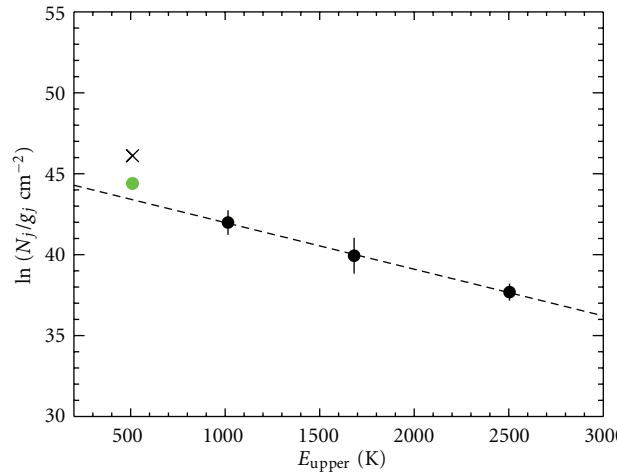


FIGURE 6: H_2 rotation diagram: for each transition the column density N_j normalized to the statistical weight of the state (g_j) is plotted as a function of the upper level energy [34]. The dashed line represents the best LTE fit obtained for S(1), S(2), and S(3) transitions (black points). For completeness we plot also the S(0) observation (black cross) and the measurement corrected for aperture effects (green point).

4. Discussion

We can now compile all the different aspects of the previous analysis in order to construct a comprehensive picture of the absorbing/emitting medium in IRAS 20551-4250. A stratified structure of the circumnuclear material, involving the different spatial scales (see Figure 1), can well explain all the observational evidence. The basic ingredients are summarized as follows: (i) the hot dust component, where the grains are transiently heated to temperatures close to the sublimation limit, can be associated with both the inner surface of the AGN torus and the starburst environment; however, due to the different spatial concentration of the hot dust in the two cases, the resulting nearly power-law continuum is much more intense for an AGN. (ii) A cold dust component and a large amount of gas are required to produce the continuum reddening, the deep absorption features (aliphatic hydrocarbon, CO, HAC, silicates), as well as the X-ray absorption. Consequently, the inferred properties of the circumnuclear absorber point to an optically thick screen along the line of sight towards a point-like source such as a bright AGN, rather than to a diffuse dust distribution spatially mixed with the energy source (as in a starburst). Moreover, this dust screen must be geometrically thick, since a large covering factor would be consistent with the absence of high-ionization coronal lines (e.g., [Ne V], [O IV]). These properties are typical of the AGN putative torus, which is located at the spatial scales ranging from a few pc to several tens of pc from the central engine. The obscuring medium is also expected to be sufficiently close to the central AGN (i.e., with the inner edge of the torus falling within the dust sublimation radius) to allow for the observed gas over-abundance. On farther scales (several hundreds to a few thousands pc), molecular clouds are associated with the starburst event. Here, in addition to warm thermal dust, the PAH grains can survive and give rise to the typical set of emission features usually employed as SF tracers. Furthermore, with the increasing optical depth within the individual star-forming clouds, photodissociation becomes eventually slow and inefficient, so that hydrogen also appears in the molecular state. This explains the unextincted H_2 pure rotational lines detected in the mid-IR. A cartoon of the circumnuclear environment is shown in Figure 7.

Of course, the qualitative considerations driven by the mid-IR spectral properties are not sufficient to fully understand the multiple physical conditions characterizing such an extreme source. In order to probe the nuclear environment and its surroundings, a detailed spectral analysis at different wavelengths is needed, possibly resolving and disentangling the different spatial scale. This would make it possible to address the problems connected to the uncertain shape of both the intrinsic and the observed AGN continuum, and therefore to better constrain the actual extinction law.

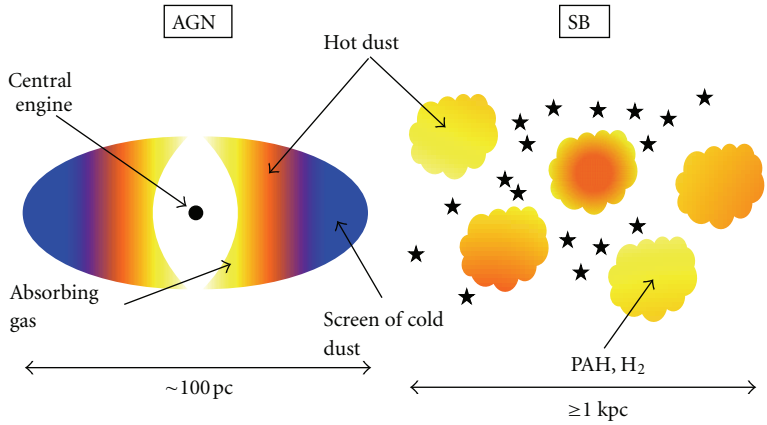


FIGURE 7: Schematic view of the possible spatial distribution of the absorbing/emitting medium in IRAS 20551-4250. The X-ray absorbing gas (in light yellow) is located within the dust sublimation radius and represents the inner edge of the axisymmetric (toroidal) absorber. The dusty regions emitting the thermal radiation are in orange (for both AGN torus and SB clouds). A compact, dense screen of cold dusty gas (in blue) forms the outer layer of the AGN absorber and is responsible for both continuum reddening and deep absorption troughs. PAH features and H_2 lines come from diffuse material spatially mixed with young stars (yellow/orange clouds). Due to the large spatial extent, only moderate internal obscuration is affecting the SB environment.

At present, even the joint modelling of the ~2–20 μm spectral energy distribution (SED) is frustrated by the spread of the signal-to-noise ratio (S/N) and the relative flux calibration among ground-based and space facilities involved in the observations. The forthcoming *James Webb Space Telescope* (*JWST*) is the ideal instrument to probe the mid-IR SED of local ULIRGs, offering the opportunity of high-quality data obtained with relatively short exposures. For example, a high-resolution ($R \sim 2700$) observation of IRAS 20551-4250 with NIRspec [35] centred at 3.5 μm requires only ~300 sec of exposure time (We used the exposure time calculator available at http://jwstetc.stsci.edu/etc/input/nirspec/spectroscopic/ with the following settings: G395H grating plus F290LP filter, average thermal background, and zodiacal light.) to reach a S/N ~ 150 per resolution element. At longer wavelengths, the medium-resolution spectrometer MIRI [36] will ensure similarly high performances. Besides the unique settings available (among which integral field unit and multi-shutter array), *JWST* will fully cover the ~1–25 μm range, allowing us to detect and resolve even faint and/or blended features. In this context, the separation of highly excited rotational levels of the CO $\nu = 1 - 0$ band would be particularly suitable to constrain the dense gas temperature, density, and kinematics within the circumnuclear environment (see, e.g., [37]).

5. Conclusions and Remarks

In the present work we have first reviewed the properties of IRAS 20551-4250, a prototypical local ULIRG observed by our group in the *L*- and *M*-band with ISAAC at the *VLT*. The spectral analysis also includes the 5–8 μm spectrum obtained by *Spitzer*/IRS. According to the AGN/SB decomposition method we have developed in several previous papers [7, 10, 17], IRAS 20551-4250 turns out to be a composite source, dominated in the mid-IR by hot dust emission associated with deeply embedded BH accretion and characterized by a vigorous circumnuclear starburst which provides the main power supply to the whole system. We have then interpreted the key spectral properties of the source over the ~3–20 μm wavelength range (e.g., the reddening of the continuum, the presence of deep absorption features, the lack of high-ionization coronal lines, and the detection of H_2 rotational transitions) in the framework of dust and gas spatial distribution and physical conditions. Our main results are the following. (i) The shape of the AGN intrinsic continuum is partly degenerate with the form of the extinction law. This is mainly evident beyond 5 μm. (ii) Given the gas amount inferred from X-ray observations, the central regions of IRAS 20551-4250 seem to have a dust-to-gas ratio much lower than the Galactic interstellar medium. (iii) Aliphatic hydrocarbon and HAC grains are over-abundant with respect to the local molecular clouds. (iv) A large covering of the nuclear engine likely prevents the ionization of the AGN narrow-line region and the excitation of fine-structure lines. Therefore, a screen of cold, dusty gas lies along the line of sight to the AGN, heavily extinguishing its spatially compact primary emission. (v) A large amount ($M_{H_2} = 6.8 \times 10^8\ M_\odot$) of warm ($T_{ex} = 347$ K) molecular hydrogen and PAH grains are associated with the starburst environment on typical scales of a few kpc. The findings have been *qualitatively* interpreted by means of a simple geometrical configuration as the one sketched in Figure 7. We have finally described the great improvement in terms of sensitivity, spectral coverage, and resolution that will be achieved in the near future with the advent of *JWST*. This will also allow us to separate the different spatial scales and explore in larger detail the connection between the AGN and SB environments and the mutual feedback between the two physical processes.

Acknowledgments

The authors thank the anonymous referee for the constructive comments and suggestions. E. Sani is grateful to Dr. F.

Fontani for precious discussions on the star forming environment. This work has made use of the NASA/IPAC extagalactic database (NED). E. Sani acknowledges financial support from ASI under Grant I/009/10/0/. E. Nardini acknowledges financial support from NASA Grants NNX11AG99G and GO0-11017X.

References

[1] E. Sani, A. Marconi, L. K. Hunt, and G. Risaliti, "The Spitzer/IRAC view of black hole-bulge scaling relations," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 2, pp. 1479–1494, 2011.

[2] R. Genzel, D. Lutz, E. Sturm et al., "What powers ultraluminous iras galaxies?" *The Astrophysical Journal*, vol. 498, no. 2, pp. 579–605, 1998.

[3] R. Cid Fernandes, Q. Gu, J. Melnick et al., "The star formation history of Seyfert 2 nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 355, no. 1, pp. 273–296, 2004.

[4] M. Schweitzer, D. Lutz, E. Sturm et al., "Spitzer quasar and ulirg evolution study (quest). I. The origin of the far-infrared continuum of QSOs," *The Astrophysical Journal*, vol. 649, no. 1, pp. 79–90, 2006.

[5] E. Sani, D. Lutz, G. Risaliti et al., "Enhanced star formation in narrow-line Seyfert 1 active galactic nuclei revealed by Spitzer," *Monthly Notices of the Royal Astronomical Society*, vol. 403, no. 3, pp. 1246–1260, 2010.

[6] D. B. Sanders and I. F. Mirabel, "Luminous infrared galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 34, no. 1, pp. 749–792, 1996.

[7] G. Risaliti, R. Maiolino, A. Marconi et al., "Unveiling the nature of ultraluminous infrared galaxies with 3–4 μm spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 303–320, 2006.

[8] B. R. Brandl, J. Bernard-Salas, H. W. W. Spoon et al., "The mid-infrared properties of starburst galaxies from Spitzer-IRS spectroscopy," *The Astrophysical Journal*, vol. 653, no. 2, pp. 1129–1144, 2006.

[9] H. Netzer, D. Lutz, M. Schweitzer et al., "Spitzer quasar and ulirg evolution study (quest). II. The spectral energy distributions of palomar-green quasars," *The Astrophysical Journal*, vol. 666, no. 2, pp. 806–816, 2007.

[10] E. Nardini, G. Risaliti, M. Salvati et al., "Spectral decomposition of starbursts and active galactic nuclei in 5–8 μm Spitzer-IRS spectra of local ultraluminous infrared galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 1, pp. L130–L134, 2008.

[11] E. Nardini, G. Risaliti, M. Salvati et al., "Exploring the active galactic nucleus and starburst content of local ultraluminous infrared galaxies through 5–8 μm spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 3, pp. 1373–1402, 2009.

[12] S. Haan, J. A. Surace, L. Armus et al., "The nuclear structure in nearby luminous infrared galaxies: hubble space telescope NICMOS imaging of the goals sample," *The Astronomical Journal*, vol. 141, no. 3, article 100, 2011.

[13] L. J. Kewley, C. A. Heisler, M. A. Dopita, and S. Lumsden, "Optical classification of southern warm infrared galaxies," *The Astrophysical Journal, Supplement Series*, vol. 132, no. 1, pp. 37–71, 2001.

[14] A. Franceschini, V. Braito, M. Persic et al., "An XMM—Newton hard X-ray survey of ultraluminous infrared galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 343, no. 4, pp. 1181–1194, 2003.

[15] D. Farrah, J. Bernard-Salas, H. W. W. Spoon et al., "High-resolution mid-infrared spectroscopy of ultraluminous infrared galaxies," *The Astrophysical Journal*, vol. 667, no. 1, pp. 149–169, 2007.

[16] G. Risaliti, M. Imanishi, and E. Sani, "A quantitative determination of the AGN content in local ULIRGs through L-band spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 1, pp. 197–203, 2010.

[17] E. Sani, G. Risaliti, M. Salvati et al., "3–5 μm Spectroscopy of obscured AGNs in ULIRGs," *The Astrophysical Journal*, vol. 675, no. 1, pp. 96–105, 2008.

[18] B.T. Draine, "Interstellar extinction in the infrared," in *Proceedings of the 22nd ESLAB Symposium on Infrared Spectroscopy in Astronomy*, vol. 290, pp. 93–98, 1989.

[19] E. Nardini, G. Risaliti, Y. Watabe, M. Salvati, and E. Sani, "The role of nuclear activity as the power source of ultraluminous infrared galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 405, no. 4, pp. 2505–2520, 2010.

[20] M. Imanishi, K. Imase, N. Oi, and K. Ichikawa, "Subaru and Gemini high spatial resolution infrared 18 μm imaging observations of nearby luminous infrared galaxies," *The Astronomical Journal*, vol. 141, article 156, 2011.

[21] H. W. W. Spoon, J. A. Marshall, J. R. Houck et al., "Mid-infrared galaxy classification based on silicate obscuration and pah equivalent width," *The Astrophysical Journal*, vol. 654, no. 1, pp. L49–L52, 2007.

[22] R. Antonucci, "Unified models for active galactic nuclei and quasars," *Annual Review of Astronomy and Astrophysics*, vol. 31, no. 1, pp. 473–521, 1993.

[23] S. J. U. Higdon, L. Armus, J. L. Higdon, B. T. Soifer, and H. W. W. Spoon, "A spitzer space telescope," *The Astrophysical Journal*, vol. 648, no. 1, pp. 323–339, 2006.

[24] T. Díaz-Santos, V. Charmandaris, L. Armus et al., "The spatial extent of (U)LIRGs in the mid-infrared. I. The continuum emission," *The Astrophysical Journal*, vol. 723, no. 2, pp. 993–1005, 2010.

[25] T. Díaz-Santos, V. Charmandaris, L. Armus et al., "The spatial extent of (U)Lirgs in the mid-infrared. II. Feature emission," *The Astrophysical Journal*, vol. 741, 32, 2011.

[26] R. Maiolino, A. Marconi, M. Salvati et al., "Dust in active nuclei I. Evidence for "anomalous" properties," *Astronomy and Astrophysics*, vol. 365, no. 2, pp. 28–36, 2001.

[27] E. Nardini and G. Risaliti, "Compton-thick active galactic nuclei inside local ultraluminous infrared galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 1, pp. 619–628, 2011.

[28] J. E. Chiar and A. G. G. M. Tielens, "Pixie dust: the silicate features in the diffuse interstellar medium," *The Astrophysical Journal*, vol. 637, no. 2, pp. 774–784, 2006.

[29] S. Nishiyama, T. Nagata, M. Tamura et al., "The interstellar extinction law toward the Galactic center. II. V, J, H, and K," *The Astrophysical Journal*, vol. 680, no. 2, pp. 1174–1179, 2008.

[30] S. Nishiyama, M. Tamura, and H. Hatano, "Interstellar extinction law toward the galactic center III: J, H, KS bands in the 2MASS and the MKO systems, and 3.6, 4.5, 5.8, 8.0 μm in the Spitzer/IRAC system," *The Astrophysical Journal*, vol. 696, article 1407, 2009.

[31] R. C. Bohlin, B. D. Savage, and J. F. Drake, "A survey of interstellar H I from L-alpha absorption measurements. II," *The Astrophysical Journal*, vol. 224, pp. 132–142, 1978.

[32] Y. J. Pendleton, S. A. Sandford, L. J. Allamandola, A. G. G. M. Tielens, and K. Sellgren, "Near-infrared absorption spectroscopy of interstellar hydrocarbon grains," *The Astrophysical Journal*, vol. 437, no. 2, pp. 683–696, 1994.

[33] S. Veilleux, D. S. N. Rupke, D. C. Kim et al., "Spitzer quasar and ulirg evolution study (quest). IV. Comparison of 1 Jy ultraluminous infrared galaxies with palomar-green quasars," *The Astrophysical Journal Supplement Series*, vol. 182, no. 2, pp. 628–666, 2009.

[34] B. Nisini, T. Giannini, D. A. Neufeld et al., "Spitzer spectral line mapping of protostellar outflows. II. H," *The Astrophysical Journal*, vol. 724, no. 1, pp. 69–79, 2010.

[35] G. Bagnasco, M. Kolm, P. Ferruit et al., "Overview of the Near Infrared Spectrograph (NIRSpec) instrument on-board the James Webb Space Telescope (JWST)," in *Cryogenic Optical Systems and Instrumentation Conference*, vol. 6692 of *Proceedings of SPIE*, August 2007.

[36] G. S. Wright, G. Reike, L. Colina et al., "The JWST MIRI instrument concept," in *Optical, Infrared, and Millimeter Space Telecopes*, vol. 5487 of *Proceedings of SPIE*, pp. 653–663, June 2004.

[37] M. Shirahata, T. Nakagawa, M. Goto et al., "Detection of warm molecular clouds toward the obscured AGN IRAS 08572+ 3915," in *The Central Engine of Active Galactic Nuclei*, L. C. Ho and J.-M. Wang, Eds., vol. 373, p. 505, 2007.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 178060, 14 pages
doi:10.1155/2012/178060

Research Article

M94 as a Unique Testbed for Black Hole Mass Estimates and AGN Activity at Low Luminosities

Anca Constantin[1] and Anil C. Seth[2]

[1] *Department of Physics and Astronomy, James Madison University, Harrisonburg, VA 22807, USA*
[2] *Department of Physics and Astronomy,University of Utah, Salt Lake City, UT 84112, USA*

Correspondence should be addressed to Anca Constantin, constaax@jmu.edu

Received 28 September 2011; Accepted 12 December 2011

Academic Editor: Antonis Georgakakis

We discuss the peculiar nature of the nucleus of M94 (NGC 4736) in the context of new measurements of the broad Hα emission from *HST*-STIS observations. We show that this component is unambiguously associated with the high-resolution X-ray, radio, and variable UV sources detected at the optical nucleus of this galaxy. These multiwavelength observations suggest that NGC 4736 is one of the least luminous broad-line (type 1) LINERs, with $L_{\rm bol} = 2.5 \times 10^{40}$ erg s^{-1}. This LINER galaxy has also possibly the least luminous broad-line region known ($L_{\rm H\alpha} = 2.2 \times 10^{37}$ erg s^{-1}). We compare black hole mass estimates of this system to the recently measured $\sim 7 \times 10^6 M_\odot$ dynamical black hole mass measurement. The fundamental plane and M-σ^* relationship roughly agree with the measured black hole mass, while other accretion-based estimates (the M-FWHM(Hα) relation, empirical correlation of BH mass with high-ionization mid-IR emission lines, and the X-ray excess variance) provide much lower estimates ($\sim 10^5 M_\odot$). An energy budget test shows that the AGN in this system may be deficient in ionizing radiation relative to the observed emission-line activity. This deficiency may result from source variability or the superposition of multiple sources including supernovae.

1. Introduction: LINERs and M94

Most high-mass galaxies are known to host massive black holes (BH), some passively lurking in their centers while others are actively accreting surrounding material (e.g., [1] and references therein). The mechanism that causes a BH's activity to turn on and off is still largely unknown. Understanding the structure of the active galactic nuclei (AGN) at their lowest luminosities is crucial to determining the physical and possibly evolutionary links between the most luminous galaxy centers and the passive ones. However, at low luminosities it is difficult to disentangle the various emission mechanisms that could be concurrently present in galaxy centers. As a consequence, the dominant power source of a large majority of actively line-emitting galaxies remains ambiguous (e.g., [2] for a review).

Diagnostic diagrams (e.g., [3–6]) are relatively successful in separating out bona fide accretion sources (Seyferts) from nuclei whose emission-line activity is mainly powered by young, hot stars (H II galaxies), based on emission line ratios. At least 50% of the strong line emitters fall easily onto the H II locus; however, only less than 10% are of the Seyfert type [5, 7, 8]. A large fraction of the objects "in between" these two categories, that exhibit relatively low levels of ionization (i.e., low values of [O III]/Hβ), maintain reasonably strong forbidden line activity (i.e., high values of [N II]/Hα) and are classified as low-ionization nuclear emission regions (LINERs). The reminder are usually called transition objects (Ts). Whether Ts and Ls are powered, at least partly, by accreting BHs, and thus could be called AGN, is a matter of continuous debate [2].

There are some typical emission characteristics that are considered to be particularly good indications that accretion onto a massive BH is an important, if not the dominant source of ionization in some Ls, and possibly Ts. The detection of broad Hα emission, regardless of its strength or luminosity, is generally considered to be *the* clue to AGN emission. Some LINERs (and maybe Ts as well) exhibit these features; however, the majority of these systems show only narrow emission, which could be generated by shocks,

poststarbursts, or other processes unrelated with accretion. Observations outside of the optical wavelengths often reveal AGN signatures in ambiguous and even starburst galaxy nuclei. X-rays are particularly good tracers of accretion; however, they are not efficient in distinguishing AGN at $L_X \lesssim 10^{42}\,\mathrm{erg\,s^{-1}}$, where contamination by X-ray binaries can be significant. X-rays are also unlikely to detect heavily absorbed AGN (i.e., Compton thick; $N_\mathrm{H} > 1.5 \times 10^{24}\,\mathrm{cm^{-2}}$). On the other hand, mid-IR high-ionization emission lines like [Ne V] $\lambda 14.32\,\mu$m, $24.32\,\mu$m (97.1 eV) appear to be a trustworthy indicator of AGN activity [9–11] due to the extreme conditions (i.e., very hard ionizing radiation) required to produce them; because these features have considerably lower optical depth, their detection can also reveal Compton-thick AGN ([12], e.g., NGC 1068). This technique has now been applied to reveal new and large numbers of optically unidentified AGN [13, 14], providing thus sensitive improvements on previous AGN censuses. These studies, along with X-ray and radio studies of nearby galaxies (e.g., [15–17]). suggest that a majority of LINERs and a large fraction of transition galaxies might in fact host accreting black holes. The presence of an accreting black hole does not however guarantee that the accretion power is the dominant source of ionization of those galaxy nuclei. A more recent assessment of the energy budget of LINERs by Eracleous et al. [18] argues that in 85% of LINERs the AGN ionizing photons are not sufficient for producing the observed nebular emission, and thus other power sources are likely to dominate.

Some new potentially powerful insights into the excitation mechanism of the low-luminosity AGN (LLAGN), and in particular the ambiguous sources, come from recent studies of large statistical samples of nearby galaxy nuclei, which reveal a potential H II → *S* → *T* → *LINER* → *Passive Galaxies* evolutionary sequence in the process of BH growth within galaxies [19, 20]. This sequence traces trends in (1) increasing host halo mass, (2) increasing environmental density, (3) increasing central BH mass and host stellar mass, (4) decreasing BH accretion rate, (5) aging of the stellar population associated with their nuclei, and (6) decreasing in the amount of dust obscuration, which might translate into a decrease in the amount of material available for star forming or accretion. In this picture, Seyferts and Ts are transition phases between the initial onset of accretion, usually swamped by the star-forming gas and associated dust, which is seen optically as an H II system, and the final phase of accretion observed as LINERs of already massive BHs. While this idea is supported by various other independent observational studies of low-luminosity AGN and starburst galaxies [21, 22], along with state-of-the-art hydrodynamical models for the life cycles of the most luminous AGN (e.g., [23, 24]) it is very probable that not *all* sources fit into this scenario. It is also important that this evolutionary sequence idea is tested on samples that span a narrow distribution in Hubble types; because most of the H IIs and Ts are relatively late-type disk galaxies that, likely, never experienced a (recent) major merger, it is possible that the trigger of such a sequence is different from the merger that initiates a similar life cycle at high luminosities. Nevertheless, investigating the challenges certain objects bring to this idea is useful for identifying and quantifying the caveats associated with this sequence. These objects may also be at interesting stages in their galaxy evolution.

NGC 4736 (or M94, UGC 7996) is a captivating example of an ambiguous galaxy nucleus, which poses challenges to the general understanding of AGN phenomena, including the above-mentioned sequence. This object is one of the closest ($d = 4.3$ Mpc; [25]) nearly face-on spiral, with a SAab Hubble Type. Its proximity enables study of details that would be unobservable in more distant systems. Its nucleus has a low-luminosity LINER spectrum, but has been also included in catalogs of transition objects [26] or Seyfert 2s (e.g., [27]). The AGN nature of this object has been constantly debated, an aging starburst being a compelling alternative [28–31]. The galaxy presents a ring of H II regions at a radius of $\sim 50''$, red arcs at $\sim 15''$, a high-surface-brightness nuclear region, and high far-infrared bulge emission [32, 33]. Its intricate structure of off-nuclear compact source detections in X-ray [30], radio [31], and UV [29], that do not necessarily match with each other, certainly increases the ambiguity associated with the nature of the main nuclear ionization mechanism. A common implied scenario in all of these studies is that this system is probably in the final stages of a merger.

We reexamine here NGC 4736 in the context of additional evidence for its AGN nature, which is the detection of a broad Hα component in its nuclear spectrum, as observed by the *Hubble Space Telescope* (*HST*) with the Space Telescope Imaging Spectrograph (STIS). We also gather multiwavelength data and show that source of the broad H-alpha emission line is coincident with a compact X-ray and radio source. These observations suggest that NGC 4736 hosts a broad-line region of significantly low luminosity, which makes this object one of the least luminous LINERs with strong evidence for BH accretion. Interestingly, BH mass indicators calibrated on rapidly accreting Seyfert galaxies give highly discrepant mass estimates in this more quiescent system. The $\sim 10^7\,M_\odot$ value given by the M-σ^* is two orders of magnitude higher than the values obtained via estimators based on the observed emission (X-ray variability, scaling relations, mid-IR emission) which, although physically independent of each other, give a consistent result of $\sim 10^5\,M_\odot$. We are thus facing the following conundrum: either (a) the standard AGN BH mass indicators do not necessarily apply to sources emitting in this low-luminosity regime as, probably, the emission mechanism is fundamentally different from that associated with higher-luminosity AGN, or (b) the emission signatures do not trace accretion onto the central BH. We discuss possible resolutions of this discrepancy in Section 5, where we propose some rather exotic scenarios.

2. Data Compilation and Analysis

NGC 4736 has been quite extensively observed across the whole electromagnetic spectrum. We present in this section the multiwavelength observations of this galaxy nucleus, in connection to new measurements of the broad Hα emission detected with *HST*-STIS.

2.1. The Broad Hα Emission Observed with HST-STIS. High-resolution optical spectra of the NGC 4736 nucleus were obtained with *HST*-STIS on July 2002. Data are publicly available (Prop ID 8591) but have not been published. The observations were obtained with the $52'' \times 0.''1$ aperture oriented at PA = 49.65°, with the slit centered along the major axis of the starlight distribution; two cosmic-ray split exposures were obtained, one being slightly shifted in the slit direction. The total combined exposure time is ~4000 s. The G750M grating was set at 6581 Å, with a scale of 0.″05/pixel, with no binning. We reduced the spectra using IRAF (IRAF is distributed by NOAO, which is operated by AURA Inc., under contract with the National Science Foundation) and the STIS reduction pipeline maintained by the Space Telescope Science Institute [34]. This reduction included image combination and cosmic-ray rejection, flux calibration, and correction of the wavelength to the heliocentric frame. To measure the nuclear nebular line-emission properties we extracted the 1-dimensional aperture spectrum five pixels wide (0.″25) centered on the continuum peak. The extracted spectrum thus consists of the central emission convolved with the STIS spatial point-spread function (PSF) and sampled over a rectangular aperture of $0.''25 \times 0.''1$. The measurements span 6295–6867 Å with a resolution of 0.87 Å ($\sigma_{\rm inst} = 17\,{\rm km\,s^{-1}}$).

Figure 1 shows the resulting spectrum together with the best matching continuum model, spectral fits of the pure emission-line component, and the associated residuals. The continuum model of the underlying stellar population is obtained via a χ^2 minimization of a nonnegative least-squares fit between the observed spectrum and a sum of discrete star bursts of different ages, adopted from the Bruzual and Charlot [35] stellar population synthesis templates, together with dust attenuation modeled as an additional free parameter. The continuum fitting is performed using an adaptation to our data of Christy Tremonti's code [36]. The pure emission-line spectrum, obtained by subtracting the modeled continuum from the observed spectrum, is fit by a combination of linear continuum and Gaussian components. In addition to narrow emission, the Hα + [N II] feature shows clear evidence for a broader feature. A flux ratio of 1 : 3 was assumed for the [N II] doublet, as dictated by the branching ratio [37]; the [O I] feature and the [N II] and [S II] doublets were assumed to share common velocity (red) shifts and widths. The best-fitting Gaussian parameters were derived via an interactive χ^2 minimization, using SPECFIT [38]. Because of the generally low signal to noise of the 2-d spectrum we are not able to test whether our measurements of the broad Hα feature could be corrupted by a possible rotating disk as in the case of M84 (e.g., [39]); the match in the widths of the [S II] and [N II] narrow features (as long as that of the narrow Hα) argue however against such a significant effect.

Measurements related to the broad Hα feature from the STIS measurements are listed in Table 1. The fractional contribution of this broad component to the total flux of the Hα + [N II] blend is 80%. Compared to other galaxies with broad-line emission in the Palomar survey [40], the width of the broad Hα is typical ($1570\,{\rm km\,s^{-1}}$). However, at just $9.1 \times 10^{37}\,{\rm ergs\,s^{-1}}$, the Hα luminosity is lower than any known broad-line sources in the Palomar sample or the Sloan Digital Sky Survey [40, 41]. This includes the famous low luminosity Seyfert 1 galaxy, NGC 4395, which has a luminosity of $1.2 \times 10^{38}\,{\rm ergs\,s^{-1}}$ [42] and is at a similar distance to NGC 4736. Note also that the broad Hα feature is redshifted by $\sim 140 \pm 20\,{\rm km\,s^{-1}}$ relative to the narrow

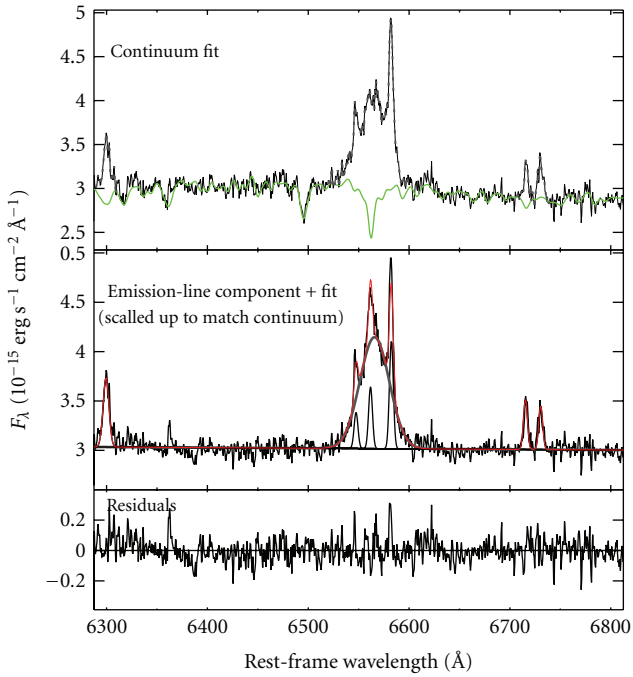


Figure 1: Spectral fits to the *HST*-STIS spectrum of NGC 4736, based on a 5-pixel extraction aperture, which corresponds to a total emission area of 0.025 arcsec^2. *Upper panel*: The thin continuous line shows the observed data, while the green line shows the best-fit stellar population model obtained by clipping out the hashed emission-line region. *Middle panel*: The continuum subtracted spectrum, that is, the emission-line component, together with the corresponding spectral fit, that have been shifted by an additive constant to match the initial continuum level, for illustration purposes. The thin continuous lines show the individual Gaussian components (1 per narrow line), the thick grey continuous line represents the broad Hα feature, and the red line is the final fit to the observed spectrum. *Lower panel*: The residuals after the subtraction of the model fit.

Table 1: Broad Hα emission measurements.

Quantity	Value	Notes
$f_{\rm blend}$	0.80	Fraction of broad Hα to Hα + [N II] blend.
$f_{\rm H\alpha}$	0.93	Fraction of broad Hα to total Hα
FWHM(H$\alpha^{\rm broad}$)	1570	±110; in km s^{-1}
Δv	140	±20; in km s^{-1}; broad relative to narrow Hα
$\log F$(H$\alpha^{\rm broad}$)	−13.38	Observed flux (erg s^{-1}cm^{-2})
$\log L$(H$\alpha^{\rm broad}$)	37.96	Observed luminosity (erg s^{-1})[a]

[a]assuming distance $d = 4.3$ Mpc.

emission, that is locked at the systemic velocity of the object; velocity offsets between the broad Hα and narrow lines are not unusual [43].

Evidence for broad Hα emission has also been presented in a recent PCA tomography study applied to this nucleus [44]. The observations reported in this case come from the Gemini Multi Object Spectrograph (GMOS)-IFU data cube and have been obtained 4 years after the *HST*-STIS spectrum had been acquired. Thus, the broad Hα line associated with the nucleus of M94 appears to be persistent for at least 4 years. They derive a broad Hα luminosity of 6×10^{38} ergs s^{-1}, $\gtrsim$6 times brighter in the GMOS data than the one we detect in the STIS spectrum. The difference could result from true variability in the broad-line luminosity or could be due to differences in aperture size or placement. We note that Steiner et al. [44] suggest the broad-line region is offset from the photocenter of the galaxy by 0.″15, while our spectral 0.″25 × 0.″1 aperture coincides with the photocenter. Even with their higher Hα luminosity, this source is among the least luminous known broad-line AGN.

2.2. Ground-Based Optical Spectroscopy and the Spectral Classification. Ground-based optical spectra of the NGC 4736 nucleus are available from the Palomar spectroscopic survey by Ho et al. [43] and the integrated spectrophotometric survey of Moustakas and Kennicutt [45] conducted with the 2.3 m Bok telescope, as well as from new data we acquired on February 2008 at the MMT Observatory. The MMT spectrum is obtained with the Blue Chanel Spectrograph, the 500 grooves/mm grating used in first order with the 1″ slit, and covers $\lambda\lambda$3800–7000 with 3.6 Å resolution. None of these data show indications of broad Hα emission. This outcome is not surprising given the feature's flux in the *HST* spectrum, which would be very difficult to discern in the 1″ or 2″ apertures employed in these ground-based observations, which are at least one order of magnitude larger than that used in the *HST*-STIS observations. Through a 2″ aperture (∼40 pc), NGC 4736's emission complex Hα + [N II] is generally heavily swamped by the host stellar light.

Probably as expected, measurements of NGC 4736's narrow-line emission are consistent with this system being powered at least partially by a nonstellar source, regardless of the resolution of its observation. Figure 2 shows the location of this nucleus in the 3-dimensional diagnostic diagram usually employed in classifying emission-line galaxies (e.g., [3, 4, 6]), for different sets of data mapping ∼10 to ∼100 pc. As with the *HST*-STIS observations, the ground-based measurements are performed after the host stellar contribution has been subtracted from the observed spectrum by means of absorption galaxy template fits. Our MMT spectrum offers only an upper limit for the [O III]/Hβ while the Palomar line flux measurements are at least 50% uncertain. We list the measurements of the nebular emission of this nucleus as observed by Palomar, Bok, MMTO, and *HST*-STIS in Table 2. Because measurements of the [O III]/Hβ ratio are not available in the *HST* spectrum, we show the high-resolution measurements in Figure 2 using the [O III]/Hβ from the Palomar catalog; it is readily apparent that the high-resolution data are consistent with a LINER or a Seyfert classification for this object and reveal its type 1 (broad-lined) character.

Table 2: Emission line measurements[a].

line name	*HST*-STIS	Palomar[b]	Bok[c]	MMTO[d]
[O I] λ6300	5.03 ± 0.31	6.03	2.8 ± 0.9	7.0 ± 1.0
Hα (narrow)	3.00 ± 0.32	> 25.12	13.5 ± 1.9	22.0 ± 2.4
[N II] λ6583	5.25 ± 0.25	54.01	21.5 ± 1.5	39.3 ± 2.9
[S II] λ6716	2.48 ± 0.18	18.59	8.3 ± 1.1	15.5 ± 0.3
[S II] λ6731	2.16 ± 0.18	16.33	7.2 ± 1.1	9.8 ± 0.3

[a]All fluxes are in units of 10^{-15} erg s^{-1} cm^{-2} and represent the observed values, not corrected for reddening.
[b]nonphotometric conditions; line ratios are at least 50% uncertain.
[c]2.″5 slit, 2.3 m telescope; photometric conditions [45].
[d]1″ slit; fluxes are in units of 10^{-15} erg s^{-1} cm^{-2}.

2.3. The Radio, UV, and X-Ray Observations and the Astrometric Coincidence with the Broad Hα *Emission.* Across the electromagnetic spectrum, the center of this galaxy exhibits a complex morphology. There is a plethora of bright radio, UV, and X-ray sources detected in the center of NGC 4736. There is a nuclear compact (15 GHz) radio source measured by Nagar et al. [46] that appears to be associated with that of the brightest of the two close (8.49 GHz) radio sources revealed by Körding et al. [31], and with that of the brightest of the two ultraviolet point sources, which varies on a 10-year time scale [29]. *Chandra* observations [30] reveal numerous discrete X-ray sources in the inner galaxy; the second brightest X-ray source, X2 ($L_{X,2\text{–}10\,\mathrm{keV}} = 5.9 \times 10^{38}$ erg s^{-1}), coincides within the errors with the nucleus position. The off-nuclear radio, UV, and X-ray sources are apparently unrelated to each other.

The relation to possible optical counterparts of these observations is not well constrained in the literature. We thus investigated this issue, with a particular interest in the degree to which our newly detected broad Hα coincides in position with the multi-λ detections. We used for this purpose a variety of archival images from the *HST*: WFPC2/PC data in the F555W filter data (PID: 5741 and 10402), HRC data in the F250W and F330W filters (PID: 9454), and NICMOS/NIC3 data in the F160W filter (PID: 9360). All data were downloaded from the *HST* archive, and images were drizzled together when required. Absolute astrometry was performed on these data as well as the STIS observations (taken with the slit out) using a ground-based V-band image obtained from the *Spitzer Infrared Nearby Galaxies Survey* (SINGS) ancillary data by Kennicutt et al. [48]. These SINGS observations were aligned to the USNO-B system using ∼80 stars, and then astrometry of the *HST* data was obtained by degrading the resolution of the F555W image to match the SINGS image. All other *HST* images were then matched with the astrometry-corrected F555W frame to an accuracy of < 0.″05. The absolute error on this astrometry is about 0.″2 and is dominated by scatter of stars in the SINGS image relative to the USNO-B positions.

Figure 3 illustrates the result of this data compilation and the corresponding radio, UV, and X-ray source matches. *It is*

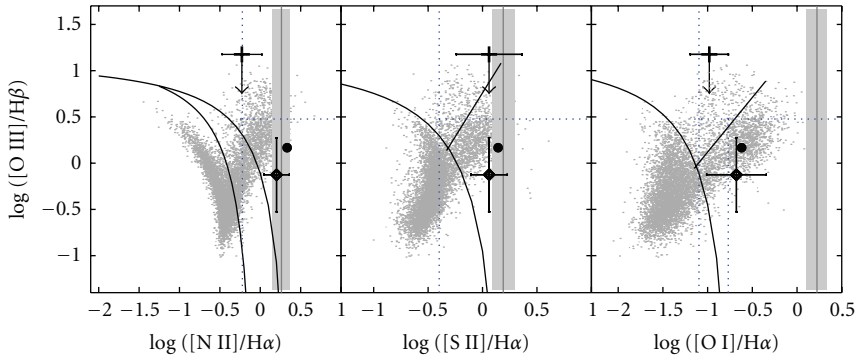


Figure 2: NGC 4736 within emission-line diagnostic diagrams. Filled circles reflect measurements based on the Palomar observations. Diamonds represent measurements from Moustakas and Kennicutt Jr. [45], while crosses show measurements from our MMT spectrum. The vertical bands indicate the *HST*-STIS measurements and the associated errors (with no correction for reddening; [O III] and Hβ are not available in the *HST* spectra). The solid (black) curves indicate the Kewley et al. [6] classification, while the dotted (blue) lines indicate criteria used by Ho et al. [43]. The background grey points correspond to measurements of SDSS nearby galaxies from Constantin et al. [19].

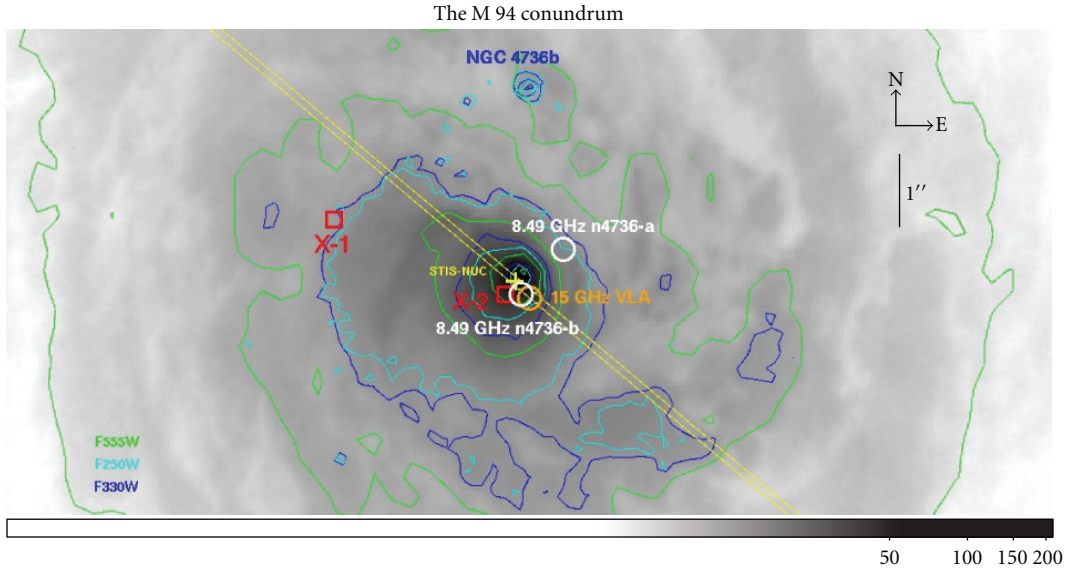


Figure 3: The mosaic-ed F555W (V-band; combination of 5 WFPC2 images taken at two different epochs) image of the NGC4736 nucleus, with intensity contours overlaid (green). The position of the STIS-nucleus is indicated (by the yellow slit) along with the locations of the 2 most luminous and closest to the nucleus hard X-ray sources, X1 and X2, as detected with Chandra (red squares), along with the position of the 15 GHz VLA radio core ([46], orange circle), and those of the two radio compact sources detected by Körding et al. [31] (white circles). The blue and cyan contours correspond to the F330W and F250W observations from Maoz et al. [29]; note the presence of NGC 4736b, $\approx 2.''5$ to the north of the nucleus. The absolute astrometry is good to within $0.''2$ (pixel size is $0.''05$ for F555W, and $0.''025$ for F250W and F330W).

clear that, within $< 0.''2$ *(4 pc), there is an obvious astrometric match in the nuclear X-ray, UV, optical, radio compact sources, and the newly detected broad* Hα *emission line.* The nucleus position as observed by STIS has RA: 12 h 50 m 53 s.20, DEC: 41°07′13.″40. The off-nuclear X1, that is the brightest compact X-ray source detected in this galaxy nucleus, has no counterpart at other wavelengths. Same is true for the off-nuclear 8.49 GHz detection, that is only $1''$ ($\sim$20 pc) away from the nucleus. The off-nuclear UV source is the only one detected in optical light.

3. NGC 4736's Nuclear Emission across the Electromagnetic Spectrum and L_{Bol}

With the wealth of data available for this galaxy nucleus, we are able to build its least contaminated X-ray to radio nuclear spectral energy distribution (SED). The multiwavelength observations of the sources detected in the very central regions of NGC 4736, plotted in $\nu L\nu$ units, are displayed in Figure 4. The X-ray detection X2 is represented here by a power law, estimated based on its photon indices $\Gamma = 1.6$,

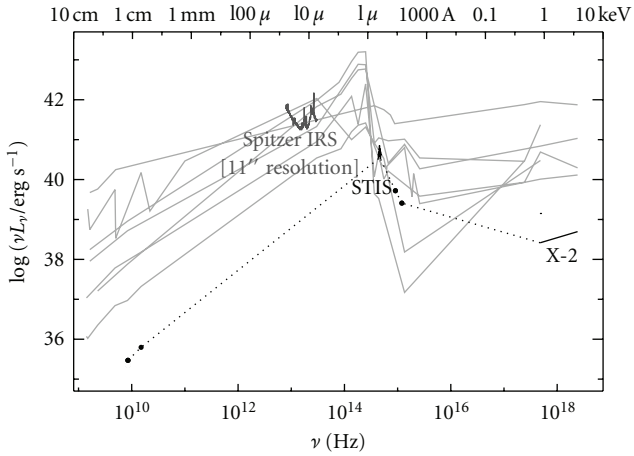


Figure 4: Radio to X-ray nuclear SED of NGC 4736 superposed on SEDs of LLAGN from Ho [47]. The optical STIS spectrum stands out as an amorphous blob featuring the strong Hα emission feature. The 2–10 keV Chandra X-ray detection X2 is depicted as a power-law corresponding to $\Gamma = 1.6$. The low spatial resolution Spitzer IRS observations are shown for comparison.

where the absorbing column density is fixed at the Galactic value. Both F250W and F330W UV observations are plotted, only for the nuclear detections. The optical data are represented by the *HST*-STIS spectrum, featuring the strong Hα emission line. Measurements corresponding to all the radio observations discussed above are indicated. For the sake of completeness, we also include in this plot lower spatial-resolution observations from SINGS IRS [48] as well. We show all of these measurements superposed on the SEDs of LLAGN from Ho [47]. No artificial normalization has been performed, and there is no correction for absorption in either the Galactic extinction (except for the X-ray data) or intrinsic to NGC 4736.

With X2 and n4736-b as the X-ray and the 8.5 GHz counterparts of the broad Hα detection, respectively and assuming that the continuum could be described as simple power laws between the points we present data for, this object's nuclear SED corresponds to $L_{\rm bol} \approx 2.5 \times 10^{40}\,{\rm erg\,s^{-1}}$. Spitzer data is not included in estimating $L_{\rm bol}$ because they do not reflect the IR emission of the uncontaminated nucleus; the aperture used in these observations is very large (see Table 3), including the entire field of view shown in Figure 3. This luminosity makes NGC 4736 one of the least luminous LINERs with strong evidence of BH accretion. This source is thus a critical signpost of BH accretion at extremely low levels. Note that this object's weak emission is most likely not caused by obscuration; Ho et al. [43] list a Balmer decrement of Hα/H$\beta = 3.1$ (albeit highly uncertain, with a probable error of $\pm 100\%$), and Eracleous et al. [30] provide an upper limit for the neutral hydrogen column density of $N_H < 3.3 \times 10^{20}\,{\rm cm^{-2}}$.

3.1. Comparison with Other Observed LLAGN. Figure 4 shows that NGC 4736's SED is very similar to those of other LLAGN [47]. This suggests that, in spite of the significant difference in $L_{\rm bol}$, there is no fundamental transition in the accretion mode in this source compared to other LLAGN.

Table 3: Nuclear SED data.

log(ν/Hz)	log(νL_ν/erg/s)	Resolution	Obs. date	Instr.
18.38	38.69	0.″15	5/2000	Chandra-ACIS
17.68	38.42	0.″15	5/2000	Chandra-ACIS
15.08	39.41	0.″5	6/2003	ACS F250W
14.96	39.72	0.″5	6/2003	ACS F330W
14.68	39.91	0.″1	6/2002	*HST*-STIS
14.64	39.93	0.″1	6/2002	*HST*-STIS
13.48	40.58	4.″7	5/2004	IRS-SH
13.19	40.33	11.″1	5/2004	IRS-SH/LH
12.91	40.87	11.″1	5/2004	IRS-LH
10.18	35.79	0.″15	1/2001	VLA, A config.
9.93	35.47	0.″24	6–10/2003	VLA, A config.

The similarity of the SED means that previously proposed bolometric corrections for LLAGN appear to work quite well for NGC 4736. The value estimated based on the correction to the [O III] line luminosity that Heckman et al. [49] proposed to work well for Seyferts is $L_{\rm bol} \approx 3500 \times L_{[\rm O\,iii]} = 4.5 \times 10^{40}\,{\rm erg\,s^{-1}}$ (here $L_{[\rm O\,III]}$ is measured from the ground within the Palomar survey and is not corrected for intrinsic reddening). If the more recent assessment of the $L_{\rm bol} = 600 \times L_{[\rm O\,III]}$ bolometric correction of kauffmann and Heckman [50] is used, with $L_{[\rm O\,III]}$ corrected for reddening based on the Balmer decrement listed above and a $\tau \propto \lambda^{-0.7}$ attenuation law [51], then $L_{\rm bol} = 0.93 \times 10^{40}\,{\rm erg\,s^{-1}}$, which compares satisfactorily with the measured $L_{\rm bol}$. The average bolometric correction to the observed 2–10 keV X-ray luminosity proposed by Ho [2], with $L/L_X = 16$, results in a somewhat lower value, $0.94 \times 10^{40}\,{\rm erg\,s^{-1}}$; with the more recent $L/L_X = 50$ average bolometric correction to the 2–10 keV luminosity of Eracleous et al. [52], $L_{\rm bol} = 1.8 \times 10^{40}\,{\rm erg\,s^{-1}}$ and, thus, very close to our integrated value. Within a typical uncertainty of ~50%, all of these estimates are consistent with the measured $L_{\rm bol}$ value. Other, more uncertain bolometric indicators are also consistent with our integration: the correlation between the mid-IR [Ne V] emission-line luminosity and bolometric luminosity derived by Satyapal et al. [53] for a small sample of much brighter nearby AGN, that has a large scatter (~1 order of magnitude), gives a $L_{\rm bol} = 1.6 \times 10^{40}\,{\rm erg\,s^{-1}}$, which is very close to our measured value.

This good match among quite a variety of bolometric estimators derived independently from the multiwavelength properties of AGN supports the AGN interpretation for this system. However, there are two details of its emission spectrum which are unlike typical broad-lined LLAGN.

First, despite the fact that the broad Hα emission of NGC 4736 is one of the weakest measured among type 1 AGN, the total Hα emission is peculiarly strong relative to the X-ray counterpart. The ratio $L_X(2\text{–}10\,{\rm keV})/L_{\rm H\alpha}$ is ~6, and thus *lower* than the median of ~15 exhibited by the type 1 AGN (and low-z quasars) included in, for example, Ho et al. [54] study that revealed a relatively tight correlation

between the two types of emission. With the higher broad Hα flux from Steiner et al. [44], the ratio is even lower, $L_X(2\text{–}10\,\mathrm{keV})/L_{\mathrm{H}\alpha} \approx 1$. This finding is surprising given that the low-luminosity nearby galaxies that deviate from the $L_X - L_{\mathrm{H}\alpha}$ linear scaling toward lower values of the $L_X/L_{\mathrm{H}\alpha}$ ratios are the type 2 sources, mostly transition objects, where the ionization mechanism is not necessarily dominated by an AGN type of source. Following, for example, Ho et al. [54] arguments, the unusually low $L_X/L_{\mathrm{H}\alpha}$ ratio measured in the nucleus of NGC 4736 suggests that either: (1) the optical line emission is not powered exclusively by a central AGN, or (2) the X-ray emission in this system arises, at least partially, from a non-AGN source, for example, an X-ray binary. A final possibility is that the source is highly variable (as suggested by the UV, X-ray, and Hα observations), and thus, the unusual $L_X/L_{\mathrm{H}\alpha}$ is the result of measurements made at different times.

Second, NGC 4736's nebular emission shows a number of peculiarities. To start with, the electron density in NGC 4736's nebular emission is very low, and it does not show the typical gradient exhibited by AGN, that is, increase toward the more nuclear regions (Constantin et al. 2012, in prep.). In both the Palomar and *HST*-STIS observations, the ratio [S II] $\lambda\lambda$ 6716/6731 is ~1.26 (Table 2), implying an electron density of $n_e \sim 10^2\,\mathrm{cm}^{-3}$, which lies at the lowest end in the distribution of particle densities measured in the Palomar objects [55].

Given the lack of a density gradient in the line-emitting region it is then rather peculiar to observe a significant increase (by a factor of ~6) in the [O I]/Hα line flux ratio in the *HST* spectrum relative to the large aperture observations, which would usually be interpreted as an indication for a more pronounced AGN-like ionization in the more central regions. There are two reasons for this: (1) [O I] requires a significantly hard radiation field, that is, that of an AGN, to sustain a sufficiently extensive partially ionized zone in clouds optically thick to Lyman continuum and thus to produce such a strong feature. Since the ionization potential of [O I] matches that of H very well, large differences in the [O I]/Hα ratios are expected between accretion and nonaccretion sources. (2) given the lack of a density stratification in this nucleus, the degree of ionization of the emitting gas is expected to diminish with radius as the ionizing radiation emerging from a nuclear source falls off in density as r^{-2} and thus produce strong gradients in the [O I]/Hα ratio.

In summary, NGC 4736 may be a broad-lined AGN, with an SED similar to other LLAGN, but it is atypical in several aspects whose physical origin remains unclear. Given the above-listed possible explanations for the peculiar $L_X/L_{\mathrm{H}\alpha}$ ratio, we evaluate in the next subsection whether photoionization by the weak AGN in this system is sufficiently powerful to balance the emission cooling in this system.

3.2. Comparison of Ionizing and Emission-Line Power. The multiwavelength observations of this system allow for a relatively rigorous assessment of whether the photoionization by this system's AGN can power the measured emission-line luminosities and in particular that of the broad Hα component. Following Eracleous et al. [18], we can run an energy budget test via a direct comparison of the Hα luminosity and count rate with the ionizing luminosity $L_i = L_{1\,\mathrm{Ry}-100\,\mathrm{keV}}$ and the ionizing photon rate $Q_i = Q_{1\,\mathrm{Ry}-100\,\mathrm{keV}}$.

It is important to treat separately the broad and narrow emission line features as there is strong evidence that they originate from regions of significantly different physical conditions. The broad Hα comes from a much more compact and much denser emitting gas than the narrow Balmer and forbidden lines. The difference in density is at least 3 orders of magnitude; the critical density for collisional excitation of the [O I] ($\sim 2 \times 10^6\,\mathrm{cm}^{-3}$) can be used to estimate the gas density of the broad-line region, as [O I] does not exhibit a broad component in this nucleus' emission. The difference in size is expected to be ~2 orders of magnitude (e.g., reverberation mapping of nearby Seyfert galaxies measured to be <1 week, [56]). Thus, the mechanisms that can operate in these two regions are expected to be qualitatively different. The photon and energy balance conditions should reflect these differences and thus should differ as well; in particular, potential contributions to the Hα emission via collisional excitation should be minimal for the narrow-line-emitting region, but important for the broad component.

Based on photoionization models (i.e., Cloudy, v94.0; see [57]) computed by Lewis et al. [58] for a wide range of ionization parameters, densities, and metallicities, Eracleous et al. [18] find that energy balance in a line-emitting nebula requires that $L_i > 18(\pm 2)L_{\mathrm{H}\alpha}/f_c$, where f_c is the covering factor or the fraction of the ionizing luminosity of the AGN that is absorbed by the line-emitting gas. These models are covering electron densities that appear to encompass both the broad- and the narrow-line-emitting regions in NGC 4736; thus, with only a fraction of 10% of ionizing photons being absorbed by the line-emitting gas, the minimum energy balance condition for AGN ionization is given by $L_i/L_{\mathrm{H}\alpha} > 180$, for both narrow and broad emission features. In the same time, however, a minimum requirement for photon balance can be quite different for the two emission regions: for the narrow-line component $Q_i > 2.2Q_{\mathrm{H}\alpha}$, corresponding to the case B recombination (i.e., one Hα photon is emitted for every 2.2 recombinations); for the denser broad-line-emitting region, the number of H-alpha photons that can be produced for each ionization can be at least 7-8 times higher than the standard case B estimate or $Q_i \gtrsim 0.25Q_{\mathrm{H}\alpha}$ (e.g., [37]).

With the L_i and Q_i values already calculated by Eracleous et al. [18] by integrating M94's nuclear SED assuming that pairs of points could be connected by a power law, and with the Hα measurements from all *HST* and ground-based observations presented above, we can proceed with the comparison. Table 4 lists the $L_{\mathrm{H}\alpha}$, $Q_{\mathrm{H}\alpha}$, and associated ratios $L_i/L_{\mathrm{H}\alpha}$, $Q_i/Q_{\mathrm{H}\alpha}$ for all of these optical spectroscopic observations, where the narrow and broad Hα measurements are shown separately.

It is readily apparent that while $L_i/L_{\mathrm{H}\alpha}$ is in general >18 ($f_c = 1$), it is almost never >180 ($f_c = 0.1$); the only exception is the narrow Hα emission measured in the *HST* aperture. Thus, a dominant AGN ionization of the narrow-line

Table 4: Ionizing and emission-line power.

	$\log L_{H\alpha}$ [a]	$\log Q_{H\alpha}$ [a]	$L_i = L_{H\alpha}$ [b]	$Q_i = Q_{H\alpha}$ [c]
Palomar	>37.7	>49.3	<93	<2.75
Bok	37.5	49.0	165.9	4.9
MMTO	37.5	49.2	104.7	3.1
HST-STIS, narrow	36.8	48.4	759	22.4
HST-STIS, broad	37.9	49.5	57.5	1.6
PCA tomography	38.8	50.3	8.7	0.25

[a]the luminosities are measured in erg s^{-1} and the photon rates in s^{-1}.
[b]the minimum energy balance condition for AGN ionization is given by $L_i/L_{H\alpha} \gtrsim 180$ (when a fraction of 10% of ionizing photons photons is absorbed by the line-emitting gas).
[c]the minimum photon balance condition for AGN ionization is given by $Q_i/Q_{H\alpha} > 2.2$ for the narrow line regions, and $Q_i/Q_{H\alpha} \gtrsim 0.25$ for the broad component.

region in this nucleus is definitely possible in the *HST* aperture, but only for $f_c \gtrsim 18\%$ at larger radial distances. For the *HST* broad Hα feature, an AGN ionization is possible only if $f_c \gtrsim 30\%$. The PCA tomography measurement of the broad Hα argues, however, against a balanced energy budget originating entirely in an AGN-like power mechanism, even when a maximum covering factor is considered; the AGN-produced energy falls short of the required amount by at least 50%. Interestingly, the measured $Q_i/Q_{H\alpha}$ ratio is well within the required photon balance corresponding to an AGN excitation for both the narrow and the broad emission features. Thus, with one clear exception, the AGN in NGC 4736 appears to be capable of providing enough photons to explain the observed Hα luminosity but only for relatively high f_c values.

Simply because the AGN-like SED of NGC 4736 can explain the majority of the ionization energy and photon rate does not imply that the actual mechanism is an AGN, alternative excitation mechanisms must be explored. Possible options are (i) we are missing ionizing photons from accretion onto the central BH, or (ii) there are other power sources that could make up the power deficit in this system, particularly for producing the broad Hα feature.

The first alternative could be possible if we were observing an "echo" of a previous epoch of more violent accretion, a few hundred years ago. This idea has been explored by Eracleous et al. [59] who showed that the reverberation of an ionizing flare in the nebula can produce LINER-like emission-line ratios. In this scenario, it is expected that the central UV source and the [O III] line would also follow the decay of the ionizing continuum; while the UV observations are not providing clear evidence for such a decay over the course of one decade, the multiple optical spectroscopic observations that we present in Figure 2 are consistent with a possible decrease in time (a few years) in the [O III] flux (in this model, Hβ, [S II], [O I] are expected to decay very slowly, in 60–250 years). Nevertheless, this duty-cycle hypothesis also requires that the broad-line emission fades immediately if the ionizing continuum declines, and the PCA tomography observations show that this is not happening in this source. The echo of such an ionizing continuum flare should also be detectable in a narrow-band [O III] λ5007 image in the form of a ring, which is not readily observed in M94.

The second alternative of power sources other than accretion has been often proposed in explaining the emission-line spectra of LINERs, with mixed success (see Section 1). The most probable alternative sources appear to be the mechanical power delivered by compact radio jets, along with photoionization by young or post-AGB stars from old or intermediate-age stellar populations. Shock models are highly unlikely to produce broad emission features and, thus, are not favored in this case. Recent star-formation activity remains, however, a viable option, particularly in light of relatively new discoveries of peculiar supernovae with broad Hα features that do not appear to fade in time. This idea is discussed in more detail in Section 5.

4. The Black Hole Mass

Given the unusual energetics of M94's nucleus, it is important to investigate whether the BH mass estimators derived for rapidly accreting Seyferts, which appear to be widely used for AGN, also work in the low-luminosity regime flagged by this particular system.

There are a variety of indirect methods that can be used to estimate the mass of BHs in galaxy centers. The available multiwavelength measurements of the nuclear emission for NGC 4736 allow the calculation of M_{BH} based on four different techniques, as well as with the M-σ^* relation. In this section, we explore and compare the results of these five methods, along with their consequences for this object's energetics, and a comparison with a recent dynamical measurement of $6.68(5.14\text{–}8.22)\times10^6\,M_\odot$ for this object, which is listed in Kormendy et al. [60] as obtained from (Gebhardt et al. 2012, in prep.). We first present the methods and associated M_{BH} calculations and then discuss the shortcomings of each measurement.

(1) Using the M-σ^* relation, established for quiescent nearby galaxies, including both ellipticals and spirals with classical bulges, as quantified by Gültekin et al. [1], for $\sigma^* = 110 \pm 5\,\mathrm{km\,s^{-1}}$ [61], $M_{BH} = 1.05 \pm 0.64 \times 10^7\,M_\odot$; within errors, this value agrees well with Gebhardt et al. dynamical measurement. With this value, $L_{bol}/L_{Edd} \approx 2 \times 10^{-5}$, which is consistent with the range of values within LINERs are expected to (e.g., [2, 62]).

(2) Using the *HST* measurements of the FWHM and the luminosity of the broad Hα component (see Table 1) within the scaling relation based solely on observations of this broad emission feature, as derived by Greene and Ho [63], we obtain $M_{BH} \approx 3\times10^4\,M_\odot$; the fractional uncertainty associated with this measurement is ~30%, and includes both the scatter in the scaling relation and the errors in the line measurements. With the broad Hα line luminosity measured in the PCA tomography study [44], the BH mass would increase (by a factor of 6.5) to $M_{BH} = 1.9 \times 10^5\,M_\odot$. The difference in these two values could be considered the most conservative uncertainty associated with this BH mass estimate. The corresponding Eddington ratios for these BH masses are $L_{bol}/L_{Edd} \approx 1\text{–}7 \times 10^{-3}$.

(3) Mid IR detection and measurements of the [Ne V] (14.32 μm) and [O IV] (25.89 μm) emission lines [64] give, via the empirical correlations between the MIR line luminosities and reverberation mapping-based $M_{\rm BH}$ values presented by Dasyra et al. [65], a black hole mass of $2.3\pm0.4\times 10^5\,M_\odot$ and $1.7\pm0.5\times10^5\,M_\odot$, respectively. The corresponding Eddington ratio is in this case $L_{\rm bol}/L_{\rm Edd}\approx 1.0\times10^{-3}$. Note that the scatter adopted for these relations is only a lower limit of the real value; thus, the uncertainty may be larger.

(4) The normalized X-ray excess variance method, as described in Papadakis et al. [66], applied to Chandra observations [30] gives $M_{\rm BH}=2.5\pm1.7\times10^5\,M_\odot$. To be specific, we used for this calculation the excess variance $\sigma=0.06\pm0.04$, $L_{\rm bol}=2.5\times10^{40}\,{\rm erg\,s^{-1}}$ as derived in Section 3, and $\nu_{\rm lf}=1/T=1/14\ {\rm h}^{-1}$ (where T is the length of the light curve), to estimate the break frequency $\nu_{\rm bf}$ and then $M_{\rm BH}$ via equations 4 and 6, respectively, of Papadakis et al. [66].

(5) The $M_{\rm BH}$ of this system can also be obtained via the "fundamental plane of black hole activity" that relates black hole mass to the emitted compact radio $L_R=\nu L_\nu(5\,{\rm GHz})$ and hard X-ray luminosities L_X(2–10 keV) and spans nine orders of magnitude in black hole mass. We calculate $L_R=1.7\times10^{35}\,{\rm erg\,s^{-1}}$ using the two 8.5 GHz and 15 GHz radio measurements presented in this paper (Figure 3, Table 3), with a flux modeled by a powerlaw $S_\nu\propto\nu^{-\alpha}$. Using the empirical fits of Merloni et al. [67] for the fundamental plane relation, $\log(M_{\rm BH}/M_\odot)=5.9\pm1.1$ for the black hole in this system; note that this measurement and its associated errors embrace all of the estimates presented above and, thus, do not provide any additional constraint to the $M_{\rm BH}$. Interestingly, the latest derivation of the fundamental plane relation [68], applied to very low nuclear galactic luminosities, provides a $\log(M_{\rm BH}/M_\odot)=7.2\pm0.4$ for this system, which remains consistent only with the M-σ^* value, as it departs considerably from those provided by the relations employing AGN emission.

There is a significant inconsistency between the value given by the M-σ^* relation, which is supported by Gebhardt's dynamical measurement, and those based on the AGN emission. The M-σ^* relation suggests a BH mass in NGC 4736 of $1\times10^7\,M_\odot$ and a correspondingly low $L_{\rm bol}/L_{\rm Edd}$ of $\sim\ 2\times10^{-5}$. Three of the estimates based on the AGN emission converge to black hole masses of $\sim10^5\,M_\odot$ (and $L_{\rm bol}/L_{\rm Edd}\sim10^{-3}$), showing a surprisingly consistent departure of two orders of magnitude from the M-σ^* estimate. The fundamental plane relation provides a $M_{\rm BH}$ value right in between these two different situations, however, with no real additional constraint, due to its associated large uncertainty. These differences are somewhat puzzling given that the BH mass estimates based on nuclear emission properties are all calibrated to follow the M-σ^* for high-mass black holes ($M_{\rm BH}\gtrsim10^6\,M_\odot$). It is however true that the calibrators are biased toward nearby Seyfert galaxies with much higher Eddington ratios than that of NGC 4736. We briefly discuss in the following subsections more specific weaknesses of each measurement.

4.1. Caveats of the M-σ^* Relation. The M-σ^* relation is expected to provide a reliable estimate of the BH mass as it is based on the strong correlation between dynamical mass measurements of supermassive BHs and their host properties [1, 69–71]. This relationship is derived primarily from ellipticals and spirals with classical bulges (formed during major mergers). Recent observations suggest it may not be valid for samples of later-type spirals which more commonly host pseudobulges (formed via secular disk processes) (e.g., [72]). Because the distinction between classical and pseudobulges is based on formation, it does not simply correlate with observable properties [73, 74]. In general pseudobulges are less luminous, have lower bulge-to-total ratios, and have ongoing star formation and lower Sérsic indices than classical bulges. The possibility that the NGC 4736 is in fact a pseudobulge provides a solution to the apparent conflict in BH mass estimates.

Nevertheless, pseudobulges are difficult to identify, and there is not yet a consensus as to what defines them. To complicate things further, classical and pseudobulges can exist within the same galaxy (e.g., NGC 2787; [75]). The classification of the NGC 4736 bulge is ambiguous: Fisher and Drory [74] classifies it as a pseudobulge based on its nuclear spiral and bar and low Sérsic index ($n=1.3$). However, they also find it has a low star formation rate, more typical of classical bulges, and thus classify it as an "Inactive Pseudobulge;" if NGC 4736 hosts a pseudobulge, it is an atypical one. We have created a surface brightness profile from NICMOS and 2MASS H-band data from the Large Galaxy Atlas [76] and found results that conflict with the fits of Fisher and Drory [74]. Specifically, we find Sérsic indices for the bulge of $n=2.3$–3.0 depending on the radial range and type of fit (single versus double Sérsic), which are consistent with a classical one (as shown in [74]).

Our surface brightness profile fits also reveal the presence of a nuclear star cluster within the central $\sim0.6''$ ($\sim$12 pc), with an H-band magnitude of $\sim$12.5. Such nuclear star clusters are common in early-type spiral galaxies [77]. The luminosity and mass of nuclear star clusters are known to scale with bulge luminosity and mass [78–80], and the NGC 4736 nuclear star cluster has a luminosity that is 0.1% of its bulge, typical for nuclear star clusters [81]. Nuclear star clusters commonly coexist with black holes, but there are a very limited number of cases where masses for both can be estimated [82]. In these cases, including the Milky Way, the BH mass is similar to the mass of the nuclear star cluster within an order of magnitude. There is also some evidence that the ratio of BH mass to nuclear star cluster mass increases with spheroid mass [83]. For NGC 4736, assuming an old population with an H-band $M/L\sim0.7$, the nuclear cluster would have a mass of $\sim2\times10^7\,M_\odot$. This mass is quite similar to the M-σ^* BH mass estimate, suggesting thus the presence of a similar-sized BH.

4.2. Caveats of the BLR Scaling Relation. The BLR scaling relations were derived using high-luminosity systems (i.e., Seyferts, not LINERs) and have been scaled to match the dynamical black hole detections of BHs with masses $>10^6\,M_\odot$. Thus, their applicability to NGC 4736, where they yield an estimate of $\sim10^5\,M_\odot$, is a (perhaps unwise) extrapolation. These relations assume that the BLR is virialized due

to proximity to the BH [84] and thus the 5100 Å continuum luminosity correlates with the emissivity-weighted radius of the BLR and thus with the BH mass [85]. The overall errors associated with BH mass measurements based on these relations do not exceed 0.5 dex [86–88], and there appears to be good consistency with BH masses obtained via the M-σ^* relation for relatively bright AGN ($L > 10^{42}\,\mathrm{erg\,s^{-1}}$), with $M_{\mathrm{BH}} \gtrsim 10^6\,M_\odot$ [89]. More recent calibrations of the radius-luminosity relationship on which these techniques are based infer that BH masses have been overestimated, however, only by up to a factor of ∼3 [90]. This latter study also indicates a trend toward larger uncertainties and larger amount of overestimation in the BH with decreasing luminosity; however, no conclusive results are yet available for this regime.

4.3. Caveats of the MIR Line Correlation. The Dasyra et al. [65] empirical relation between the MIR line emission properties and the BH mass is derived using reverberation mapping BH masses and, thus, as with the previous method, may not apply to systems with $M_{\mathrm{BH}} < 10^6\,M_\odot$. The relation holds for systems with $L_{\mathrm{bol}}/L_{\mathrm{Edd}} > 0.003$, but not necessarily beyond this range. The $M_{\mathrm{BH}} \approx 10^5\,M_\odot$ given by this method places this object at the low end of this $L_{\mathrm{bol}}/L_{\mathrm{Edd}}$ range; on the other hand, a more massive and thus quiescent BH would correspond to a [Ne V] luminosity of a few orders of magnitude lower than that measured. The scatter associated with this relation remains in average 0.5 dex and thus cannot account for the ∼2-3 orders of magnitude difference between this value and that obtained via M-σ^*.

4.4. Issues with Estimating the BH Mass from X-Rays. The BH mass derivation based on X-ray variability, or more precisely on the relation between the excess variance and M_{BH}, relies on the hypothesis of a universal power spectral density function (PSD) shape and amplitude in AGN, which is based on the idea that the X-ray variability mechanism and the accretion efficiency are the same for all AGN, at all redshifts. These assumptions appear to hold for the objects involved in deriving and testing this relation [66]; that sample is, however, small and rather biased toward luminous (X-ray) sources, with $\log L_X/(\mathrm{erg\,s^{-1}}) > 41.5$. None of the 2–10 keV sources detected in the nucleus of NGC 4736 are in this luminosity range. Nevertheless, Galactic BHs ($M_{\mathrm{bh}} < 10^2\,M_\odot$) in their hard states show variability properties that match well those of AGN, both of these types of sources falling on the same projection of the T_B-M_{bh}-$L_{\mathrm{bol}}/L_{\mathrm{Edd}}$ plane, where $T_B = 1/\nu_{\mathrm{bf}}$, suggesting that an extrapolation to the intermediate mass BH (or lower L) level is practicable.

4.5. Problems with the Fundamental Plane. Finally, it appears that the main problem with the fundamental plane is that the relation is not sharpened enough to provide strong constraints on BH masses. Because the relation spans nine orders of magnitude in BH mass, it is expected to equally apply to any value of M_{bh} in the range we are interested in. It is also the case that a wide variety of BH accretion models (e.g., with efficient and inefficient flows for the X-ray emission or associating X-ray flux with synchrotron emission near the base of a jet) are consistent with this relation (e.g., [91, 92]), suggesting that a large diversity of accretion modes or rates are accommodated.

Nevertheless, the fundamental plane remains best constrained only for systems with $M_{\mathrm{BH}} \gtrsim 10^6\,M_\odot$ and for very low nuclear luminosities, that is, with negligible or zero AGN contribution that allow a dynamical measurement of their BH mass [68]. In this regime, the latest derivation of the fundamental plane relation provides for M94 a M_{BH} value consistent with the M-σ^* estimate and the dynamical measurement, while it would not reliably constrain a $M_{\mathrm{BH}} \lesssim 10^6\,M_\odot$ value. The scatter in the fundamental plane relation increases for lower BH masses, with higher Eddington ratios [68, 93].

5. Discussion: Alternative Power Generation Mechanisms

The detection of broad Hα emission, combined with the spatial coincidence of this emission with the detection of X-ray, UV, and radio compact sources, provides strong evidence that this galaxy hosts an accreting massive black hole. The SED of this nucleus makes this object one of the lowest luminosity LINER with a distinct contribution to the total emission by black hole accretion. In this scenario, the nucleus is an AGN, and the presence of off-nuclear sources, particularly the radio and the UV detections, may result from remnant jet activity emerging from the nucleus. It is definitely exciting to detect AGN activity at energy levels equal to that of several young supernova remnants of the Cas A variety [94], an OB association which hosts a high-mass X-ray binary [30] or simply a group of five late O supergiants [28].

Nevertheless, the presence of a number of unusual off-nuclear sources, coupled with the apparent deficit in the photoionizing photons, and with the fact that the BH mass estimates based on AGN emission appear to fail for this object, encourages us to explore alternative scenarios for the NGC 4736 nucleus. In particular, it is very likely that the BH mass estimates that exploit the multiwavelength AGN characteristics do not work for this system because at least some of this emission is not the result of BH accretion.

There are certain (peculiar) kinds of core-collapse SNe that present multiwavelength observations and in particular broad Hα emission components with characteristics that are very similar to those we measure in the nucleus of M94. The SN 2005ip, presented by Smith et al. [95], is a very good example. The so-called intermediate Hα component associated with this SN ejecta presents the same FWHM and brightness level as the broad Hα detected in the STIS aperture, and it does not show any sign of diminishing its strength over more than 3-year period (see their Figure 7; a *very* broad component with FWHM $\gtrsim 10000\,\mathrm{km\,s^{-1}}$ like the one exhibited by SN 2005ip would not be measurable in the galaxy spectrum as it would be completely swamped in the continuum stellar light, even in the *HST*-STIS observations).

SNe as luminous as X2, with intrinsically hard X-ray spectra (photon index $\Gamma \lesssim 1$), have certainly been encountered, for cases observed few years after the explosion

(e.g., ATe #1023; [96]); for these cases, $L_X/L_{\rm bol}$ ratios are high relative to more standard SN cases [97]. Another particular example of a SN that matches well the measurements of the nuclear emission in NGC 4736 is SDSS J09529.56 + 214313.3 [98], which is believed (but not confirmed) to be a SN type IIn. For this system, the broad Hα component stays strong for at least three years, and its $L_X(2\text{–}10\,\text{keV})/L_{\text{H}\alpha} \approx 1$, matching thus very well the surprisingly low value measured for NGC 4736.

The peculiar nebular, radio, and UV characteristics of the nucleus of NGC 4736 appear to also compare well the SN phenomenology. The electron densities (or [S II] line flux ratios) measured in this galaxy center are matching exactly the ones measured in the environments of the extraordinary type IIn SNe we compare here with, for example, SDSS J09529.56 + 214313.3. The two different band radio observations of this nucleus are consistent with emission from extragalactic SNe, in both intensity and decline rate in flux density [99–101]. Moreover, the compact radio off-nuclear detection, only 1″ (20 pc) away from the nuclear one [31], could be interpreted as the result of shock emission associated with a nuclear core-collapse SN. Measurements of the brightness of the off-nuclear UV detection in F250W and F330W place this object into the O star of late (5-ish) type spectral category, fitting thus well into the idea that this nucleus could simply be a star-forming site and that NGC 4736a's emission includes significant contributions from a SN which exploded close to the weakly active (and massive, $M_{\rm BH} \sim 10^{(6\text{-}7)} M_\odot$) central BH.

This SN contribution scenario may have its own drawbacks. If the broad Hα emission has actually gotten brighter by a factor of 6, as suggested by the comparison between the *HST*-STIS and GMOS observations, the SN interpretation becomes problematic; variability of the broad-line region originating in the AGN could, in principle, account for this effect. Also, the observed intraday (hour-scale) X-ray variability measured for X2 remains yet to be detected in a supernova and conflicts with the physical scale over which X-ray emission is expected in SN remnants; thus, this particular behavior may remain strictly associated with the AGN. Additional observations would be necessary to fully confirm or rule out the SN scenario. Specifically, new high-resolution observations in the optical, UV, or X-rays would be able to confirm if there is a fading in the light curve, as expected from a SN, and would also allow accurate localization of the source. High S/N optical spectra would much better resolve the emission-line profile and a possible temporal evolution.

6. Conclusion

We have presented here an exhaustive multiwavelength analysis of the nuclear emission properties of NGC 4736, prompted by new measurements of a broad Hα emission component detected in its high-resolution *HST*-STIS optical spectrum. This broad Hα component, with a luminosity of $9 \times 10^{37}\,\text{ergs}\,\text{s}^{-1}$, is one of the lowest luminosity broad line known. This broad Hα is coincident with a compact bright X-ray and radio source. Our measurements of this object's spectral energy distribution reveal a bolometric luminosity of $L_{\rm bol} \approx 2.5 \times 10^{40}\,\text{erg}\,\text{s}^{-1}$, that categorizes NGC 4736 as one of the least luminous LINERs with strong evidence for BH accretion. Our comparison of five independent BH mass estimates reveals a discrepancy of two orders of magnitude between the value $\sim 10^7\,M_\odot$ predicted by the M-σ^* relation and the value $\sim 10^5\,M_\odot$ toward which methods based on AGN emission activity in optical, mid-IR, and X-ray, seem to converge; the fifth method is provided by the fundamental plane relation, which, however, due to its large associated uncertainties, does not offer any additional constraint to this comparison.

We conclude that this system's BH mass cannot be reliably estimated via standard AGN BH mass indicators because the nuclear emission in this system is not entirely tracing the accretion onto the central BH. Our assessment of the energy budgets of the ionizing and emission-line power suggests a possible deficit in the AGN ionization and production of a broad Hα emission feature which can be made up by a peculiar kind of Type IIn SN that matches well the nuclear emission of NGC 4736 over the whole electromagnetic spectrum and supports this galaxy nucleus' general aging starburst-like appearance.

Acknowledgments

The authors thank the anonymous referee for constructive comments that helped them improved the paper. Support for this work was provided by NASA through Grant no. HST-AR-11749.01-A from the STScI, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA Contract NAS5-26555, based on observations made with the NASA/ESA *Hubble Space Telescope*, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA Contract NAS 5-26555. Some observations reported here were obtained at the MMT Observatory, a joint facility of the Smithsonian Institution and the University of Arizona.

References

[1] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "the m-σ and m-l relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal*, vol. 698, no. 1, p. 198, 2009.

[2] L. C. Ho, "Nuclear activity in nearby galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 46, pp. 475–539, 2008.

[3] J. A. Baldwin, M. M. Phillips, and R. Terlevich, "Classification parameters for the emission-line spectra of extragalactic objects," *Astronomical Society of the Pacific*, vol. 93, pp. 5–19, 1981.

[4] S. Veilleux and D. E. Osterbrock, "Spectral classification of emission-line galaxies," *Astrophysical Journal, Supplement Series*, vol. 63, pp. 295–301, 1987.

[5] G. Kauffmann, T. M. Heckman, C. Tremonti et al., "The host galaxies of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 346, no. 4, pp. 1055–1077, 2003.

[6] L. J. Kewley, B. Groves, G. Kauffmann, and T. Heckman, "The host galaxies and classification of active galactic nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 3, pp. 961–976, 2006.

[7] L. C. Ho, A. V. Filippenko, and W. L. W. Sargent, "A search for "dwarf" seyfert nuclei. V. Demographics of nuclear activity in nearby galaxies," *Astrophysical Journal*, vol. 487, no. 2, pp. 568–578, 1997.

[8] A. Constantin and M. S. Vogeley, "The clustering of low-luminosity active galactic nuclei," *Astrophysical Journal*, vol. 650, no. 2 I, pp. 727–748, 2006.

[9] D. W. Weedman, L. Hao, S. J. U. Higdon et al., "Mid-infrared spectra of classical AGNs observed with the spitzer space telescope," *Astrophysical Journal*, vol. 633, no. 2 I, pp. 706–716, 2005.

[10] L. Armus, J. Bernard-Salas, H. W. W. Spoon et al., "Detection of the buried active galactic nucleus in NGC 6240 with the infrared spectrograph on the Spitzer Space Telescope," *Astrophysical Journal*, vol. 640, no. 1 I, pp. 204–210, 2006.

[11] N. P. Abel and S. Satyapal, "[Ne v] Emission in optically classified starbursts," *Astrophysical Journal*, vol. 678, no. 2, pp. 686–692, 2008.

[12] E. Sturm, D. Lutz, A. Verma et al., "Mid-infrared line diagnostics of active galaxies: a spectroscopic AGN survey with ISO-SWS," *Astronomy and Astrophysics*, vol. 393, no. 3, pp. 821–841, 2002.

[13] S. Satyapal, D. Vega, R. P. Dudik, N. P. Abel, and T. Heckman, "Spitzer uncovers active galactic nuclei missed by optical surveys in seven late-type galaxies," *Astrophysical Journal*, vol. 677, no. 2, pp. 926–942, 2008.

[14] A. D. Goulding and D. M. Alexander, "Towards a complete census of AGN in nearby Galaxies: a large population of optically unidentified AGN," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 3, pp. 1165–1193, 2009.

[15] N. M. Nagar, H. Falcke, and A. S. Wilson, "Radio sources in low-luminosity active galactic nuclei IV. Radio luminosity function, importance of jet power, and radio properties of the complete Palomar sample," *Astronomy and Astrophysics*, vol. 435, no. 2, pp. 521–543, 2005.

[16] W. M. Zhang, R. Soria, S. N. Zhang, D. A. Swartz, and J. Liu, "A census of X-ray nuclear activity in nearby galaxies," *Astrophysical Journal*, vol. 699, no. 1, pp. 281–297, 2009.

[17] L.-B. Desroches, J. E. Greene, and L. C. Ho, "X-ray properties of intermediate-mass black holes in active galaxies. II. X-ray-bright accretion and possible evidence for slim disks," *Astrophysical Journal*, vol. 698, p. 1515, 2009.

[18] M. Eracleous, J. A. Hwang, and H. M. L. G. Flohic, "An assessment of the energy budgets of low-ionization nuclear emission regions," *Astrophysical Journal*, vol. 711, no. 2, pp. 796–810, 2010.

[19] A. Constantin, F. Hoyle, and M. S. Vogeley, "Active galactic nuclei in void regions," *Astrophysical Journal*, vol. 673, no. 2, pp. 715–729, 2008.

[20] A. Constantin, P. Green, T. Aldcroft et al., "Probing the balance of AGN and star-forming activity in the local universe with ChaMP," *Astrophysical Journal*, vol. 705, no. 2, pp. 1336–1355, 2009.

[21] K. Schawinski, D. Thomas, M. Sarzi et al., "Observational evidence for AGN feedback in early-type galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 4, pp. 1415–1431, 2007.

[22] K. Schawinski, C. M. Urry, S. Virani et al., "Galaxy zoo: the fundamentally different co-evolution of supermassive black holes and their early- and late-type host galaxies," *Astrophysical Journal*, vol. 711, no. 1, pp. 284–302, 2010.

[23] T. Di Matteo, V. Springel, and L. Ilernquist, "Energy input from quasars regulates the growth and activity of black holes and their host galaxies," *Nature*, vol. 433, no. 7026, pp. 604–607, 2005.

[24] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes, and galaxy spheroids," *Astrophysical Journal, Supplement Series*, vol. 163, no. 1, pp. 1–49, 2006.

[25] R. B. Tully and J. R. Fisher, *Catalog of Nearby Galaxies*, Cambridge University Press, Cambridge, UK, 1988.

[26] A. V. Filippenko and W. L. W. Sargent, "A search for "Dwarf" seyfert 1 nuclei. I. The initial data and results," *Astrophysical Journal, Supplement Series*, vol. 57, pp. 503–522, 1985.

[27] P. F. Spinelli, T. Storchi-Bergmann, C. H. Brandt, and D. Calzetti, "An atlas of Hubble space telescope stis spectra of Seyfert galaxies," *Astrophysical Journal, Supplement Series*, vol. 166, no. 2, pp. 498–504, 2006.

[28] D. Maoz, A. V. Filippenko, L. C. Ho et al., "Detection of compact ultraviolet nuclear emission in liner galaxies," *Astrophysical Journal*, vol. 440, no. 1, pp. 91–99, 1995.

[29] D. Maoz, N. M. Nagar, H. Falcke, and A. S. Wilson, "The murmur of the sleeping black hole: detection of nuclear ultraviolet variability in LINER galaxies," *Astrophysical Journal*, vol. 625, no. 2 I, pp. 699–715, 2005.

[30] M. Eracleous, J. C. Shields, G. Chartas, and E. C. Moran, "Three LINERs under the Chandra X-ray microscope," *Astrophysical Journal*, vol. 565, no. 1 I, pp. 108–124, 2002.

[31] E. Körding, E. Colbert, and H. Falcke, "A radio monitoring survey of ultra-luminous X-ray sources," *Astronomy and Astrophysics*, vol. 436, no. 2, pp. 427–436, 2005.

[32] A. L. Kinney, R. C. Bohlin, D. Calzetti, N. Panagia, and R. F. G. Wyse, "An atlas of ultraviolet spectra of star-forming galaxies," *Astrophysical Journal, Supplement Series*, vol. 86, no. 1, pp. 5–93, 1993.

[33] B. J. Smith, P. M. Harvey, C. Colomé, C. Y. Zhang, J. DiFrancesco, and R. W. Pogge, "Far-infrared emission from the bulges of early-type spirals: kao observations of NGC 4736 (M94) and NGC 3627 (M66)," *Astrophysical Journal*, vol. 425, no. 1, pp. 91–102, 1994.

[34] L. Dressel, P. Hodge, and P. Barrett, "wx2d: a PyRAF routine to resample spectral images," *Instrument Science Report STIS 2007-04*, 2007.

[35] G. Bruzual and S. Charlot, "Stellar population synthesis at the resolution of 2003," *Monthly Notices of the Royal Astronomical Society*, vol. 344, no. 4, pp. 1000–1028, 2003.

[36] C. A. Tremonti, T. M. Heckman, G. Kauffmann et al., "The origin of the mass-metallicity relation: insights from 53,000 star-forming galaxies in the sloan digital sky survey," *Astrophysical Journal*, vol. 613, no. 2 I, pp. 898–913, 2004.

[37] D. E. Osterbrock, *Astrophysics of Gaseus Nebulae and Active Galactic Nuclei*, University Science Books, 1989.

[38] G. Kriss, "Fitting models to UV and optical spectral data," in *Astronomical Data Analysis Software and Systems III*, D. R. Crabtree, R. J. Hanisch, and J. Barnes, Eds., vol. 61 of *Astronomical Society of the Pacific Conference Series*, p. 437, 1994.

[39] J. L. Walsh, A. J. Barth, and M. Sarzi, "The supermassive black hole in M84 revisited," *Astrophysical Journal*, vol. 721, no. 1, pp. 762–776, 2010.

[40] L. C. Ho, A. V. Filippenko, W. L. W. Sargent, and C. Y. Peng, "A search for "dwarf" seyfert nuclei. IV. Nuclei with broad Hα emission," *Astrophysical Journal, Supplement Series*, vol. 112, no. 2, pp. 391–414, 1997.

[41] J. E. Greene and L. C. Ho, "A new sample of low-mass black holes in active galaxies," *Astrophysical Journal Letters*, vol. 670, no. 1, pp. 92–104, 2007.

[42] A. V. Filippenko and W. L. W. Sargent, "Discovery of an extremely low luminosity Seyfert 1 nucleus in the dwarf galaxy NGC 4395," *The Astrophysical Journal*, vol. 342, pp. 11–14, 1989.

[43] L. C. Ho, A. V. Filippenko, and W. L. W. Sargent, "A search for "dwarf" seyfert nuclei. III. Spectroscopic parameters and properties of the host galaxies," *Astrophysical Journal, Supplement Series*, vol. 112, no. 2, pp. 315–390, 1997.

[44] J. E. Steiner, R. B. Menezes, T. V. Ricci, and A. S. Oliveira, "PCA Tomography: how to extract information from data cubes," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 1, pp. 64–75, 2009.

[45] J. Moustakas and R. C. Kennicutt Jr., "An integrated spectrophotometric survey of nearby star-forming galaxies," *Astrophysical Journal, Supplement Series*, vol. 164, no. 1, pp. 81–98, 2006.

[46] N. M. Nagar, H. Falcke, A. S. Wilson, and J. S. Ulvesta, "Radio sources in low-luminosity active galactic nuclei. III. "AGNs" in a distance-limited sample of "LLAGNs"," *Astronomy and Astrophysics*, vol. 392, no. 1, pp. 53–82, 2002.

[47] L. C. Ho, "The spectral energy distributions of low-luminosity active galactic nuclei," *Astrophysical Journal*, vol. 516, no. 2, pp. 672–682, 1999.

[48] R. C. Kennicutt, L. Armus, G. Bendo et al., "SINGS: the SIRTF nearby galaxies survey," *Publications of the Astronomical Society of the Pacific*, vol. 115, no. 810, pp. 928–952, 2003.

[49] T. M. Heckman, G. Kauffmann, J. Brinchmann, S. Charlot, C. Tremonti, and S. D. M. White, "Present-day growth of black holes and bulges: the sloan digital sky survey perspective," *Astrophysical Journal*, vol. 613, no. 1 I, pp. 109–118, 2004.

[50] G. Kauffmann and T. M. Heckman, "Feast and Famine: regulation of black hole growth in low-redshift galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 1, pp. 135–147, 2009.

[51] S. Charlot and S. M. Fall, "A simple model for the absorption of starlight by dust in galaxies," *Astrophysical Journal*, vol. 539, no. 2, pp. 718–731, 2000.

[52] M. Eracleous, J. A. Hwang, and H. M. L. G. Flohic, "Spectral energy distributions of weak active galactic nuclei associated with low-ionization nuclear emission regions," *Astrophysical Journal, Supplement Series*, vol. 187, no. 1, pp. 135–148, 2010.

[53] S. Satyapal, D. Vega, T. Heckman, B. O'Halloran, and R. Dudik, "The discovery of an active galactic nucleus in the late-type galaxy NGC 3621: spitzer spectroscopic observations," *Astrophysical Journal*, vol. 663, no. 1, pp. L9–L12, 2007.

[54] L. C. Ho, E. D. Feigelson, L. K. Townsley et al., "Detection of nuclear X-ray sources in nearby galaxies with chandra," *Astrophysical Journal*, vol. 549, no. 1, pp. L51–L54, 2001.

[55] L. C. Ho, A. V. Filippenko, and W. L. W. Sargent, "A search for "dwarf" Seyfert nuclei. VI. Properties of emission-line nuclei in nearby galaxies," *Astrophysical Journal*, vol. 583, no. 1 I, pp. 159–177, 2003.

[56] K. D. Denney, B. M. Peterson, R. W. Pogge et al., "Reverberation mapping measurements of black hole masses in six local Seyfert galaxies," *Astrophysical Journal*, vol. 721, no. 1, pp. 715–737, 2010.

[57] G. J. Ferland, K. T. Korista, D. A. Verner, J. W. Ferguson, J. B. Kingdon, and E. M. Verner, "CLOUDY 90: Numerical simulation of plasmas and their spectra," *Publications of the Astronomical Society of the Pacific*, vol. 110, no. 749, pp. 761–778, 1998.

[58] K. T. Lewis, M. Eracleous, and R. M. Sambruna, "Emission-line diagnostics of the central engines of weak-line radio galaxies," *Astrophysical Journal*, vol. 593, no. 1 I, pp. 115–126, 2003.

[59] M. Eracleous, M. Livio, and L. Binette, "A duty cycle hypothesis for the central engines of LINERs," *Astrophysical Journal Letters*, vol. 445, p. L1, 1995.

[60] J. Kormendy, R. Bender, and M. E. Cornell, "Supermassive black holes do not correlate with galaxy disks or pseudobulges," *Nature*, vol. 469, no. 7330, pp. 374–376, 2011.

[61] A. J. Barth, L. C. Ho, and W. L. W. Sargent, "A study of the direct fitting method for measurement of galaxy velocity dispersions," *Astronomical Journal*, vol. 124, no. 5, pp. 2607–2614, 2002.

[62] L. C. Ho, "Black hole demography from nearby active galactic nuclei," in *Coevolution of Black Holes and Galaxies*, L. C. Ho, Ed., the Carnegie Observatories Astrophysics Series, p. 293, Cambridge University Press, 2004.

[63] J. E. Greene and L. C. Ho, "Estimating black hole masses in active galaxies using the Hα emission line," *Astrophysical Journal*, vol. 630, no. 1 I, pp. 122–129, 2005.

[64] R. P. Dudik, S. Satyapal, and D. Marcu, "A *spitzer* spectroscopic survey of low-ionization nuclear emission-line regions: characterization of the central source," *Astrophysical Journal*, vol. 691, p. 1501, 2009.

[65] K. M. Dasyra, L. C. Ho, L. Armus et al., "High-ionization mid-infrared lines as black hole mass and bolometric luminosity indicators in active galactic nuclei," *Astrophysical Journal*, vol. 674, no. 1, pp. L9–L12, 2008.

[66] I. E. Papadakis, E. Chatzopoulos, D. Athanasiadis, A. Markowitz, and I. Georgantopoulos, "The long-term X-ray variability properties of AGNs in the Lockman Hole region," *Astronomy and Astrophysics*, vol. 487, no. 2, pp. 475–483, 2008.

[67] A. Merloni, S. Heinz, and T. Di Matteo, "A fundamental plane of black hole activity," *Monthly Notices of the Royal Astronomical Society*, vol. 345, no. 4, pp. 1057–1076, 2003.

[68] K. Gültekin, E. M. Cackett, J. M. Miller et al., "the fundamental plane of accretion onto black holes with dynamical masses," *The Astrophysical Journal*, vol. 706, no. 1, p. 404, 2009.

[69] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.

[70] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.

[71] S. Tremaine, K. Gebhardt, R. Bender et al., "The slope of the black hole mass versus velocity dispersion correlation," *Astrophysical Journal*, vol. 574, no. 2 I, pp. 740–753, 2002.

[72] J. E. Greene, C. Y. Peng, M. Kim et al., "Precise black hole masses from megamaser disks: black hole-bulge relations at low mass," *Astrophysical Journal*, vol. 721, no. 1, pp. 26–45, 2010.

[73] J. Kormendy and R. C. Kennicutt, "Secular evolution and the formation of pseudobulges in disk galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 42, pp. 603–683, 2004.

[74] D. B. Fisher and N. Drory, "Bulges of nearby galaxies with Spitzer: scaling relations in pseudobulges and classical bulges," *Astrophysical Journal*, vol. 716, no. 2, pp. 942–969, 2010.

[75] P. Erwin, J. C. Vega Beltrán, A. W. Graham, and J. E. Beckman, "When is a bulge not a bulge? Inner disks masquerading as bulges in NGC 2787 and NGC 3945," *Astrophysical Journal Letters*, vol. 597, no. 2 I, pp. 929–947, 2003.

[76] T. H. Jarrett, T. Chester, R. Cutri, S. E. Schneider, and J. P. Huchra, "The 2mass large galaxy atlas," *Astronomical Journal*, vol. 125, no. 2, pp. 525–554, 2003.

[77] C. M. Carollo, M. Stiavelli, M. Seigar, T. P. De Zeeuw, and H. Dejonghe, "Spiral galaxies with HST/NICMOS. I. Nuclear morphologies, color maps, and distinct nuclei," *Astronomical Journal*, vol. 123, no. 1, pp. 159–183, 2002.

[78] M. Balcells, A. W. Graham, L. Domínguez-Palmero, and R. F. Peletier, "Galactic bulges from Hubble Space Telescope near-infrared camera multi-object spectrometer observations: the lack of r1/4 bulges," *Astrophysical Journal*, vol. 582, no. 2, pp. L79–L82, 2003.

[79] L. Ferrarese, P. Côté, E. D. Bontà et al., "A fundamental relation between compact stellar nuclei, supermassive black holes, and their host galaxies," *Astrophysical Journal*, vol. 644, pp. 21–24, 2006.

[80] J. Rossa, R. P. Van Der Marel, T. Böker et al., "Hubble space telescope stis spectra of nuclear star clusters in spiral galaxies: dependence of age and mass on hubble type," *Astronomical Journal*, vol. 132, no. 3, pp. 1074–1099, 2006.

[81] P. Côté, S. Piatek, L. Ferrarese et al., "The ACS Virgo Cluster Survey. VIII. The nuclei of early-type galaxies," *Astrophysical Journal, Supplement Series*, vol. 165, no. 1, pp. 57–94, 2006.

[82] A. Seth, M. Agüeros, D. Lee, and A. Basu-Zych, "The coincidence of nuclear star clusters and active galactic nuclei," *Astrophysical Journal*, vol. 678, no. 1, pp. 116–130, 2008.

[83] A. W. Graham and L. R. Spitler, "Quantifying the coexistence of massive black holes and dense nuclear star clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 4, pp. 2148–2162, 2009.

[84] B. M. Peterson and A. Wandel, "Keplerian motion of broad-line region gas as evidence for supermassive black holes in active galactic nuclei," *Astrophysical Journal*, vol. 521, no. 2, pp. L95–L98, 1999.

[85] S. Kaspi, P. S. Smith, H. Netzer, D. Maoz, B. T. Jannuzi, and U. Giveon, "Reverberation measurements for 17 quasars and the size-mass-luminosity relations in active galactic nuclei," *Astrophysical Journal*, vol. 533, no. 2, pp. 631–649, 2000.

[86] M. Vestergaard, "Determining central black hole masses in distant active galaxies," *Astrophysical Journal*, vol. 571, no. 2 I, pp. 733–752, 2002.

[87] C. H. Nelson, R. F. Green, G. Bower, K. Gebhardt, and D. Weistrop, "The relationship between black hole mass and velocity dispersion in Seyfert 1 galaxies," *Astrophysical Journal*, vol. 615, no. 2 I, pp. 652–661, 2004.

[88] C. A. Onken, L. Ferrarese, D. Merritt et al., "Supermassive black holes in active galactic nuclei. II. Calibration of the black hole mass-velocity dispersion relationship for active galactic nuclei," *Astrophysical Journal*, vol. 615, no. 2 I, pp. 645–651, 2004.

[89] A. J. Barth, J. E. Greene, and C. H. O. Luis, "Dwarf Seyfert 1 nuclei and the low-mass end of the MBH-σ relation," *Astrophysical Journal*, vol. 619, no. 2, pp. L151–L154, 2005.

[90] M. C. Bentz, B. M. Peterson, H. Netzer, R. W. Pogge, and M. Vestergaard, "The radius-luminosity relationship for active galactic nuclei: the effect of host-galaxy starlight on luminosity measurements. II. the full sample of reverberation-mapped agns," *Astrophysical Journal*, vol. 697, no. 1, pp. 160–181, 2009.

[91] H. Falcke and P. L. Biermann, "The jet-disk symbiosis. I. Radio to X-ray emission models for quasars," *Astronomy and Astrophysics*, vol. 293, pp. 665–682, 1995.

[92] S. Heinz and R. A. Sunyaev, "The non-linear dependence of flux on black hole mass and accretion rate in core-dominated jets," *Monthly Notices of the Royal Astronomical Society*, vol. 343, no. 3, pp. L59–L64, 2003.

[93] E. Körding, H. Falcke, and S. Corbel, "Refining the fundamental plane of accreting black holes," *Astronomy and Astrophysics*, vol. 456, no. 2, pp. 439–450, 2006.

[94] J. L. Turner and P. T. P. Ho, "Bright radio continuum emission from star formation in the cores of nearby spiral galaxies," *Astrophysical Journal*, vol. 421, no. 1, pp. 122–139, 1994.

[95] N. Smith, J. M. Silverman, R. Chornock et al., "Coronal lines and dust formation in SN 2005ip: not the brightest, but the hottest type IIn supernova," *Astrophysical Journal*, vol. 695, no. 2, pp. 1334–1350, 2009.

[96] D. Pooley, S. Immler, and A. V. Filippenko, "Chandra observation of SN 2005kd: very luminous and hard X-ray emission," *The Astronomer's Telegram*, vol. 1023, p. 1, 2007.

[97] S. Immler, P. J. Brown, P. Milne et al., "X-ray, UV, and optical observations of supernova 2006bp with Swift: detection of early X-ray emission," *Astrophysical Journal*, vol. 664, no. 1 I, pp. 435–442, 2007.

[98] S. Komossa, H. Zhou, A. Rau et al., "NTT, spitzer, and chandra spectroscopy of SDSSJ095209.56+214313.3: the most luminous coronal-line supernova ever observed, or a stellar tidal disruption event?" *Astrophysical Journal*, vol. 701, no. 1, pp. 105–121, 2009.

[99] S. D. Van Dyk, K. W. Weiler, R. A. Sramek, and N. Panagia, "SN 1988Z: the most distant radio supernova," *Astrophysical Journal*, vol. 419, no. 2, pp. L69–L72, 1993.

[100] C. L. Williams, N. Panagia, S. D. Van Dyk, C. K. Lacey, K. W. Weiler, and R. A. Sramek, "Radio emission from SN 1988Z and very massive star evolution," *Astrophysical Journal*, vol. 581, no. 1 I, pp. 396–403, 2002.

[101] P. Chandra, V. V. Dwarkadas, A. Ray, S. Immler, and D. Pooley, "X-rays from the explosion site: 15 years of light curves of SN 1993J," *Astrophysical Journal*, vol. 699, no. 1, pp. 388–399, 2009.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 782528, 11 pages
doi:10.1155/2012/782528

Research Article

The Low-Mass End of the $M_{\mathrm{BH}}/M_{\mathrm{host}}$ Relation in Quasars

Roberto Decarli,[1] Renato Falomo,[2] Jari K. Kotilainen,[3] Tomi Hyvönen,[3] Michela Uslenghi,[4] and Aldo Treves[4,5,6]

[1] *Galaxy and Cosmology Department, Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany*
[2] *Osservatorio Astronomico di Padova, INAF, Vicolo dell' Osservatorio 5, 35122 Padova, Italy*
[3] *Finnish Centre for Astronomy with ESO (FINCA), University of Turku, Väisäläntie 20, 21500 Piikkiö, Finland*
[4] *INAF-IASF Milano, Via E. Bassini 15, 20133 Milano, Italy*
[5] *Dipartimento di Fisica e Matematica, Università degli Studi dell'Insubria, Via Valleggio 11, 22100 Como, Italy*
[6] *INFN Milano-Bicocca, Dipartimento di Fisica University, Milano Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

Correspondence should be addressed to Roberto Decarli, decarli@mpia.de

Received 28 July 2011; Accepted 4 October 2011

Academic Editor: Angela Bongiorno

The M_{BH}-M_{host} relation in quasars has been probed only in a limited parameter space, namely, at $M_{\mathrm{BH}} \sim 10^9\,\mathrm{M}_\odot$ and $M_{\mathrm{host}} \sim 10^{12}\,\mathrm{M}_\odot$. Here we present a study of 26 quasars laying in the low-mass end of the relation, down to $M_{\mathrm{BH}} \sim 10^7\,\mathrm{M}_\odot$. We selected quasars from the SDSS and HST-FOS archives, requiring modest M_{BH} (as derived through the virial paradigm). We imaged our sources in H band from the Nordic Optical Telescope. The quasar host galaxies have been resolved in 25 out of 26 observed targets. Host galaxy luminosities and stellar masses are computed, under reasonable assumptions on their star formation histories. Combining these results with those from our previous studies, we manage to extend the sampled parameter space of the M_{BH}-M_{host} relation in quasars. The relation holds over 2 dex in both the parameters. For the first time, we are able to measure the slope of the M_{BH}-M_{host} relation in quasars. We find that it is consistent with the linear case (similarly to what observed in quiescent galaxies). We do not find any evidence of a population of massive black holes lying below the relation.

1. Introduction

Massive black holes (BHs) are ubiquitously found in the centre of massive galaxies [1–3]. Their masses (M_{BH}) show strong correlations with large-scale properties of the host galaxies, namely, the stellar velocity dispersion (σ_*, [4–7]), the luminosity, and mass of the spheroidal component (L_{host}, M_{host}; see [8–10]). These relations have been interpreted as the outcome of a joint evolution between BHs and their host galaxies [11–18]. In this scenario, the growth of BHs through accretion regulates the gas cooling in the outskirt of host galaxies through energy or momentum injection (feedback), thus quenching the formation of stars. Galaxy mergers may also play a role in this scenario, as gravitationally induced dynamical instabilities may trigger both star formation bursts and gas inflows fuelling the BH activity ([12, 19–21], see also [22]).

The M_{BH}-host galaxy relations have been pinned down on an albeit small set of local, mostly quiescent galaxies. The sampled parameter space ranges over 3 dex in masses, from a few million to a few billion solar masses in terms of M_{BH}. Extending these studies beyond the local Universe is challenging. On one side, the influence radius of BHs, R_{inf}, that is, the radius where the gravitational potential is dominated by the singularity, is resolved only in very nearby objects (distances < few tens Mpc) with high M_{BH} values. For any other sources, indirect tracers of M_{BH} are required. The most commonly adopted indirect estimator of M_{BH} is based on the width of broad emission lines and the size of the broad line region (BLR). This can be done only in type-1 AGN, where broad lines are observed [23–26]. This approach allows to estimate M_{BH} from single-epoch spectra in ∼100 000 quasars up to $z \sim 5$ from SDSS spectra [27], and in most of the $z \sim 6$ quasars known to day [28–31]. (Caveats to this technique

arise as BLR clouds may be supported by radiation pressure [32] or move nonvirially; projection effects depending on the geometry and orientation of the BLR may hinder our ability to actually measure the orbital velocity of clouds [33, 34]; different emission lines may be produced in regions where the gas dynamics are different [35–39].)

On the other side, the properties of host galaxies are hard to measure in distant sources. Bright active nuclei (necessary to measure $M_{\rm BH}$) can easily outshine the light of their host galaxies. Observations in excellent seeing conditions (e.g., [40, 41])or based on adaptive optics (e.g., [42, 43]) are required. Diffraction-limited observations with HST have also significantly contributed in this field [44–54], although some concerns about the reproducibility of the PSF have been arisen [51]. Up to now, $\sim$ 300 quasar host galaxies have been resolved up to $z \approx 3$, and most of them at $z < 0.5$ (see [40], and references therein).

In order to understand the processes and timescales which led to the onset of the BH-host galaxy relations, two key observational tests are required. The first one consists in *tracing the evolution of the* BH*-to-host galaxy mass ratio* ($\Gamma \equiv M_{\rm BH}/M_{\rm host}$) *as a function of Cosmic Time*. If, for example, $\Gamma(z > 0) < \Gamma(z \approx 0)$, then we can argue that the BHs in these systems still have to accrete in already formed bulges. Vice versa, $\Gamma(z > 0) > \Gamma(z \approx 0)$ could suggest a rapid growth of the BHs, followed by a slower build-up of the spheroids. Most of the studies on the evolution of the BH-host galaxy relations suggest that at high redshift, for a given mass of the host spheroid, the harbored BH is more massive than at low-z [50, 53, 55–58], with $\Gamma(z) \propto (1+z)^{3/2}$. It is interesting to note that the host galaxy of J1148 + 5251, the highest-z SDSS quasar, at $z = 6.42$, shows an $M_{\rm BH}$-host galaxy dynamical mass ratio of $\approx$ 0.13 [59, 60], in an order-of-magnitude agreement with the extrapolation from the $z < 3$ studies.

A second observational test to probe the onset of the BH-host galaxy relations is to *trace the low-mass end of the BH-host galaxy relations*. Different initial host galaxy mass, BH seed mass, and build-up processes produce different slopes of the relations, especially at low mass. Light seeds ($10 - 100\,{\rm M}_\odot$) are expected as the remnants of metal-pure stars in the early Universe, while heavier seeds (up to $\sim 10^5\,{\rm M}_\odot$) can result from the direct collapse of primordial gas clouds. The former ones would produce a larger scatter in the $M_{\rm BH}$-host relations, a higher occupation fration in relatively small galaxies, and a lower cutoff in the minimum $M_{\rm BH}$ with respect to the latter (see, e.g., [61], and references therein). Some authors even claim that the $M_{\rm BH} - \sigma_*$ relation itself is just the upper limit of a broader distribution, with a number of (hard-to-detect) modest-mass BHs embedded in relatively massive galaxies (e.g., [62]).

The low-mass end of the $M_{\rm BH}$-host galaxy relations has been probed down to few 10^5 $M_{\rm BH}$ in quiescent galaxies or in low-luminosity AGN at low-z [54, 63–67], suggesting that the $M_{\rm BH}$-host relations hold in quiescent or mildly active galaxies. However, no effort has been attempted so far to extend this test to higher luminosity AGN. Quasars are ideal probes of the BH-host galaxy relations at $z > 0$. However, while Γ in quasars has been measured up to very high redshift, the ranges of $M_{\rm BH}$ and $M_{\rm host}$ investigated up to date are limited, and comparable with the observed scatter of the BH-host galaxy relations. Filling the low-mass end of the $M_{\rm BH}$-$M_{\rm host}$ relation therefore represents a main challenge and a fundamental step in our comprehension of the BH-host galaxy evolution. In this paper, we present ground-based NIR observations of quasars at $z < 0.5$ selected so that virial $M_{\rm BH} < 10^9$ $M_{\rm BH}$. Our imaging campaign successfully resolved 25 quasar host galaxies. This enables us to directly probe the slope of the $M_{\rm BH}$-$M_{\rm host}$ relation in quasars.

The structure of this work is the following: in Section 2 we describe the sample. In Section 3 we present the analysis of the spectra and we derive $M_{\rm BH}$ in all our sources. The new observations, the data reduction and analysis, and the results from the imaging campaign are presented in Section 4. In Section 5 we discuss our results. Conclusions are summarized in Section 6. Throughout the paper we will assume a standard cosmology with $H_0 = 70\,{\rm kms}^{-1}\,{\rm Mpc}^{-1}$, $\Omega_{\rm m} = 0.3$, and $\Omega_\Lambda = 0.7$.

2. The Sample

We selected quasars at $z < 0.5$ with available Hβ observations in the Sloan Digital Sky Survey (SDSS [68]) spectroscopic database or $C_{\rm IV}$ or $Mg_{\rm II}$ observations in the HST-Faint Object Spectrograph (FOS) archive. We require that black hole virial masses, computed as described in Section 3, range between 10^7 and $10^9\,{\rm M}_\odot$. Note that, out of the 62 $z < 0.5$ quasars examined in our previous study [26], only 8 (13%) had $M_{\rm BH} < 10^9\,{\rm M}_\odot$. We then selected quasars having at least 3 relatively bright ($m_H = 10$–15 mag) stars within $2'$ (corresponding to the half size of the NOTCam field of view), and a number of fainter stars in order to have an accurate description of the Point Spread Function (PSF). Observability constraints and modest weather losses further limited our analysis to 26 targets (see Table 1), mostly detected in radio wavelengths according to the [69] catalogue. This new sample is then matched with the 62 $z < 0.5$ targets presented in our previous study of the $M_{\rm BH}$-$M_{\rm host}$ relation [26, 58]. Figure 1 compares the distribution of $M_{\rm BH}$ in the general SDSS sample at $z < 0.5$, computed from the continuum luminosities and Hβ width estimates reported by [27], with the ones presented in [26] and in the present study.

3. The Spectroscopic Dataset

The spectroscopic dataset consists of pipeline-processed, publicly available spectra from the SDSS or FOS archive. SDSS spectra have $\lambda/\Delta\lambda \sim 2000$ and a spectral range between 3800 and 9000 Å. Uncertainties on wavelength calibration amount to 0.05 Å, while flux calibration formal errors account to 5%. FOS spectra are taken from the compilation of recalibrated quasar and AGN FOS spectra by [70]. Observations were performed with a number of different gratings at both high spectral resolution (1–6 Å diode^{-1}, $\lambda/\Delta\lambda \approx 1300$) and low spectral resolution (6–25 Å diode^{-1}, $\lambda/\Delta\lambda \approx 250$) covering various spectral windows from 1140 Å to 3275 Å. Photometric uncertainties are usually 5–10%, while typical wavelength calibration uncertainties are around 0.5 channels rms (see [70]).

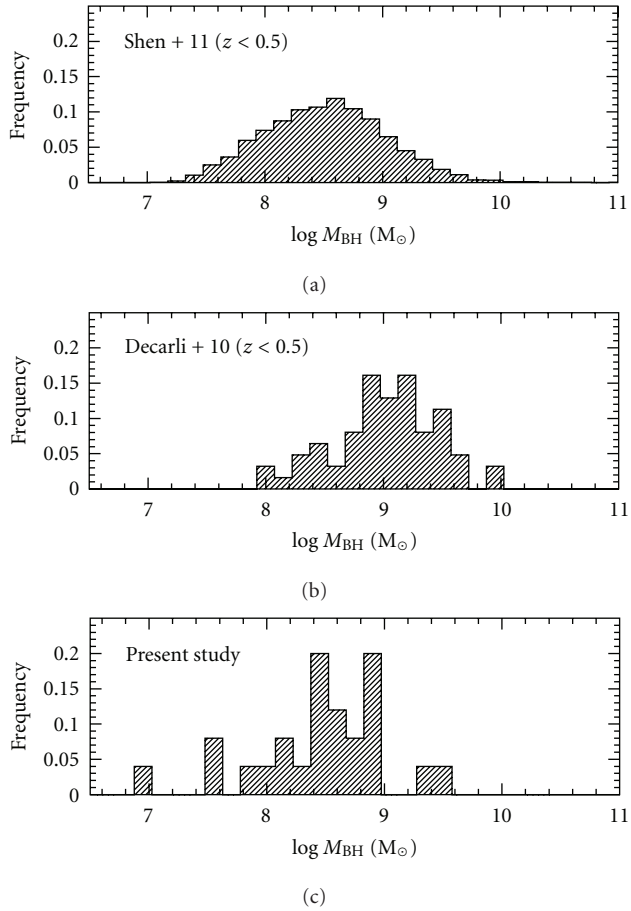


Figure 1: Comparison among the M_{BH} distributions in the SDSS quasars at $z < 0.5$ [27] (a), in our 2010 study (b) and in the present work (c). The latter is clearly more representative of the general quasar population at low redshift.

Table 1: The sample. (1) Target name. (2-3) Right ascension and declination (J2000). (4) Is the target detected in radio wavelengths, according to the [69] catalogue? (5) Catalogue redshift. (6) Apparent visual magnitude of the quasar.

Name	RA	DEC	Radio?	z	V [mag]
(1)	(2)	(3)	(4)	(5)	(6)
PB5723	00 05 47.5	+02 03 02	N	0.234	16.60
PG0026 + 12	00 29 13.7	+13 16 04	Y	0.145	15.41
PG0052 + 251	00 54 52.1	+25 25 39	Y	0.155	15.43
B20110 + 29	01 13 24.2	+29 58 16	Y	0.363	17.00
PKS0214 + 10	02 17 07.6	+11 04 10	Y	0.408	16.46
J02321 + 0008	02 32 11.8	+00 08 03	Y	0.432	19.10
J02331 − 0909	02 33 10.6	−09 09 40	Y	0.388	18.45
J03010 + 0004	03 01 00.2	+00 04 29	Y	0.486	19.33
J03323 + 0106	03 32 18.0	+01 06 48	N	0.482	18.91
J03579 − 0550	03 57 59.0	−05 50 15	Y	0.439	18.93
B20752 + 25*A*	07 55 37.0	+25 42 39	Y	0.446	18.00
J08044 + 1904	08 04 42.1	+19 04 26	Y	0.346	19.40
J08285 + 2748	08 28 53.5	+27 48 33	N	0.330	20.00
J08305 + 0802	08 30 57.4	+08 02 34	Y	0.319	19.20
PG0844 + 349	08 47 42.4	+34 45 03	Y	0.064	14.50
J09010 + 3538	09 01 00.9	+35 38 09	N	0.302	19.20
PG0947 + 396	09 50 48.3	+39 26 51	Y	0.206	16.39
PG0953 + 415	09 56 52.4	+41 15 23	Y	0.234	15.32
TON1187	10 13 03.1	+35 51 23	N	0.079	14.75
TEX1156 + 213	11 59 26.2	+21 06 56	Y	0.349	16.90
Q1214 + 1804	12 16 49.0	+17 48 04	N	0.374	17.30
PG1404 + 226	14 06 21.9	+22 23 47	Y	0.098	15.82
PG1415 + 451	14 17 00.8	+44 56 06	Y	0.114	15.24
PG1626 + 554	16 27 56.1	+55 22 31	Y	0.132	15.68
4C73.18	19 27 48.5	+73 58 02	Y	0.302	16.50
PKS2251 + 11	22 54 10.4	+11 36 39	Y	0.325	15.82

Spectral analysis follows the same approach described in [26] and [31]. Briefly, continua are modeled as a superposition of a power-law component, the host galaxy emission (only in the optical spectra; the Elliptical model by [71] was adopted as a template), and the Fe_{II} multiplets (modeled on the template by [72]). Relevant broad emission lines (Hβ, Mg_{II}, and C_{IV}) are modeled with the superposition of two Gaussian curves with the same peak wavelength. The spectral resolution of the SDSS data is adequate to allow an easy identification of the narrow Hβ emission with respect to the broad line. Examples of the line fitting are shown in Figure 2.

We measure the continuum luminosity at 1350, 3000, and 5100 Å from the fitted power-law. The broad line luminosity and FWHM are measured from the line model (see [36], for a discussion on the fitting technique and the parametrization of the relevant quantities). Virial black hole masses are computed with the same recipes used in [26], adopting geometrical factors of 1.6 for Hβ and Mg_{II} and 2.4 for C_{IV}. Table 2 lists the main measurements and inferred quantities from the spectroscopic analysis. (We note that, after a careful reanalysis of the spectra, two objects (B20110+29 and B20752 + 25*A*) show values of M_{BH} slightly exceeding the initial selection criteria. Nevertheless, we include these sources in the present analysis.)

4. The Imaging Dataset

All the objects in our study have been observed in *H* band in a campaign at the 2.5 m Nordic Optical Telescope (Roque de Los Muchachos, Spain) using NOTCam. Observations were carried out during three observing runs in May 2007 and April and November 2008. The average seeing in *H* band was 0.7″. The 1024 × 1024 pixel NOTCam detector has a pixel scale of 0.235″ pxl^{-1}, yielding a field of view size of ∼4 × 4 $arcmin^2$. Usual observing techniques for broad-band NIR imaging of point-like sources were adopted. Observations were split in 1 min long individual frames. Random jittering patterns within a 20″ box were adopted in order to perform optimal sky subtraction. The total time on each source was 45 min.

Data reduction was performed using IRAF. (IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement

Table 2: Results from the spectroscopic analysis. (1) Quasar name. (2) Redshift. (3) Line used in the M_{BH} estimate. (4) Continuum monochromatic luminosity at 1350 (for C_{IV}), 3000 (for Mg_{II}) or 5100 Å (for Hβ), in ergs^{-1}. (5) Line FWHM in km s^{-1}. (6) Virial estimate of the black hole mass, in solar units.

Name	z	Line	$\log \lambda L_\lambda$ [erg s^{-1}]	FWHM [km s^{-1}]	$\log M_{BH}$ [$M_\odot$]
(1)	(2)	(3)	(4)	(5)	(6)
PB5723	0.234	C_{IV}	44.71	3715	8.15
PG0026 + 12	0.145	C_{IV}	45.22	2062	7.92
PG0052 + 251	0.155	C_{IV}	45.33	5914	8.90
B20110 + 29	0.363	Hβ	44.81	6149	9.33
PKS0214 + 10	0.408	C_{IV}	45.71	4122	8.79
J02321 + 0008	0.432	Hβ	44.41	1727	7.56
J02331 − 0909	0.388	Hβ	44.63	1863	7.77
J03010 + 0004	0.486	Hβ	44.45	6634	8.76
J03323 + 0106	0.482	Hβ	44.77	4282	8.59
J03579 − 0550	0.439	Hβ	44.62	4005	8.43
B20752 + 25A	0.446	Hβ	45.06	9738	9.50
J08044 + 1904	0.346	Hβ	44.09	7322	8.60
J08285 + 2748	0.330	Hβ	44.14	5385	8.37
J08305 + 0802	0.319	Hβ	44.14	6149	8.48
PG0844 + 349	0.064	Mg_{II}	44.57	3209	8.20
J09010 + 3538	0.302	Hβ	44.29	8495	8.86
PG0947 + 396	0.206	C_{IV}	45.17	4090	8.49
PG0953 + 415	0.234	C_{IV}	45.45	3490	8.50
TON1187	0.079	Hβ	44.20	2141	7.60
TEX1156 + 213	0.349	Hβ	44.93	5663	8.94
Q1214 + 1804	0.374	Hβ	45.04	3728	8.65
PG1404 + 226	0.098	Hβ	44.14	1036	6.93
PG1415 + 451	0.114	Hβ	44.19	3244	7.96
PG1626 + 554	0.132	C_{IV}	45.18	4057	8.49
4C73.18	0.302	C_{IV}	45.92	4155	8.91
PKS2251 + 11	0.325	C_{IV}	45.54	5028	8.87

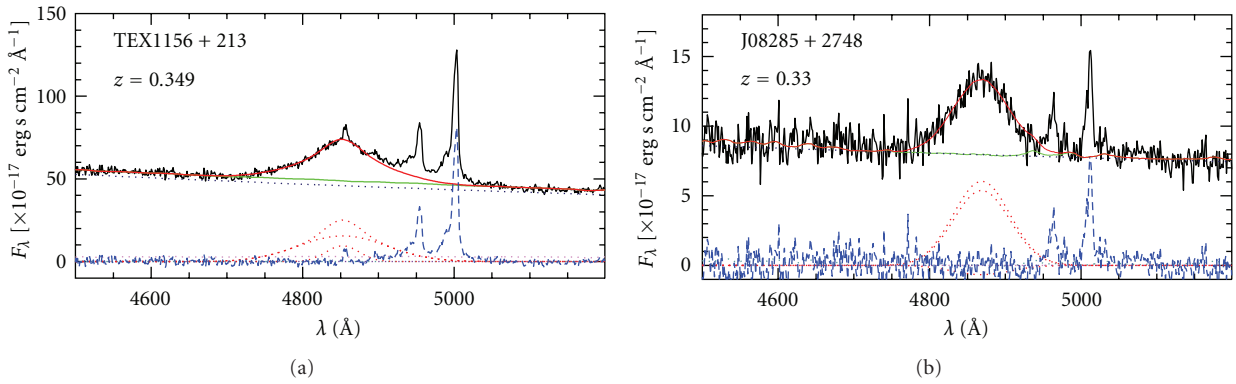


Figure 2: Examples of Hβ line fitting. The observed spectra, shifted to the rest frame, are plotted in solid, black lines. The various components of the models are shown: the power-law (black, dotted line), the host galaxy and Fe_{II} templates (green, solid line), the line model (red, solid line), and its components (red, dotted lines). Fit residuals are shown in blue, dashed lines.

with the National Science Foundation.) Bad pixels were corrected for in each image using a mask made from the ratio of two sky flats with different illumination levels. Sky subtraction was performed for each science image using a median averaged frame of all the other temporally close frames in a grid of eight exposures. Flat fielding was made using normalized median averaged twilight sky frames with different illumination levels. Finally, images were aligned to subpixel accuracy using field stars as reference points and combined after removing spurious pixel values to obtain the final reduced coadded image. Zero point calibration is achieved by cross-matching the photometry of field stars with the 2MASS database. This approach yields typical photometric uncertainties <0.1 mag.

We analyzed our data using the Astronomical Image Decomposition and Analysis package (AIDA; [73]), an IDL-based software designed to perform two-dimensional model fitting of quasar images. Details on the procedure are presented in [40, 74] and briefly summarized here.

A careful modeling of the Point Spread Function (PSF) is crucial to disentangle the extended host galaxy light from the nuclear source. To model the PSF shape, we used suitable field stars. Each star was modeled with four two-dimensional Gaussians, representing the core of the PSF, and an exponential feature, representing the extended wing of the PSF. Regions contaminated by nearby sources, saturated pixels, and other defects affecting the images were masked out.

In order to discriminate between resolved and unresolved targets, we first fit the images of our sources with the pure PSF model. In most of the cases, an extended halo was clearly observed in the residuals. We then reperformed the fits using a Sersic law, (describing the host galaxy) plus a point-source (the nucleus), convolved to the PSF model. In all but one case (J02331 − 0909), this second fit is significantly better than the fit with the pure PSF, as confirmed by the χ^2 ratio between the two fits (see Table 3).

An example of the outputs of our analysis is shown in Figure 3.

Host galaxy apparent H-band magnitudes are then converted into rest-frame R-band. We use the Elliptical galaxy template by [71] to estimate k-corrections. The host galaxy R-band absolute magnitude is then converted into a stellar mass by adopting the mass-to-light ratio (M/L) of a single stellar population originated at $z_{\rm burst} = 5$ and passively evolving down to $z = 0$ (see [58], for details). Table 3 summarizes the relevant results from the modeling of the quasar host galaxies described here.

5. Discussion

In Figure 4 we show the $M_{\rm BH}$-$M_{\rm host}$ relation for quasar host galaxies at $z < 0.5$. The dataset (62 objects from [26, 58], plus 25 objects with resolved host galaxies from the present study) span over 2 dex in both $M_{\rm BH}$ and $M_{\rm host}$. The same $M_{\rm BH}$-$M_{\rm host}$ relation observed for inactive galaxies appears to hold through all the sampled range, from $\sim 3 \times 10^{10}$ to $\sim 3 \times 10^{12}\,{\rm M}_\odot$ in terms of $M_{\rm host}$. We find that $\langle \log \Gamma \rangle = -2.843$ (in excellent agreement with the $M_{\rm BH} = 0.0015\, M_{\rm host}$ value reported by [9], for inactive galaxies) with a 0.44 dex scatter. Only 3 sources (J02321+0008, PG1404+226 from this study; 1001 + 291 from the old sample) lie more than 2-σ *below* the relation. Since the sampled parameter space is about 5 times larger than the dispersion of the relation, we can exclude that the observed $M_{\rm BH}$-$M_{\rm host}$ relation is the upper envelope of a population of quasars with relatively small black holes hosted by very massive galaxies. The best bilinear regression fit of the relation is

$$\log \frac{M_{\rm BH}}{10^9\,{\rm M}_\odot} = (1.26 \pm 0.29) \times \log \frac{M_{\rm host}}{7 \cdot 10^{11}\,{\rm M}_\odot} + (0.04 \pm 0.03), \quad (1)$$

consistent with the relations with a constant $M_{\rm BH}/M_{\rm host}$ ratio, as observed in quiescent galaxies in the local Universe [9, 10].

When considering subsets of our data, the high-mass end shows a slightly smaller scatter (see Table 4).

For twelve objects the host galaxies are found to be best described by a Sersic law with small index ($n_{\rm s} < 1.5$), suggesting the presence of significant disc components. In particular, PG1404 + 226 (incidentally, the object showing the smallest $M_{\rm BH}$ and the largest deviation with respect to the $M_{\rm BH}$-$M_{\rm host}$ relation in our sample) shows clear spiral arms in the residuals of the host galaxy model. From local galaxy studies, $M_{\rm BH}$ is found to be more sensitive to the properties of the spheroidal stellar component rather than of the whole galaxy. On the other hand, a bulge + disc decomposition is practically impossible with ground-based images of quasar host galaxies at relatively high redshift. Here we attempt a rule-of-thumb correction starting from the Sersic index value. We assume that the bulge-to-total luminosity ratio in the rest-frame R band, B/T, scales with the Sersic index as follows:

$$B/T = \begin{cases} (n_{\rm s} - 0.5)/3.5 & \text{if } n_{\rm s} < 4, \\ 1 & \text{if } n_{\rm s} \geq 4. \end{cases} \quad (2)$$

This simple analytical form roughly traces the bulk of the B/T values for $n_{\rm s} < 4$, as found by [75], who performed accurate image deconvolution for ∼1 million galaxies from the SDSS. Furthermore, it is consistent with the operative hypothesis that all the galaxies well described by a de Vaucouleurs profile ($n_{\rm s} = 4$) are bulge dominated ($B/T \approx 1$), as assumed in our previous study. The effect of this correction is to move disc-dominated host galaxies towards the left side of Figure 4. In particular, all but one source at $M_{\rm bulge} < 10^{10}\,{\rm M}_\odot$ would lie *above* the local relation. The best fit relation is indeed flatter (0.88 ± 0.18 instead of 1.26 ± 0.29), but still consistent with the linear case. The scatter is also increased (0.53 dex, computed over the whole sample; 0.61 and 0.55 dex in the small $M_{\rm BH}$ and small $M_{\rm bulge}$ subsets, resp.). We stress that the correction reported in (2) is uncertain, because of the wide range of B/T values reported for any given $n_{\rm s}$. However, we remark that any correction for the B/T would make the case against a population of under-massive black holes in very massive galaxies even stronger.

A similar argument can be used to evaluate how our results are affected by different assumptions on the star formation history. In our study, we adopted the mass-to-light ratio (M/L) of a single stellar population originated at $z_{\rm burst} = 5$ and passively evolving down to $z = 0$. However, objects with significant disc contaminations are expected to have a younger stellar population than old, passive spheroids. This would imply smaller M/L, that is, less massive host galaxies for a given observed host luminosity. This would make the case against a quasar population lying *below* the observed $M_{\rm BH}$-$M_{\rm host}$ relation even more robust.

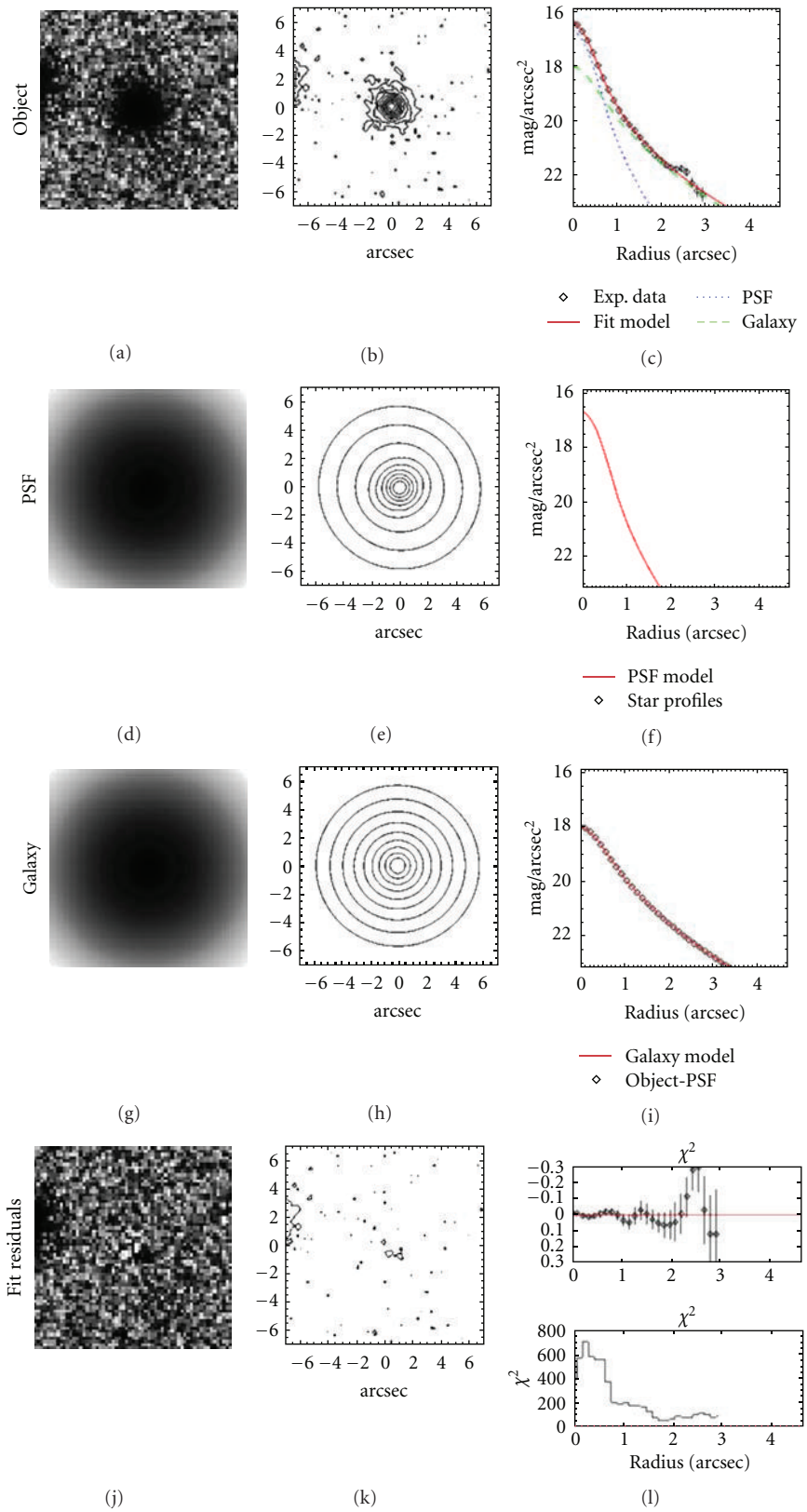


FIGURE 3: An example of the analysis of the quasar host galaxies, shown in the case of the quasar J08285 + 2748. *Top panels:* observed quasar image (left), contour plot (center), and light profile (right; the PSF and galaxy models are plotted in dotted and dashed lines, resp.). *Central panels:* similarly, the images (left), contour plots (center), and light profile (right) of the PSF and the galaxy model. *Bottom panels:* image (left), contour plot (center), and light profile (top right) of residuals after model subtraction. The radial distribution of χ^2 values is also plotted (bottom right).

Table 3: Results from the imaging analysis. (1) Quasar name. (2) Redshift. (3) Apparent observed H-band magnitude of the nucleus. (4) Apparent observed H-band magnitude of the host galaxy. (5) Apparent central surface brightness, as extrapolated from the host galaxy model. (6) Chi-square ratio between the best fit using a pure PSF and the best fit using a galaxy+quasar model. (7) Effective radius of the host galaxy. (8) Ellipticity of the host galaxy model. (9) Sersic index of the host galaxy model. (10) Rest-frame R-observed H color correction. (11) Resulting absolute rest-frame R-band magnitude of the host galaxy. (12) Adopted mass-to-light ratio. (13) Stellar mass of the host galaxy. (14) BH-to-host galaxy mass ratio.

Name	z	m_{nuc} [mag]	m_{host} [mag]	μ_0 [mag $''^{-2}$]	χ_{psf}/χ_{gq}	R_e	Ell[$''$]	n_s	R-H [mag]	M_R [mag]	log M/L [$M_\odot/L_\odot$]	log M_{host}[$M_\odot$]	log Γ
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11	(12)	(13)	(14)
PB5723	0.234	15.99	15.81	5.41	1.41	0.32	0.57	5.00	2.59	−21.93	0.71	11.38	−3.61
PG0026 + 12	0.145	12.90	14.94	16.10	1.77	1.44	0.15	1.20	2.58	−21.64	0.75	11.30	−3.76
PG0052 + 251	0.155	13.46	14.37	16.34	6.55	2.19	0.19	1.22	2.58	−22.38	0.74	11.59	−3.07
B20110 + 29	0.363	16.04	16.21	12.07	2.51	3.89	0.09	5.00	2.60	−22.61	0.67	11.61	−2.48
PKS0214 + 10	0.408	14.60	16.10	13.98	1.88	1.16	0.00	2.77	2.59	−23.03	0.65	11.77	−3.36
J02321 + 0008	0.432	99.90	16.80	9.93	1.43	0.37	0.00	3.89	2.58	−22.50	0.64	11.54	−4.18
J02331 − 0909	0.388	15.98	—	—	1.01	—	—	—	—	—	—	—	—
J03010 + 0004	0.486	99.90	17.35	13.42	1.17	0.26	0.24	1.89	2.57	−22.25	0.63	11.43	−2.87
J03323 + 0106	0.482	18.66	17.15	11.55	1.91	0.54	0.61	3.17	2.57	−22.43	0.63	11.50	−3.11
J03579 − 0550	0.439	18.78	17.22	8.06	1.14	0.45	0.35	5.00	2.58	−22.12	0.64	11.39	−3.16
B20752 + 25A	0.446	14.80	16.14	12.51	1.39	0.71	0.07	2.95	2.58	−23.24	0.64	11.83	−2.53
J08044 + 1904	0.346	17.07	17.40	17.73	1.55	0.98	0.37	1.00	2.60	−21.30	0.67	11.09	−2.69
J08285 + 2748	0.330	16.98	17.20	16.63	1.58	0.81	0.06	1.53	2.60	−21.38	0.68	11.13	−2.96
J08305 + 0802	0.319	17.50	18.25	17.66	1.19	0.57	0.33	0.90	2.60	−20.25	0.68	10.68	−2.40
PG0844 + 349	0.064	13.34	14.59	17.55	1.92	2.82	0.27	0.90	2.56	−20.13	0.78	10.73	−2.73
J09010 + 3538	0.302	17.08	16.35	14.10	2.79	0.91	0.20	2.43	2.60	−22.01	0.69	11.39	−2.73
PG0947 + 396	0.206	14.42	15.18	16.17	4.45	1.92	0.20	1.59	2.59	−22.24	0.72	11.52	−3.41
PG0953 + 415	0.234	12.93	15.80	19.36	1.33	3.39	0.12	0.90	2.59	−21.93	0.71	11.38	−3.26
TON1187	0.079	13.97	14.95	16.83	3.70	1.91	0.20	1.10	2.56	−20.25	0.77	10.77	−3.37
TEX1156 + 213	0.349	15.38	15.78	14.86	2.50	1.07	0.12	1.98	2.60	−22.95	0.67	11.75	−3.01
Q1214 + 1804	0.374	16.36	16.48	12.05	1.11	0.25	0.40	1.97	2.59	−22.44	0.66	11.54	−3.09
PG1404 + 226	0.098	14.35	14.70	17.08	7.62	2.99	0.52	1.04	2.57	−20.97	0.77	11.05	−4.32
PG1415 + 451	0.114	13.65	14.02	15.22	10.56	1.50	0.04	1.30	2.57	−22.01	0.76	11.46	−3.70
PG1626 + 554	0.132	13.80	14.78	16.23	8.55	2.04	0.04	1.53	2.57	−21.60	0.75	11.29	−3.18
4C73.18	0.302	13.55	16.42	17.58	1.44	1.16	0.18	0.90	2.60	−21.94	0.69	11.36	−2.83
PKS2251 + 11	0.325	13.53	15.70	18.62	2.94	2.43	0.05	0.90	2.60	−22.86	0.68	11.72	−3.23

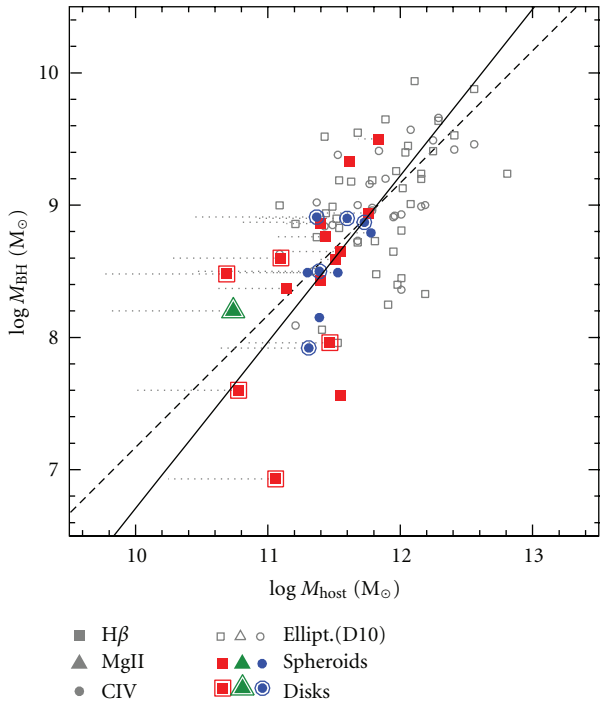


Figure 4: The M_{BH}-M_{host} relation in the objects in our sample (filled symbols), as compared with the $z < 0.5$ sample from [58] (empty symbols) and the local relation ($M_{BH}/M_{host} = 0.0015$; see [9]). Squares, triangles, and circles refer to M_{BH} estimates derived from Hβ, Mg_{II}, and C_{IV} respectively. Our new data substantially extend the sampled range of the M_{BH}-M_{host} relation. The relation holds down to quasar host masses of $\sim 10^{11}\,M_\odot$ (BH masses of $\sim 10^{8}\,M_\odot$). There is no evidence of a population of objects lying *below* the relation, as claimed by [62]. The overall scatter of the M_{BH}-M_{host} relation is 0.44 dex, and it extends over 2 dex both in terms of host galaxy and BH mass. Dotted, horizontal lines show how our objects would move if the correction for the bulge-to-total luminosity ratio is taken into account (see the text for details). The solid line is the best fit to our data.

Table 4: Average values of the M_{BH}/M_{host} ratio in various subsets of our sample. (1) Subsample. (2) Number of objects. (3) Average value of log Γ. (4) Root Mean Square of log Γ.

Subsample	N	⟨log Γ⟩ dex	RMS dex
(1)	(2)	(3)	(4)
Without B/T correction			
ALL	87	−2.842	0.444
$M_{BH} \geq 10^9\,M_\odot$	33	−2.627	0.351
$M_{BH} < 10^9\,M_\odot$	54	−2.974	0.445
$M_{host} \geq 4 \times 10^{11}\,M_\odot$	49	−2.889	0.401
$M_{host} < 4 \times 10^{11}\,M_\odot$	38	−2.782	0.492
With B/T correction			
ALL	87	−2.697	0.524
$M_{BH} \geq 10^9\,M_\odot$	33	−2.623	0.356
$M_{BH} < 10^9\,M_\odot$	54	−2.886	0.603
$M_{bulge} \geq 4 \times 10^{11}\,M_\odot$	46	−2.886	0.419
$M_{bulge} < 4 \times 10^{11}\,M_\odot$	41	−2.485	0.554

6. Conclusions

We measured black hole masses and host galaxy luminosities in a sample of 25 low-redshift ($z < 0.5$) quasars selected to have modest ($< 10^9\,M_\odot$) black hole masses.

For each object we inferred stellar masses. This allowed us to significantly expand the sampled range of M_{host} and M_{BH} for quasars. We found the following.

(i) The M_{BH}-M_{host} relation holds over all the 2 dex both in terms of M_{BH} and M_{host}. The relation has a scatter of 0.44 dex; that is, the sampled parameter space is ∼5 times larger.

(ii) The slope of the M_{BH}-M_{host} relation in quasars is consistent with unity (in a log-log plane), consistently with what observed in quiescent galaxies.

(iii) The scatter of the relation increases by ∼0.9 dex at the low-mass end.

(iv) After applying a simplistic correction for the disc contribution in objects with low Sersic indexes, the slope of the M_{BH}-M_{host} relation is smaller but still consistent with the linear case.

(v) No evidence of a population of quasars with relatively modest M_{BH} and very high M_{host} values is found.

Further studies at even lower M_{BH} masses ($\lesssim 10^7\,M_\odot$) and at higher redshift could provide further constraints on the early black hole growth and the nature of the seeds, and to pin down the evolution of the M_{BH}-host galaxy relations along the Cosmic Time. This requires extremely high-quality imaging of quasar host galaxies that would become possible with the next generation of ELT and laser assisted AO imagers.

Acknowledgments

R. Decarli acknowledges funding from Germany's national research center for aeronautics and space (DLR, project FKZ 50 OR 1104), based on observations made with the Nordic Optical Telescope, operated on the island of La Palma jointly by Denmark, Finland, Iceland, Norway, and Sweden, in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. (Facilities: NOT (NOTCAM) SDSS HST).

References

[1] J. Kormendy and D. Richstone, "Inward bound—the search for supermassive black holes in galactic nuclei," *Annual Review of Astronomy and Astrophysics*, vol. 33, no. 1, pp. 581–624, 1995.

[2] R. Decarli, G. Gavazzi, I. Arosio et al., "The census of nuclear activity of late-type galaxies in the Virgo cluster," *Monthly Notices of the Royal Astronomical Society*, vol. 381, no. 1, pp. 136–150, 2007.

[3] E. Gallo, T. Treu, J. Jacob, J. H. Woo, P. J. Marshall, and R. Antonucci, "Amuse-Virgo—I. Supermassive black holes in low-mass spheroids," *Astrophysical Journal*, vol. 680, no. 1, pp. 154–168, 2008.

[4] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.

[5] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.

[6] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal*, vol. 698, no. 1, pp. 198–221, 2009.

[7] A. W. Graham, C. A. Onken, E. Athanassoula, and F. Combes, "An expanded Mbh-σ diagram, and a new calibration of active galactic nuclei masses," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2211–2228, 2011.

[8] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[9] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *Astrophysical Journal Letters*, vol. 589, no. 1, pp. L21–L24, 2003.

[10] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *Astrophysical Journal Letters*, vol. 604, no. 2, pp. L89–L92, 2004.

[11] J. Silk and M. J. Rees, "Quasars and galaxy formation," *Astronomy and Astrophysics*, vol. 331, no. 2, pp. L1–L4, 1998.

[12] G. Kauffmann and M. Haehnelt, "A unified model for the evolution of galaxies and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 311, no. 3, pp. 576–588, 2000.

[13] A. King, "The AGN-starburst connection, galactic superwinds, and M BH-σ," *Astrophysical Journal*, vol. 635, no. 2, pp. L121–L123, 2005.

[14] J. S. B. Wyithe and A. Loeb, "Constraints on the process that regulates the growth of supermassive black holes based on the intrinsic scatter in the Mbh-σsph relation," *Astrophysical Journal*, vol. 634, no. 2 I, pp. 910–920, 2005.

[15] B. Robertson, L. Hernquist, T. J. Cox et al., "The evolution of the MBH-σ relation," *Astrophysical Journal*, vol. 641, no. 1, pp. 90–102, 2006.

[16] P. F. Hopkins, L. Hernquist, T. J. Cox, B. Robertson, and E. Krause, "A theoretical interpretation of the black hole fundamental plane," *Astrophysical Journal*, vol. 669, no. 1, pp. 45–66, 2007.

[17] F. Shankar, "The demography of supermassive black holes: growing monsters at the heart of galaxies," *New Astronomy Reviews*, vol. 53, no. 4–6, pp. 57–77, 2009.

[18] R. K. Malbon, C. M. Baugh, C. S. Frenk, and C. G. Lacey, "Black hole growth in hierarchical galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 4, pp. 1394–1414, 2007.

[19] T. Di Matteo, V. Springel, and L. Ilernquist, "Energy input from quasars regulates the growth and activity of black holes and their host galaxies," *Nature*, vol. 433, no. 7026, pp. 604–607, 2005.

[20] G. Canalizo, N. Bennert, B. Jungwiert et al., "Spectacular shells in the host galaxy of the QSO MC2 1635 + 119," *Astrophysical Journal*, vol. 669, no. 2, pp. 801–809, 2007.

[21] N. Bennert, G. Canalizo, B. Jungwiert et al., "Evidence for merger remnants in early-type host galaxies of low-Redshift QSOs," *Astrophysical Journal*, vol. 677, no. 2, pp. 846–857, 2008.

[22] M. Cisternas, K. Jahnke, K. J. Inskip et al., "The bulk of the black hole growth since z∼1 occurs in a secular universe: no major merger-AGN connection," *Astrophysical Journal Letters*, vol. 726, no. 2, article 57, 2011.

[23] B. M. Peterson and A. Wandel, "Evidence for supermassive black holes in active galactic nuclei from emission-line reverberation," *Astrophysical Journal*, vol. 540, no. 1, pp. L13–L16, 2000.

[24] M. Vestergaard, "Determining central black hole masses in distant active galaxies," *Astrophysical Journal*, vol. 571, no. 2, pp. 733–752, 2002.

[25] M. Vestergaard and B. M. Peterson, "Determining central black hole masses in distant active galaxies and quasars—II. Improved optical and UV scaling relationships," *Astrophysical Journal*, vol. 641, no. 2, pp. 689–709, 2006.

[26] R. Decarli, R. Falomo, A. Treves, J. K. Kotilainen, M. Labita, and R. Scarpa, "The quasar MBH-Mhost relation through cosmic time—I. Data set and black hole masses," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 4, pp. 2441–2452, 2010.

[27] Y. Shen, G. T. Richards, M. A. Strauss et al., "A catalog of quasar properties from Sloan Digital Sky Survey Data Release 7," *Astrophysical Journal*, vol. 194, no. 2, article 45, 2011.

[28] C. J. Willott, R. J. McLure, and M. J. Jarvis, "A 3×109 M$\odot$ black hole in the quasar SDSS J1148 + 5251 at z = 6.41," *Astrophysical Journal*, vol. 587, no. 1, pp. L15–L18, 2003.

[29] J. D. Kurk, F. Walter, X. Fan et al., "Black hole masses and enrichment of z ∼ 6 SPSS quasars," *Astrophysical Journal*, vol. 669, no. 1, pp. 32–44, 2007.

[30] C. J. Willott, L. Albert, D. Arzoumanian et al., "Eddington-limited accretion and the black hole mass function at redshift 6," *Astronomical Journal*, vol. 140, no. 2, pp. 546–560, 2010.

[31] G. De Rosa, R. Decarli, F. Walter et al., "Evidence for non-evolving fe II/Mg II ratios in rapidly accreting z∼6 QSOs," *Astrophysical Journal*, vol. 739, no. 2, article 56, 2011.

[32] A. Marconi, D. J. Axon, R. Maiolino et al., "On the observed distributions of black hole masses and eddington ratios from radiation pressure corrected virial indicators," *The Astrophysical Journal Letters*, vol. 698, article L103, 2009.

[33] R. Decarli, M. Dotti, M. Fontana, and F. Haardt, "Are the black hole masses in narrow-line Seyfert 1 galaxies actually small?" *Monthly Notices of the Royal Astronomical Society*, vol. 386, no. 1, pp. L15–L19, 2008.

[34] R. Decarli, M. Dotti, and A. Treves, "Geometry and inclination of the broad-line region in blazars," *Monthly Notices of the Royal Astronomical Society*, vol. 416, no. 1, 2011.

[35] H. Netzer, P. Lira, B. Trakhtenbrot, O. Shemmer, and I. Cury, "Black hole mass and growth rate at high redshift," *Astrophysical Journal*, vol. 671, no. 2, pp. 1256–1263, 2007.

[36] R. Decarli, M. Labita, A. Treves, and R. Falomo, "On the geometry of broad emission region in quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 387, no. 3, pp. 1237–1247, 2008.

[37] M. Labita, R. Decarli, A. Treves, and R. Falomo, "Downsizing of supermassive black holes from the SDSS quasar survey—II. Extension to z ∼ 4," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 4, pp. 2099–2106, 2009.

[38] S. Fine, M. J. Jarvis, and T. Mauch, "Orientation effects in quasar spectra: the broad- and narrow-line regions," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 1, pp. 213–222, 2011.

[39] G. T. Richards, N. E. Kruczek, S. C. Gallagher et al., "Unification of luminous type 1 quasars through C_{iv} emission," *Astronomical Journal*, vol. 141, no. 5, article 167, 2011.

[40] J. K. Kotilainen, R. Falomo, R. Decarli, A. Treves, M. Uslenghi, and R. Scarpa, "The properties of quasar hosts at the peak of the quasar activity," *Astrophysical Journal*, vol. 703, no. 2, pp. 1663–1671, 2009.

[41] T. A. Targett, J. S. Dunlop, and R. J. McLure, "The host galaxies and black-hole:galaxy mass ratios of luminous quasars at z~4," *Monthly Notices of the Royal Astronomical Society*. In press.

[42] R. Falomo, J. K. Kotilainen, R. Scarpa, and A. Treves, "VLT adaptive optics imaging of QSO host galaxies and their close environment at z ~ 2.5: results from a pilot program," *Astronomy and Astrophysics*, vol. 434, no. 2, pp. 469–473, 2005.

[43] S. M. Croom, D. Schade, B. J. Boyle, T. Shanks, L. Miller, and R. J. Smith, "Gemini imaging of QSO host galaxies at z ~ 2," *Astrophysical Journal*, vol. 606, no. 1 I, pp. 126–138, 2004.

[44] J. N. Bahcall, S. Kirhakos, D. H. Saxe, and D. P. Schneider, "Hubble Space Telescope images of a sample of 20 nearby luminous quasars," *Astrophysical Journal*, vol. 479, no. 2, pp. 642–658, 1997.

[45] M. J. Kukula, J. S. Dunlop, R. J. McLure et al., "A NICMOS imaging study of high-z quasar host galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 326, no. 4, pp. 1533–1546, 2001.

[46] S. E. Ridgway, T. M. Heckman, D. Calzetti, and M. Lehnert, "NICMOS imaging of the host galaxies of z ~ 2-3 radio-quiet quasars," *Astrophysical Journal*, vol. 550, no. 1, pp. 122–141, 2001.

[47] T. S. Hamilton, S. Casertano, and D. A. Turnshek, "The luminosity function of QSO host galaxies," *Astrophysical Journal*, vol. 576, no. 1, pp. 61–74, 2002.

[48] J. S. Dunlop, R. J. McLure, M. J. Kukula, S. A. Baum, C. P. O'Dea, and D. H. Hughes, "Quasars, their host galaxies and their central black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 340, no. 4, pp. 1095–1135, 2003.

[49] D. J. E. Floyd, M. J. Kukula, J. S. Dunlop et al., "The host galaxies of luminous quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 355, no. 1, pp. 196–220, 2004.

[50] C. Y. Peng, C. D. Impey, L. C. Ho, E. J. Barton, and H. W. Rix, "Probing the coevolution of supermassive black holes and quasar host galaxies," *Astrophysical Journal*, vol. 640, no. 1, pp. 114–125, 2006.

[51] M. Kim, L. C. Ho, C. Y. Peng, A. J. Barth, and M. Im, "Decomposition of the host galaxies of active galactic nuclei using hubble space telescope images," *Astrophysical Journal*, vol. 179, no. 2, pp. 283–305, 2008.

[52] K. Jahnke, A. Bongiorno, M. Brusa et al., "Massive galaxies in cosmos: evolution of black hole versus bulge mass but not versus total stellar mass over the last 9 Gyr?" *Astrophysical Journal Letters*, vol. 706, no. 2, pp. L215–L220, 2009.

[53] V. N. Bennert, M. W. Auger, T. Treu, J.-H. Woo, and M. A. Malkan, "The relation between black hole mass and host spheroid stellar mass out to z~2," *Astrophysical Journal*, vol. 742, no. 2, 2011.

[54] Y.-F. Jiang, J. E. Greene, and L. C. Ho, "Black hole mass and bulge luminosity for low-mass black holes," *Astrophysical Journal Letters*, vol. 737, no. 2, 2011.

[55] R. J. McLure, M. J. Jarvis, T. A. Targett, J. S. Dunlop, and P. N. Best, "On the evolution of the black hole: spheroid mass ratio," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 3, pp. 1395–1403, 2006.

[56] C. Y. Peng, C. D. Impey, H. W. Rix et al., "Probing the coevolution of supermassive black holes and galaxies using gravitationally lensed quasar hosts," *Astrophysical Journal*, vol. 649, no. 2, pp. 616–634, 2006.

[57] A. Merloni, A. Bongiorno, M. Bolzonella et al., "On the cosmic evolution of the scaling relations between black holes and their host galaxies: broad-line active galactic nuclei in the zCosmos survey," *Astrophysical Journal*, vol. 708, no. 1, pp. 137–157, 2010.

[58] R. Decarli, R. Falomo, A. Treves, M. Labita, J. K. Kotilainen, and R. Scarpa, "The quasar MBH-Mhost relation through cosmic time—II. Evidence for evolution from z = 3 to the present age," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 4, pp. 2453–2461, 2010.

[59] F. Walter, F. Bertoidl, C. Carilli et al., "Molecular gas in the host galaxy of a quasar at redshift z = 6.42," *Nature*, vol. 424, no. 6947, pp. 406–408, 2003.

[60] F. Walter, C. Carilli, F. Bertoldi et al., "Resolved molecular gas in a quasar host galaxy at redshift z = 6.42," *Astrophysical Journal*, vol. 615, no. 1, pp. L17–L20, 2004.

[61] M. Volonteri and P. Natarajan, "Journey to the MBH-σ relation: the fate of low-mass black holes in the universe," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 4, pp. 1911–1918, 2009.

[62] D. Batcheldor, "The M•-$\sigma*$ relation derived from sphere of influence arguments," *Astrophysical Journal*, vol. 711, no. 2, pp. L108–L111, 2010.

[63] J. E. Greene and L. C. Ho, "Active galactic nuclei with candidate intermediate-mass black holes," *Astrophysical Journal*, vol. 610, no. 2, pp. 722–736, 2004.

[64] J. E. Greene and L. C. Ho, "A new sample of low-mass black holes in active galaxies," *Astrophysical Journal Letters*, vol. 670, no. 1, pp. 92–104, 2007.

[65] X. Dong, T. Wang, W. Yuan et al., "Broad-line AGNs with Candidate intermediate-mass black holes in the sloan digital sky survey," in *Proceedings of the Central Engine of Active Galactic Nuclei*, vol. 373 of *ASP Conference Series*, p. 57, 2007.

[66] C. E. Thornton, A. J. Barth, L. C. Ho, R. E. Rutledge, and J. E. Greene, "The host galaxy and central engine of the dwarf active galactic nucleus POX 52," *Astrophysical Journal*, vol. 686, no. 2, pp. 892–910, 2008.

[67] T. Xiao, A. J. Barth, J. E. Greene et al., "The sloan digital sky survey: technical summary," *The Astronomical Journal*, vol. 120, p. 1579, 2000.

[68] D. G. York, J. Adelman, J. E. Anderson Jr. et al., "The Sloan Digital Sky Survey: Technical Summary," *Astronomical Journal*, vol. 120, p. 1579, 2000.

[69] M.-P. Veron-Cetty and P. Veron, "A catalogue of quasars and active nuclei: 13th edition," *Astronomy and Astrophysics*, vol. 518, article 10, 2010.

[70] I. N. Evans and A. P. Koratkar, "A complete atlas of recalibrated Hubble Space Telescope Faint Object Spectrograph spectra of active galactic nuclei and quasars—I. Pre-costar spectra," *Astrophysical Journal*, vol. 150, no. 1, pp. 73–164, 2004.

[71] F. Mannucci, F. Basile, B. M. Poggianti et al., "Near-infrared template spectra of normal galaxies: k-corrections, galaxy models and stellar populations," *Monthly Notices of the Royal Astronomical Society*, vol. 326, no. 2, pp. 745–758, 2001.

[72] M. Vestergaard and B. J. Wilkes, "An empirical ultraviolet template for iron emission in quasars as derived from I Zwicky 1," *Astrophysical Journal*, vol. 134, no. 1, pp. 1–33, 2001.

[73] M. Uslenghi and R. Falomo, "AIDA: astronomical image decomposition and analysis," *Modelling and Simulation In Science*, pp. 313–318, 2008.

[74] J. K. Kotilainen, R. Falomo, M. Labita, A. Treves, and M. Uslenghi, "The nuclear to host galaxy relation of high-redshift quasars," *Astrophysical Journal*, vol. 660, no. 2 I, pp. 1039–1050, 2007.

[75] L. Simard, J. Trevor Mendel, D. R. Patton, S. L. Ellison, and A. W. McConnachie, "A catalog of bulge+disk decompositions and updated photometry for 1.12 million galaxies in the Sloan digital sky survey," *Astrophysical Journal*, vol. 196, no. 1, article 11, 2011.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 893984, 2 pages
doi:10.1155/2012/893984

Erratum

Erratum to "Are Nuclear Star Clusters the Precursors of Massive Black Holes?"

Nadine Neumayer[1,2] and C. Jakob Walcher[3]

[1] *European Southern Observatory, Karl-Schwarzschild Straße 2, 85748 Garching bei München, Germany*
[2] *Excellence Cluster Universe, Boltzmann Straße 2, 85748 Garching bei München, Germany*
[3] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany*

Correspondence should be addressed to Nadine Neumayer, nneumaye@eso.org

Received 2 May 2012; Accepted 16 May 2012

The original paper had an erroneous value for the Srsic index of NGC 205, which was quoted as $n = 4$. We used the surface brightness profile of [1] to derive Srsic $n = 2.05$. The correct version of Figure 2 is reproduced below. This does not influence any of the conclusions.

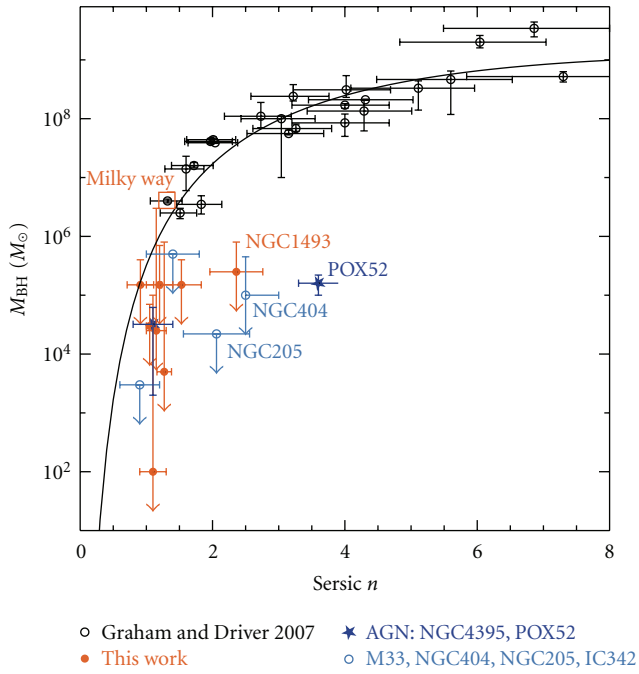


FIGURE 2: The mass of the BH against the Srsic index of the host bulge or disk. We plot the objects as listed in the text. The largest outliers are NGC205 (with $n = 2.05$), POX52 ($n = 3.6$), NGC404 ($n = 2.5$), and NGC1493 ($n = 2.4$).

References

[1] M. Valluri, L. Ferrarese, D. Merritt, and C. L. Joseph, "The low end of the supermassive black hole mass function: constraining

the mass of a nuclear black hole in NGC 205 via stellar kinematics," *The Astrophysical Journal*, vol. 628, no. 1, pp. 137–152, 2005.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 709038, 13 pages
doi:10.1155/2012/709038

Research Article

Are Nuclear Star Clusters the Precursors of Massive Black Holes?

Nadine Neumayer[1,2] and C. Jakob Walcher[3]

[1] *European Southern Observatory, Karl-Schwarzschild Straße 2, 85748 Garching bei München, Germany*
[2] *Excellence Cluster Universe, Boltzmann Straße 2, 85748 Garching bei München, Germany*
[3] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany*

Correspondence should be addressed to Nadine Neumayer, nneumaye@eso.org

Received 15 August 2011; Revised 17 November 2011; Accepted 27 November 2011

Academic Editor: Francesca Civano

We present new upper limits for black hole masses in extremely late type spiral galaxies. We confirm that this class of galaxies has black holes with masses less than $10^6 M_\odot$, if any. We also derive new upper limits for nuclear star cluster masses in massive galaxies with previously determined black hole masses. We use the newly derived upper limits and a literature compilation to study the low mass end of the global-to-nucleus relations. We find the following. (1) The $M_{\rm BH}$-σ relation cannot flatten at low masses, but may steepen. (2) The $M_{\rm BH}$-$M_{\rm bulge}$ relation may well flatten in contrast. (3) The $M_{\rm BH}$-Sersic n relation is able to account for the large scatter in black hole masses in low-mass disk galaxies. Outliers in the $M_{\rm BH}$-Sersic n relation seem to be dwarf elliptical galaxies. When plotting $M_{\rm BH}$ versus $M_{\rm NC}$ we find three different regimes: (a) nuclear cluster dominated nuclei, (b) a transition region, and (c) black hole-dominated nuclei. This is consistent with the picture, in which black holes form inside nuclear clusters with a very low-mass fraction. They subsequently grow much faster than the nuclear cluster, destroying it when the ratio $M_{\rm BH}/M_{\rm NC}$ grows above 100. Nuclear star clusters may thus be the precursors of massive black holes in galaxy nuclei.

1. Introduction

Supermassive black holes (BHs) are thought to be ubiquitous in the nuclei of massive galaxies. The discovery of a number of tight correlations between the global properties of galaxies and the properties of their nuclei (e.g., [1–3]) has led astronomers to realize that the evolution of galaxies may be closely linked to their nuclear properties. However, the nuclei of galaxies do not only host massive BHs but also host massive star clusters, commonly called nuclear star clusters (NCs). (Note that we here make the distinction between nucleus, that is, the location at the very center, and nuclear star cluster. Often the NC has been called nucleus or stellar nucleus in the past, but this seems ambiguous to us).

The overall nucleation frequency is around 75% over all Hubble types ([4–6], hereafter B02), but NCs seem to be absent in the most massive galaxies [5, 7]. NCs typically have stellar velocity dispersions of 15–35 $\rm km s^{-1}$, effective radii of a few parsecs, and dynamical masses of $\sim 10^6$–$10^7\, M_\odot$ (B02 [8, 9]). Moreover, they show stellar populations of multiple ages [10–12], pointing towards them having a complex formation history. This might be related to their special location in the galaxy, as on average, NCs appear to sit at the photometric centre of their host galaxy [6, 13]. We recently showed that for bulgeless galaxies their location also coincides with the kinematic centre, that is, the bottom of the potential well [14].

Intriguingly, NCs in late-type spirals and dwarf ellipticals follow relationships with their host galaxies that mirror the $M_{\rm BH}$-σ and $M_{\rm BH}$-$M_{\rm bulge}$ relationships of high-mass galaxies [10, 15, 16], suggesting the possibility that the fueling and growth of NCs and BHs are determined by similar processes, and that BHs and NCs should be grouped together into "central massive objects" (CMOs, [15]). The NC would then be nothing else than the failed BH [17]. In this picture, BHs would form in high-density clumps typically located in high-mass galaxies, while NCs form from lower-density clumps in lower-density disks. Recent simulation studies (e.g., [18, 19]) have been able to reproduce the formation of BHs through direct collapse models. If the collapse is quick—compared to the cooling time of the gas—a BH will form. If, however, the gas has sufficient time to cool and form stars, it will

form an NC (see also the recent review in [20]). Competing formation scenarios for NCs are, however, equally successful. For example, recent work by Hartmann et al. [21] has shown that the observed properties of NCs are well reproduced by combining mergers of star clusters with the accretion of gas at a much later time in the history of a galaxy.

A further reason for interest in NCs and their BHs is that a number of authors [22–26] have found that dense clusters of young, massive stars can experience runaway coalescence of their most massive stars, leading to an intermediate mass black hole (IMBH, but see also [27]). It would then be tempting to identify NCs with the long-sought seeds for BH formation. An observational result supporting this view is that NCs and BHs can coincide [28, 29]; this is especially well studied in our own Galaxy [30, 31]. On the other hand, parameter studies of the runaway collapse scenario (e.g., [32]) show that NCs are actually not in a region of parameter space that would be favourable to the collapse.

Of the many global-to-nucleus relations, the three most frequently referred to ones seem to be the $M_{\rm BH}$-σ relation [1, 2], the $M_{\rm BH}$-$M_{\rm bulge}$ relation [3], and the $M_{\rm BH}$-$n_{\rm Sersic}$ relation [33]. As all these relations have been initially set up for the range of massive galaxies (i.e., $M_{\rm BH} > 10^8\,M_\odot$), the low-mass range of BHs is not very well populated and holds most potential to find out which one of the three is more fundamental. A particularly interesting case is BHs and NCs in bulgeless galaxies. Indeed, while according to the $M_{\rm BH}$-σ relation one would expect late-type, bulgeless spirals to host BHs of mass $\leq 10^6\,M_\odot$, the $M_{\rm BH}$-$M_{\rm bulge}$ relation is no longer "defined" for bulgeless galaxies, as the lack of a bulge would imply the absence of a black hole. On the other hand, exploring the low-mass end of the scaling relations, Green et al. [34] have derived reliable BH masses in spiral galaxies (with bulges) from maser measurements and find that these fall below the $M_{\rm BH}$-σ relation of elliptical galaxies but seem consistent with the $M_{\rm BH}$-$M_{\rm bulge}$ relation.

In fact both NCs and BHs have been found in bulgeless galaxies. For NCs see B02; for BHs see, for example, the cases of NGC4395 [28, 35], NGC1042 [36], NGC3621 [37, 38], and probably many more (see, e.g., [39–42]). On the other hand, very tight upper limits for the BH mass exist for some galaxies such as M33 [43, 44], but direct observational constraints are scarce because such small BHs are extremely difficult targets for dynamical searches and therefore very few objects have useful measurements. While it would thus seem tempting to declare that NCs are the central spheroids in bulgeless galaxies, this could lead to a paradox. Indeed, NCs have largely been identified with CMOs in massive galaxies, on the ground that they follow similar scaling relations as BHs. Identifying the same objects with the spheroid in low-mass galaxies would imply a transition in physical properties of the NC. Many observational hints seem to point against this possibility [9], the most important being that NCs have constant radius over Hubble type. A backdoor might be that Erwin and Gadotti [45] find that BH mass correlates with bulge mass (and no correlation with disk mass exists, [46]), while NC mass correlates better with total galaxy mass.

To conclude this introduction, measurements of the demographics of the lowest-mass BHs are an important goal. Their mass distribution encodes a fossil record of the mass scale and formation efficiency of the initial BH seeds at high redshift (e.g., [47]) and they hold the power to help us distinguish between different scenarios explaining the observed global-to-nucleus relations [48–52]. In order to increase the statistics in this particularly interesting low-mass regime, we have calculated $M_{\rm BH}$ upper limits for a sample of 9 NCs, for which integrated velocity dispersions had been published previously ([9], hereafter W05). We have also placed upper limits on $M_{\rm NC}$ for a sample of 11 galaxies with measured black hole masses. We have then used these upper limits in conjunction with a literature compilation to gauge which of the different proposed global-to-nucleus relations seem to hold best at the low-mass end.

2. New Upper Limits for BHs in NCs

2.1. Data. Our sample consists of 9 NCs culled from the HST/WFPC2 snapshot survey of B02. Imaging in the F804W filter is available from B02 and we refer to this paper for all details. All 9 NCs are resolved, even if some only barely. We here use the images as available through MAST to yield the surface brightness profile through a multi-Gaussian expansion (see the following).

VLT/UVES spectra with high S/N and high spectral resolution have been obtained by W05. We use their velocity dispersion measurement. The properties of our sample are summarized in Table 1.

The sample selection for spectroscopic follow-up technically implied a slight bias to the more luminous among the NCs. Nevertheless, we expect this sample to be a fair representation of NCs in pure disk galaxies in general, as it covers the upper 2/3 of the luminosity range of NCs.

2.2. Analysis. We constructed a dynamical model to estimate the mass and *M/L* of the NCs and to put meaningful upper limits on the possible central black holes inside them. The first step in this process is developing a model for the light distribution. To parametrise the surface brightness profiles of the NCs and to deproject the surface brightness into three dimensions, we adopted a Multi-Gaussian Expansion (MGE; [53]). The MGE fit was performed with the method and software of Cappellari [54], on the HST *I*-band images deconvolved from the PSF (using a Tiny Tim PSF [55]). As most of the clusters are barely resolved in the HST images and shape measurements are therefore impossible, we assume spherical symmetry. Note that although the NCs in NGC300 and NGC7793 (the best resolved) are indeed round, this may be due to their host disks being seen face-on. Seth et al. [11] find that edge-on NCs can have quite disky outer isophotes.

We use the Jeans Anisotropic MGE (JAM) software by Cappellari [56] which implements the solution of the Jeans equations allowing for orbital anisotropy. The model has three free parameters: (i) the anisotropy, (ii) the mass of a central black hole $M_{\rm BH}$, and (iii) the *I*-band total dynamical *M/L*. From the velocity dispersion profile computed by JAM, we compute the luminosity-weighted velocity dispersion ($\sigma_{\rm LW}$) over an aperture of 1 square arcsecond. This corresponds to the width of the UVES slit on the sky. We

Table 1: Properties of the sample of NCs in bulgeless galaxies.

Galaxy	Type	NC r_e (pc)	σ (km/s)	$M/L^{\min}$ $(M/L_{I,\odot})$	$M_{\rm BH}^{\max}$ $(M_\odot)$	$M_{\rm BH}^{\rm best}$ $(M_\odot)$
NGC 300	SAd	2.9	13 ± 2	0.41	1×10^5	1×10^2
NGC 428	SABm	3.36	24.4 ± 4	0.41	7×10^4	3×10^4
NGC 1042	SABcd	1.94	32 ± 5	0.07	3×10^6	2.5×10^4
NGC 1493	SBcd	2.6	25 ± 4	0.07	8×10^5	2.5×10^5
NGC 2139	SABcd	10.3	17 ± 3	0.02	4×10^5	1.5×10^5
NGC 3423	SAcd	4.18	30 ± 5	0.87	7×10^5	1.5×10^5
NGC 7418	SABcd	12.3	34 ± 5	0.10	9×10^6	1.5×10^5
NGC 7424	SABcd	7.4	16 ± 2	0.10	4×10^5	1.5×10^5
NGC 7793	SAd	7.7	25 ± 4	0.15	8×10^5	5×10^3

iterate the computation of $\sigma_{\rm LW}$ over a grid of values for M/L and $M_{\rm BH}$. The results are shown in Figure 1 which is directly comparable to Figure 8 of Barth et al. [37]. Direct comparison with the mass-to-light ratios obtained by W05 (thin solid vertical line) shows that the ratios scatter around 1.0, with no obvious systematic outliers. The small differences in the result can be attributed to the way in which the surface brightness was modelled (Multi-Gaussian expansion here versus direct deprojection in W05).

The maximum allowed mass of the black hole will be obtained when a minimum of mass is present in the form of stars. From Figure 1 one can easily read what BH mass would result if we assumed $M/L = 0$ for the stars in the cluster. A more interesting lower limit to the M/L comes from the spectral fitting with stellar population models. We exploit the fact that the age obtained by fitting a simple single stellar population to a composite stellar population is strongly biased to the age of the youngest population in that object which contributes significantly to the total luminosity (see, e.g., W06, [57]). The relevant values are tabulated in W06 and repeated in Table 1 along with the values derived as upper limits to the mass of a putative BH from the intersection between both thick solid lines in Figure 1. This is a conservative estimate for the $M_{\rm BH}$ upper limit. A more realistic value for $M_{\rm BH}$ can be derived from the intersection or asymptotic point of the model (thick solid curved line) with the best-fit M/L from W05 (thin solid vertical line) in Figure 1. The resulting best-fit $M_{\rm BH}$ values are listed in Table 1.

We explicitly test the effect of velocity anisotropy on the modeling results and found very little change in the results—certainly below our systematic uncertainties due to the lower limit to the mass-to-light ratio that we apply (see also [37]). We therefore neglect this effect for the rest of this paper.

3. Global-to-Nucleus Relations

We now plot the upper limits we have derived into figures showing existing correlations from earlier work. In these figures we typically have a comparison sample which is taken from a larger statistical study and we add a number of objects at the low-mass end from different sources in the literature. We have tried to be complete at the very lowest mass end of the relations. Further literature does exist, but typically, the BH masses exceed values of $\sim 10^6\, M_\odot$ and the galaxies structural parameters have not been studied individually.

3.1. M_{BH}-σ Relation. For Figure 2, the $M_{\rm BH}$-σ relation, the comparison sample, and relation are as compiled by Gültekin et al. ([58], black open symbols). We extend this compilation with recent maser measurements by Greene et al. ([34], green crosses). Active AGNs are denoted by blue stars; these are NGC4395 [28, 59] and POX52 [60]. In principle, NGC1042 from the present work falls also into this category (see, [36]) but is plotted as a filled red circle. Previous upper limits for nonactive nucleated galaxies are plotted as open blue circles: M33 [43, 44], NGC205 [61], IC342 [62], and NGC404 [63]. We also plot the globular clusters G1 [64], ω Cen [51, 65, 66], and NGC6388 [67] as green open circles. The verdict on the usefulness of these measurements is still out, with strong contrasting claims by other authors that there is no evidence for a black hole in ω Cen [68, 69] and in G1 (e.g., [70]). We nevertheless use the derived values in a spirit of adventure, that is, what would it mean if these measurements were correct? Finally, the new upper limits derived in this work are denoted by filled red circles. It emerges that a flattening of the relation is not consistent with the current measurements. It may well be that a downwards bending would be necessary, if more stringent upper limits such as that for M33 would be published.

3.2. M_{BH}-M_{bulge} Relation. For Figure 3, the $M_{\rm BH}$-$M_{\rm bulge}$ relation, the comparison relation and sample are taken from Häring and Rix [3] (filled black circles), while the other data points come from the same sources as in Figure 2. There is a hint towards a flattening of the relation with the lowest spheroid masses, but it will be difficult to confirm this without much better estimates of the masses of IMBHs. On the other hand a steepening, that is, bulges that are too massive for their BHs, has been mentioned by Greene et al. [34, 71]. If there are BHs in galaxies with no bulges as well as bulges that are too massive for their BHs, it seems clear that the $M_{\rm BH}$-$M_{\rm bulge}$ relation must suffer from large scatter at small masses.

3.3. M_{BH}-n_{Sersic} Relation. For Figure 4, the $M_{\rm BH}$-$n_{\rm Sersic}$ relation, the comparison relation, and sample are taken from

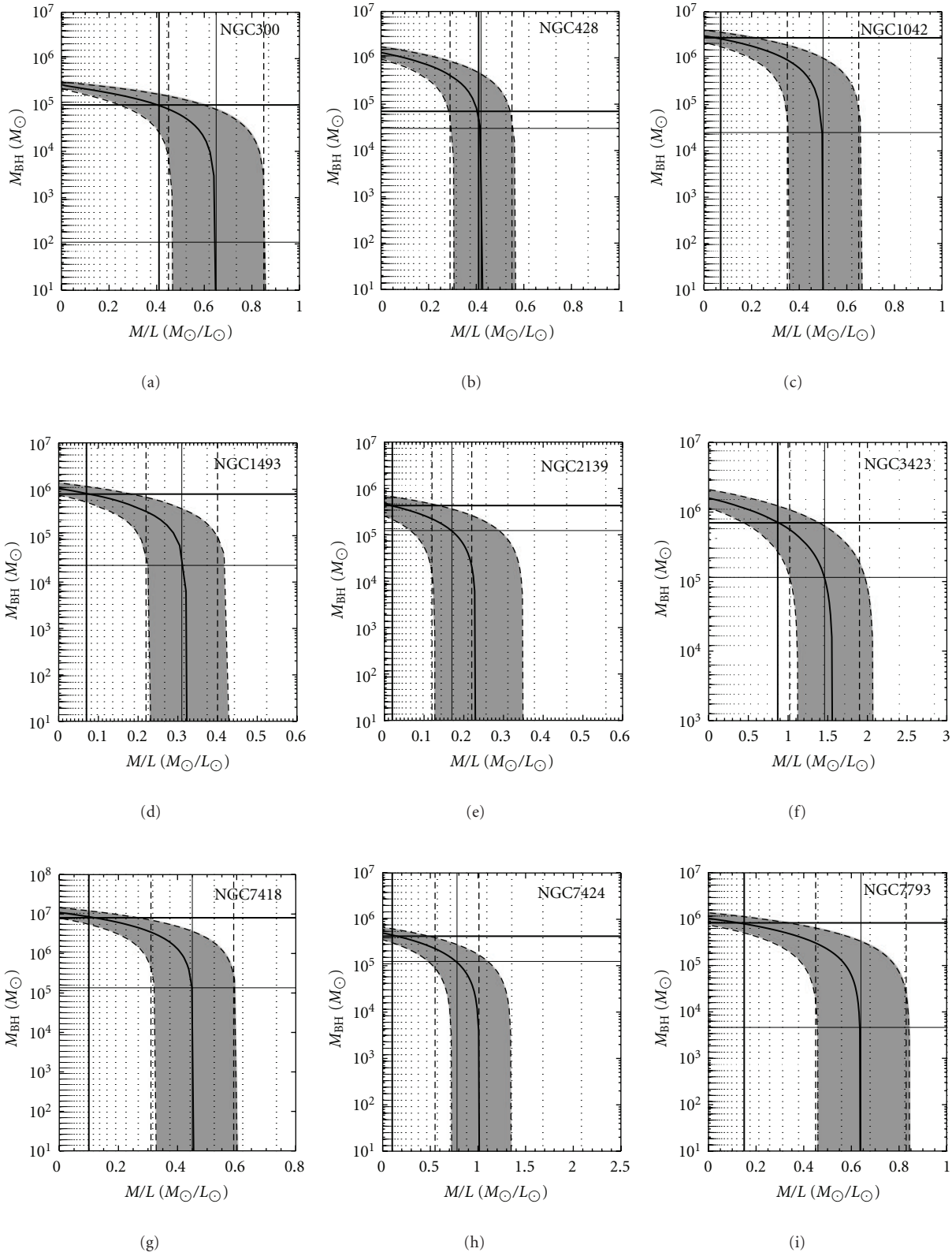


Figure 1: The M/L ratio assumed for the stellar population against the mass of the putative black hole for each of the nine nuclear clusters. Models falling onto the right solid vertical line have the same velocity dispersion as given in Table 1 for each cluster (dashed lines are upper and lower uncertainties). We also draw a vertical, full line (left) for the minimum mass-to-light ratio compatible with the observed spectrum of the stellar population in the cluster. The horizontal solid lines indicate the black hole masses referring to the two different M/L values quoted previously. For the minimum M/L we find a firm upper limit to the black hole mass M_{BH}^{max} (upper line) and for the best-fit stellar populations M/L we get M_{BH}^{best} (lower line).

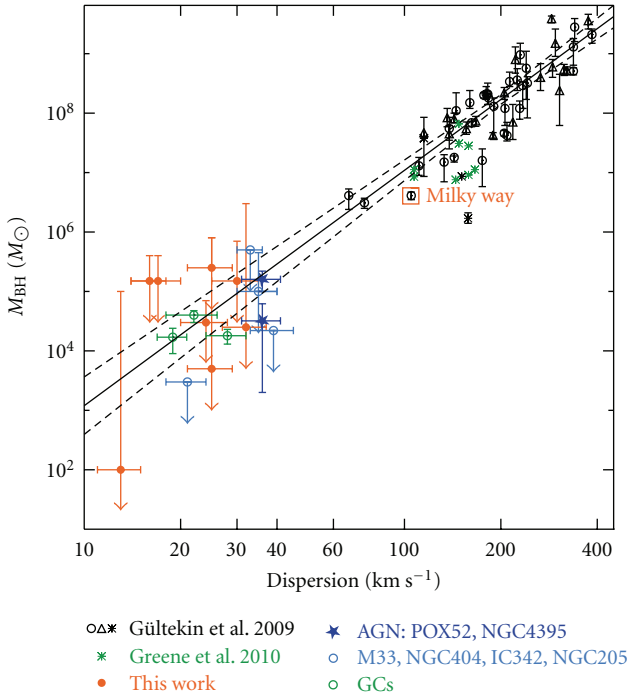


Figure 2: The relation between the mass of the BH and the velocity dispersion of the spheroid around it. We plot the objects as listed in the text. The lines give the best fit of Gültekin et al. [58].

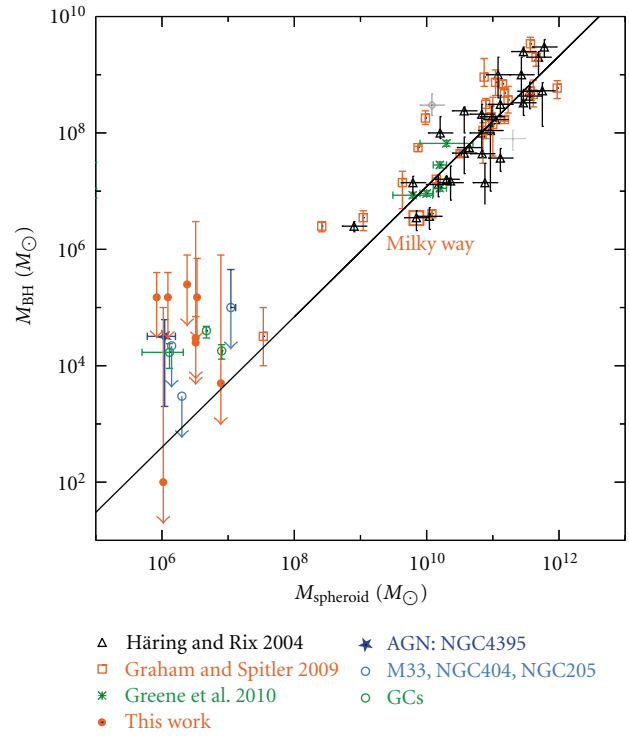


Figure 3: The BH mass versus the spheroid mass (bulge, GC, NC). We plot the objects as listed in the text. The line indicates the best-fitting relation of Häring and Rix [3].

Graham and Driver [33] (filled black circles), while the other data points come from the same sources as in Figure 2. We have also assembled measurements of the Sersic n from literature sources for all objects with published BH masses. For the galaxies with M_{BH} limits newly derived in the present paper, Sersic n was derived from the following literature sources: Ganda et al. [72] for NGC1042 and NGC3423 and Weinzirl et al. [73] for NGC2139. For NGC300 and NGC428 Spitzer IRAC 3.6 μm images were downloaded from the Spitzer Heritage Archive (http://sha.ipac.caltech.edu/) and Sersic n was determined using the GALFIT software [74]. For NGC1493, NGC7424, and NGC7793 the corresponding images were obtained from the 2MASS archive (http://irsa.ipac.caltech.edu/) and again fitted with GALFIT. All galaxies were fit using one PSF component for the central NC, one Sersic component representing the disk, and one constant sky component. All parameters were left free to be fit for. The webpages provide appropriate point spread functions; although all of our targets are nearby and therefore well resolved, the resulting Sersic n is almost independent of the PSF used in GALFIT. We caution that the resulting Sersic n may depend heavily on the radial range used in the fitting. To cite two extreme examples, the Sersic n of NGC300 is independent of the radial range used within $\Delta(n) = 0.1$. On the other extreme, the Sersic n for NGC1493 varies between $\sim$1.3 and the reported value of $\sim$2.5. It is beyond the scope of the current paper to derive a physically meaningful fit range that would put the physical meaning of the Sersic n on firmer ground. We emphasize that it is *despite* the cited uncertainties that the relation between M_{BH} and Sersic n seems to hold.

Figure 4 shows two interesting features. (1) Because the relation fitted by Graham and Driver [33] curves down at $n = 1$, a large range of BH masses is allowed in this regime, which clearly allows for the scatter that seems to emerge as a common trend in the previous two nucleus-to-global relations. (2) There are significant outliers in this plot, in the sense that some low-mass galaxies can have too high Sersic n for their BH mass.

3.4. M_{BH}-M_{NC} Relation. In Figure 5 we show the relation between M_{BH} and M_{NC} (compare [29, 75]). We have plotted objects already used above, for which determinations of both M_{BH} and M_{NC} exist. In searching for a high-mass comparison sample we have made use of the compilations by Graham and Spitler [75] and Gültekin et al. [76] from which we also take the distances. Where not available, we have then proceeded to derive upper limits to the NC masses either from the literature or from own fits to archival HST images. (Thorough work deriving consistent photometry and structural parameters for NCs across the entire Hubble sequence for large swaths of the HST archive is badly needed, but is beyond the scope of the current work. Note that one focus of such work could be the distinction (if any clear distinction exists) between NCs and nuclear disks. In the case of NGC4342, for example, the upper limit we give on the NC mass is not only observationally uncertain, but also conceptually uncertain. As Scorza and van den Bosch [77] discuss, a relation between the nuclear disk mass and the BH is as plausible as between the NC and the BH. Indeed, some

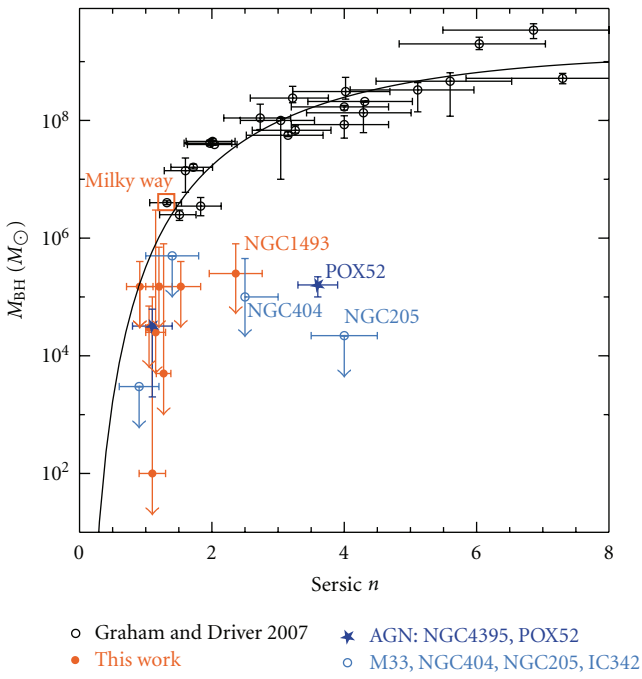


FIGURE 4: The mass of the BH against the Sersic index of the host bulge or disk. We plot the objects as listed in the text. The largest outliers are NGC205 (with $n = 4$), POX52 ($n = 3.6$), NGC404 ($n = 2.5$), and NGC1493 ($n = 2.4$).

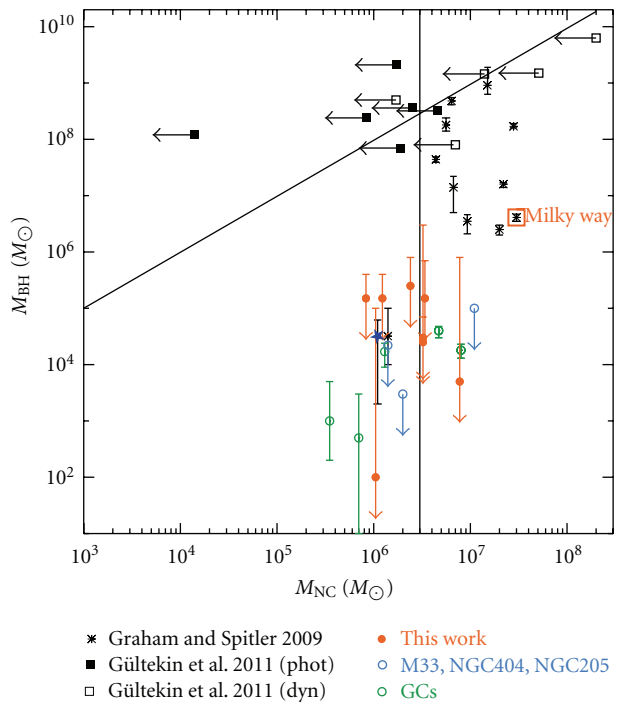


FIGURE 5: The mass of the BH mass versus the NC mass. We plot the objects as listed in the text. The two full lines indicate an NC mass of $3 \times 10^6\, M_\odot$ and a M_{BH}/M_{NC} mass ratio of 100. These lines separate NC-dominated galaxy nuclei (lower left of both lines) from BH-dominated galaxy nuclei (upper left of both lines) and a transition region (to the right of both lines).

NCs may turn out to be nuclear discs on close inspection (compare [11])).

We now discuss the ways that we have obtained upper limits for the NC masses galaxy by galaxy. We strongly emphasize that we have tried to obtain *upper limits to* rather than real measurements of the NC mass. Real measurements of NC masses can only be carried out by a combination of dynamical modeling and spectral analysis to determine the relative influence of the AGN and possible varying M/L ratios. We rather aim to be conservative with respect to all uncertainties affecting our estimates of upper limits to the NC masses. Our resulting upper limits are listed in Table 2. For the following 5 objects we estimated upper mass limits from the literature only.

NGC4486 (M87): the bright nucleus is dominated by AGN light. There is no evidence for an NC. We therefore use Figure 7 of Gebhardt and Thomas [78], which shows the enclosed stellar mass within the central arcsec to be $2 \times 10^8\, M_\odot$. This is consistent with an estimate from Young et al. [79], which gives a total of $M = 5 \times 10^9\, M_\odot$ within a radius of 100 pc and $M/L = 60$, thus leading to an estimate of the stellar mass within that radius of $3 \times 10^8\, M_\odot$, assuming that the stellar $M/L = 4$. We emphasize that this is the total stellar mass within a radius comparable to the radii of typical NCs and therefore gives an upper limit to M_{NC}. We do not claim that M87 actually hosts a stable NC at its centre.

NGC4374 (M84): an AGN has been shown to exist by Bower et al. [80], with very weak stellar features. To estimate an upper mass to the NC in NGC4374 we use the paper by Walsh et al. [81], which yields a BH mass estimate of $4 \times 10^8\, M_\odot$. Their Figure 4 shows the circular velocity profiles due to the BH and the stellar mass, respectively. Assuming a distance to M84 of 17 Mpc yields 70 pc/arcsec. Assuming an NC radius of 10 and a stellar $M/L = 4$, we obtain that at 10 pc radius the circular velocity due to the BH is 400 km/s, while the circular velocity due to the remaining stellar mass is smaller than 50 km/s. To obtain an estimate of the upper limit for a putative NC, we need to correct for the different masses and for the different spatial distribution (point-like versus extended). From the virial theorem, we can scale the velocity quadratically. From Table 2 in the current work it can be seen that a conservative factor (i.e., one that gives a lot of stellar mass) for the conversion from point-like to extended would be a factor of 10. The upper limit for the stellar mass within 10 then becomes $M_{NC}/M_{BH} = 10 * 50^2/400^2 = 0.15$, thus yielding an NC upper mass limit of $6.3 \times 10^7 M_\odot$. Walsh et al. [81] also state that stellar mass is a negligible contributor to their mass budget; it is therefore entirely possible that no NC exists in that galaxy.

NGC4261: the central luminosity distribution is complex, with a nuclear disk and a luminous nuclear source which seems to be dominated by an AGN; at least a radio jet is present [82]. There is thus no clear evidence in favour of any NC. Ferrarese et al. [82] find that $M/L_V = 2100$ within the inner 14.5 pc. A maximum M/L_V for stellar populations is 7. We thus obtain that 7/2100 of the central mass within 15 pc can be in stars, which is $5 \times 10^8/300 = 1.6 \times 10^6\, M_\odot$.

Table 2: Sample of galaxies for which new properties were derived in this paper.

Galaxy	Type	σ(km/s)	Dist (M_{pc})	M_{BH} ($M_\odot$)	M_{NC} ($M_\odot$)	Sersic n	M_{Bulge} ($M_\odot$)	Ref
group 1				×				
NGC 300	SAd	13 ± 2	2.2	$<1\times10^2$	1×10^6	1.1	—	[105]
NGC 428	SABm	24.4 ± 4	16.1	$<3\times10^4$	3×10^6	1.05	—	[105]
NGC 1042	SABcd	32 ± 5	18.2	$<2.5\times10^4$	3×10^6	1.15	—	[105]
NGC 1493	SBcd	25 ± 4	11.4	$<2.5\times10^5$	2×10^6	2.36	—	[105]
NGC 2139	SABcd	17 ± 3	23.6	$<1.5\times10^5$	8×10^5	1.53	—	[105]
NGC 3423	SAcd	30 ± 5	14.6	$<1.5\times10^5$	3×10^6	1.20	—	[105]
NGC 7418	SABcd	34 ± 5	18.4	$<1.5\times10^5$	6×10^7	—	—	[105]
NGC 7424	SABcd	16 ± 2	10.9	$<1.5\times10^5$	1×10^6	0.91	—	[105]
NGC 7793	SAd	25 ± 4	3.3	$<5\times10^3$	8×10^6	1.27	—	[105]
group 2					×			
NGC 4486	E1	375 ± 18	17.0	6.3×10^9	$<2\times10^8$	6.86	6.0×10^{11}	[68]
NGC 4374	E1	296 ± 14	17.0	1.5×10^9	$<6.3\times10^7$	5.60	3.6×10^{11}	[41]
NGC 1332	S0	321 ± 14	19.6	1.45×10^9	$<1.4\times10^7$	—	—	[60]
NGC 3031	Sb	143 ± 7	4.1	8×10^7	$<7\times10^6$	3.26	—	[37]
NGC 4261	E2	315 ± 15	33.4	5×10^8	$<1.7\times10^6$	7.30	3.6×10^{11}	[70]
group 3					×			
NGC 4649	E2	385 ± 19	16.5	2.1×10^9	$<2\times10^6$	6.04	4.9×10^{11}	[103]
NGC 3998	S0	305 ± 15	14.9	2.4×10^8	$<8.5\times10^5$	—	—	[98]
NGC 2787	SB0	189 ± 9	7.9	0.7×10^8	$<1.9\times10^6$	1.97	—	[95]
NGC 3379	E0	206 ± 10	11.7	1.2×10^8	$<1.4\times10^4$	4.29	6.8×10^{10}	[108]
NGC 4342	S0	225 ± 11	18.0	3.6×10^8	$<2.5\times10^6$	5.11	1.2×10^{10}	[13]
NGC 4291	E2	242 ± 12	25.0	3.2×10^8	$<5\times10^6$	4.02	1.3×10^{11}	[103]

Galaxies for group 1 are from W05, and we here derived upper limits on the black hole mass and Sersic n. Galaxies for groups 2 and 3 are from Gultekin et al. [76]. For group 2 objects we derived upper limits on the NC mass via dynamical arguments, while for group 3 objects we used photometry to derive M_{NC} upper limits. For groups 2 and 3 Sersic n values are taken from Graham and Driver [33], bulge masses are from Häring and Rix [3], and velocity dispersions are from Hyperleda. The newly derived quantities are marked with an × at the top of the respective column. References for black hole masses are. (1) this work, (2) Gebhardt and Thomas [78], Gebhardt et al. [109], (3) Bower et al. [110], (4) Rusli et al. [83], (5) Devereux et al. [84], (6) Ferrarese et al. [82], (7) Gebhardt et al. [111], (8) De Francesco et al. [88], (9) Sarzi et al. [112], (10) Gebhardt et al. [90]; Shapiro et al. [113], and (11) Cretton and van den Bosch [89].

NGC1332: there is no firm evidence for an NC, although the surface brightness profile of Rusli et al. [83] hints at a central luminosity excess within the central arcsec. The dynamical model of Rusli et al. [83] gives a central stellar luminosity density of $4\times10^{12}\,L_\odot\,\mathrm{kpc}^{-3}$. For an NC of 5 pc radius this yields an NC luminosity of $2\times10^6\,L_\odot$ in the R-band. With $M/L_R = 7$ (also according to [83]), M_{NC} max is $1.4\times10^7\,M_\odot$.

NGC3031: Devereux et al. [84] list values of stellar mass within radius in their Table 3. From their Figure 3, it is clear that the nuclear source is not extended; there is thus no evidence for the presence of an NC. To estimate an upper mass limit for the NC, we assume an NC radius of 7 pc (compare [5, 6]); the upper limit to M_{NC} is then $7\times10^6\,M_\odot$.

For the following 6 objects no NC mass estimate was available. We therefore turned to the HST images as downloaded from the Hubble Legacy Archive. We have then used GALFIT [74] to derive the magnitudes of the NCs. Because all NCs we treat in this last step are in early-type galaxies, we can assume that their ages range between 1 and 10 Gyr, yielding an estimate of the allowed range for the M/L ratio. For most cases we used the F814W filter on either ACS or WFPC2, setting the allowed range of M/L between 1 and 4. Much more sophisticated modeling of the photometry, while possible, would yield only marginally better estimates of the total stellar mass of the NC for several reasons. (1) The star formation histories (SFHs) of NCs are unknown and, indeed, likely to be semirandom, repetitive bursts of star formation. Therefore no strong prior can be applied to the SFH. Because the oldest stellar populations are the faintest per unit mass, the resulting uncertainty on M/L is of order factor 2. (2) The photometry of the NCs is in itself uncertain. We have made use of realistic PSFs from either Jee et al. [85] (http://acs.pha.jhu.edu/~mkjee/acs_psf/) or from Tiny Tim (http://www.stsci.edu/hst/observatory/focus/TinyTim). It is much less certain what the ideal profile for the surface brightness of the host galaxy should be though (compare [86, 87]). We have used one single Sersic, as we are only interested in subtracting the host, not in describing it. Nevertheless, we estimate that the use of different profiles (2 Sersics, Nuker, etc.) could impact the total photometry of the NC by up to 0.5 or even 1 magnitude (compare, e.g., the central extrapolations of [6]). We therefore have chosen to let these uncertainties be reflected in the errorbars of the NC mass estimate, rather than trying to hide them somewhere within a sophisticated analysis.

NGC4649: no nuclear source is visible (as also found by [75]). We first fit this galaxy with a single Sersic. When additionally forcing in a point source (GALFIT PSF component) of different magnitudes (20, 20.5, 21,21.5, 22), the resulting oversubtraction can be seen clearly in the residual image for as faint as $m_I = 21.0$. We use this value as a conservative upper limit to the NC magnitude. This results in an upper mass limit of $2 \times 10^6\ M_\odot$.

NGC4291: we attempted the same procedure as before. However, due to a flat central surface brightness profile, our simple Sersic fit by itself produced an oversubtraction of the central flux, not allowing us to use the exact same procedure as for NGC4649. Nevertheless, the HST image clearly shows the absence of any point source in the center. We therefore assumed the same limit as before, that is, 21 mag in F814W, which results in $M_{\rm NC} = 5 \times 10^6\ M_\odot$.

NGC3998: after the GALFIT fit, a clear spiral structure and a bar are seen in the residuals. The central light source was modeled as a Sersic with an effective radius of $0.2''$ and a Sersic $n = 0.1$, making us believe it is unresolved. De Francesco et al. [88] classify this galaxy as a LINER; thus the central source is AGN-light dominated. Therefore our photometrically derived NC mass of $8.5 \times 10^5\ M_\odot$ again is a conservative upper limit.

NGC4342: the fit with GALFIT was difficult, with 4 Sersic components in the final fit. The final solution was chosen to oversubtract the NC. Again we have a conservative upper limit of 21.85 mag corresponding to $M_{\rm NC} = 2.5 \times 10^6\ M_\odot$, using an M/L of 6.5 in I from Cretton and van den Bosch [89]. Contamination from AGN light is also possible, making our upper mass limit more robust.

NGC3379 (M105): the NC is visible in Gebhardt et al. [90], but not mentioned there. Graham and Spitler [75] note this galaxy as unnucleated. Two extended components with Sersic $n \sim 1$ and one very compact source with Sersic $n \approx 1/2$ (i.e., Gaussian surface brightness profile) and $r_e = 0.2''$ give a good fit to this object. The measured integrated magnitude of the central point source is 25.7 in the F814W of WFPC2, corresponding to $M_{\rm NC} = 1.4 \times 10^4\ M_\odot$. We used $M/L_F814W = 3$ as a suitable upper limit to the M/L.

NGC2787: this galaxy was analyzed in Peng et al. [74] and the nuclear photometry is taken from that source. We used $M/L_F547M = 3$ as a suitable upper limit to the M/L. Thus, the NC upper mass limit is $1.9 \times 10^6\ M_\odot$.

Note that in the galaxies NGC4486, NGC4374, and NGC3379 a luminous nuclear source is clearly seen. While this could all be AGN light, we see no way to ascertain the absence of an NC. In contrast to Graham and Spitler [75] we only claim to be able to derive an upper limit to the NC mass, rather than excluding an NC all together. Note also that a stellar cusp containing 10% of the BH mass is predicted around any BH [91].

4. Discussion

We now discuss and interpret a number of features we saw in the previous section, with the aim to discuss ideas that emerge from these figures but, to our knowledge, have not been discussed in the literature before. The ultimate aim of our study is of course to contribute to a consistent physical picture of black hole and nuclear cluster growth.

In Figure 4, one relation between the global galaxy properties and $M_{\rm BH}$ holds for a large range of n, independent of the presence of a bulge. Outliers are rather low-mass galaxies (and not low-mass black holes).

We stress that this relation is purely observational at this stage. Due to the heterogeneous assembly of the Sersic n values (literature, own fitting) the physical region represented by them is not always the same. In particular some of the galaxies do not contain a bulge, while for others the Sersic n has been explicitly measured for the bulge component. The existence of a relation seems evidence that indeed the measurement of Sersic n is meaningful. In particular, no conspiracy is obvious that would fundamentally bias our measurements in the sense of producing a spurious correlation. It thus seems to us that even independently of the exact details of the derivation of the Sersic n, it clearly describes a property of the galaxy that is relevant for the BH mass. A question to ask is then whether we fully understand the physical implications of that relation, and whether we could potentially reduce the relation to underlying intrinsic distribution of galaxy properties (e.g., if Sersic n was related to bulge mass in a very tight manner, we might be tempted to argue that bulge mass is the more fundamental measurement). We believe that the present paper cannot resolve this question but hope it provides motivation to explore these issues further.

To venture a possible physical interpretation of the outliers from the relation we note the following: it could be that the transformation process from disk galaxy to spheroid is different in this galaxy mass regime. While BHs in massive galaxies grow during the morphological transformation process of their host galaxies, BHs in low-mass galaxies are not affected (fed) during the transformation process. It might be worthwhile exploring through simulations, whether this has to do with a possible transformation dichotomy, that is, mergers versus harassment. It is worth pointing out here that such a dichotomy does not seem to be immediately apparent from the age or metallicity profiles, as these seem not to depend on mass [92].

Figure 5 has not been published previously in this form to the best of our knowledge (though compare [29, 75] for similar representations) and may yield considerable insight on the relation between NCs and BHs. An immediate conclusion from this figure is that BHs and their host NCs do not share the same intimate connection as BHs and their host spheroids. It rather seems that in galaxies with a high total mass, or alternatively a sizable spheroid, the BH has been able to grow *independently* of the NC, thus being able to reach comparable masses. In galaxies or star clusters unaffected by spheroid growth, as, for example, the GCs, M33, and others, it seems the BH, if existent, is only a very small portion of the mass of the NC.

Figure 5 (and similar figures, see [7, 29, 75] is still in an early phase and we believe that further studies in the field will attempt to fill in the high- and low-mass end of the BH mass regime with more NC masses and BH masses, respectively. Nevertheless it seems that two extreme ends can be identified,

with a transition region in between. At the low BH-mass end, there is little evidence for the presence of any BH; yet NCs seem to be common (B02). On the other hand no nuclear BH has been found that is not surrounded by an NC in this regime. If GCs with BHs are indeed the remnants of accreted satellite galaxies (e.g.,[93–96]) and if indeed they lie on the $M_{\rm BH}$-σ relation, this would imply, however, that at this stage BH growth is linked to NC growth much in the same way that BH growth is later tied to galaxy growth. A close look suggests indeed that some physical processes that occur in or with NCs, such as gas accretion [97–99] and merging [100–105], are quite similar to those experienced by galaxies. An alternative to the assumption that the process giving rise to the $M_{\rm BH}$-σ is so astoundingly generic is of course that the BH mass measurements in GCs may be subject to the "expectation bias", that is, when the measurement is in accordance with the expectations of the community they get accepted more easily.

At the very high-mass end of the BH-mass range, the BH is much more massive than the NC. On the other hand, this is the region where global-to-nucleus relations hold best. This could happen through two mechanisms: (1) either the galaxies in question never had a sizable NC, possibly because their central BHs grew early on in the age of the universe, thus stopping NC growth [106], or (2) massive BHs destroy their host NCs. Figure 5 in its current form suggests that this may happen at a mass ratio of $\geq$100 or alternatively when the BH radius of influence is of the same size as the NC radius. Loss cone depletion and core formation in early-type galaxies are well-studied mechanisms, that would amply suffice to destroy the preexisting NC [107].

Bekki and Graham [108] have examined an alternative solution for the disappearance of NCs for massive galaxies. Their argument relies on the mergers that are responsible for the morphological transformation from disk-dominated to bulge-dominated galaxies. They show that NCs can be significantly heated and thus be made susceptible to destruction during the merger event. The picture painted here differs significantly from that painted in Bekki and Graham [108] in that we put weight on the importance of the BH for destroying the NC. Indeed, what determines NC disappearance does not seem to be galaxy morphology, as most early-type galaxies have NCs. Rather, there is evidence for an upper limit to the BH/NC mass ratio, arguing strongly for a pivotal role of this ratio in leading to NC disruption.

The intermediate mass or transition regime may possibly lie between two boundaries, that is, above NC masses of $5 \times 10^6 M_\odot$ and below a $M_{\rm BH}/M_{\rm NC}$ mass ratio of 100. In this intermediate mass regime, while BHs have grown by at least 2 orders of magnitude, and probably more than 4 as compared to the NC-dominated regime, the NC grows by at most a factor of 10. There thus is some common growth; yet it does not occur in parallel. On the other hand, this is the region of most scatter in the typical global-to-nucleus relations. This phase would thus be characterized as a transition phase between NC-dominated nuclei and BH-dominated nuclei.

Does Figure 5 imply that NCs do not grow by the same processes as their BHs and is this a serious setback to the grouping together of NCs and BHs into CMOs [15]? That NCs and BHs need not grow in parallel has been emphasized by Nayakshin et al. [106], where both types of objects rather grow in competition for the same gas reservoir. Nayakshin et al. [106] ask whether the BH can prevent the NC from growing through its feedback and postulate that this is the case when the gas accretion rate is smaller than the Eddington rate. This picture is attractive in explaining Figure 5 because it naturally explains the three regimes—NC dominated, NC/BH transition, and BH dominated. Nevertheless, given the very low accretion rates observed in bulgeless galaxies and the presence of significant BHs in at least a few of them, this picture seems to break down exactly for the NC-dominated regime.

Discrimination between the different scenarios envisaged in the literature seems to be mostly an observational question at present. At low masses the error bars on BH measurements are typically very large, while NC masses are well measured. At high masses, BH masses are more accurate while the uncertainties for NC masses increase, due to resolution problems of the NCs above the underlying galaxies. We need both reliable BH and NC masses to see what the exact locus of points in this plot is. If there is a smooth transition, making the sequence look like a closed parenthesis, this would imply that the destruction of the NC due to the growing black hole is a slow process. If there really is a well-defined transition at $M_{\rm BH}/M_{\rm NC} = 100$, then this would imply either that the process of NC destruction is very fast or that these galaxies never had an NC.

5. Conclusions

We have computed new upper limits for the masses of intermediate mass black holes in 9 pure disk galaxies with very low BH masses. We also computed upper limits to the masses of nuclear star clusters in the nuclei of galaxies with previously determined massive BHs. We plot these upper limits on the three global-to-nucleus relations $M_{\rm BH}$-σ, $M_{\rm BH}$-$M_{\rm bulge}$, and $M_{\rm BH}$-Sersic n, as well as on a new figure that compares $M_{\rm BH}$ and $M_{\rm NC}$. We discuss the features we see in these figures. Two possible conclusions emerge from our discussion.

(1) In the $M_{\rm BH}$-Sersic n figure, those galaxies that lie on the relation seem to prove that there is a relation between $M_{\rm BH}$ and the morphological transformation of their host galaxies. A few notable outliers are dwarf elliptical galaxies, where the morphological transformation process does not seem to be associate with BH growth. We speculate that this difference may arise from different mechanisms, that is, mergers for high-mass galaxies and harassment for dwarfs.

(2) In the $M_{\rm BH}$-$M_{\rm NC}$ figure, we can clearly distinguish three regimes; NC dominated, BH dominated, and transition between the two. We speculate that this could imply that BHs are formed in NCs, then start to grow much faster than their host NCs, and, through a transition phase with similar masses for both components, could then ultimately destroy their host through loss cone depletion.

We expect further progress in the field to arise from better measurements of BH masses at the low-mass end of the $M_{\rm BH}$ mass function and from better measurements of NC masses at the high-mass end of the $M_{\rm BH}$ mass function. In particular, it might be useful for further research in the field if authors attempting to measure black hole masses also stated more clearly what their constraints on the NC mass are. Currently NCs are treated more or less as a nuisance to get rid of, while a clearer assessment of the constraint on their mass would benefit our understanding of the role NCs play in galaxy nuclei.

Acknowledgments

The authors thank the referee for an extensive report that significantly improved the presentation of the results in this paper, in particular on the NC upper limits. The authors acknowledge the support and hospitality of the ESO Garching office for science during the genesis of this work. N. Neumayer acknowledges support by the DFG cluster of excellence "Origin and Structure of the Universe". This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. They acknowledge the usage of the HyperLeda database (http://leda.univ-lyon1.fr). This work is based in part on observations made with the NASA/ESA Hubble Space Telescope and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA), and the Canadian Astronomy Data Centre (CADC/NRC/CSA). This work is based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This paper has made use of NASA's Astrophysics Data System Bibliographic Services.

References

[1] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal Letters*, vol. 539, no. 1, pp. L9–L12, 2000.

[2] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal Letters*, vol. 539, no. 1, pp. L13–L16, 2000.

[3] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *Astrophysical Journal Letters*, vol. 604, no. 2, pp. L89–L92, 2004.

[4] C. M. Carollo, M. Stiavelli, and J. Mack, "Spiral galaxies with WFPC2. II. The nuclear properties of 40 objects," *Astronomical Journal*, vol. 116, no. 1, pp. 68–84, 1998.

[5] P. Côté, S. Piatek, L. Ferrarese et al., "The ACS virgo cluster survey. VIII. The nuclei of early-type galaxies," *Astrophysical Journal*, vol. 165, no. 1, pp. 57–94, 2006.

[6] T. Böker, S. Laine, R. P. Van Der Marel et al., "A *Hubble Space Telescope* census of nuclear star clusters in late-type spiral galaxies. I. Observations and image analysis," *Astronomical Journal*, vol. 123, no. 3, pp. 1389–1410, 2002.

[7] A. W. Graham, C. A. Onken, E. Athanassoula, and F. Combes, "An expanded M_{bh}-σ diagram, and a new calibration of active galactic nuclei masses," *Monthly Notices of the Royal Astronomical Society*, vol. 412, no. 4, pp. 2211–2228, 2011.

[8] T. Böker, M. Sarzi, D. E. Mclaughlin et al., "A *Hubble Space Telescope* census of nuclear star clusters in late-type spiral galaxies. II. Cluster sizes and structural parameter correlations," *Astronomical Journal*, vol. 127, no. 1, pp. 105–118, 2004.

[9] C. J. Walcher, R. P. Van Der Marel, D. Mclaughlin et al., "Masses of star clusters in the nuclei of bulgeless spiral galaxies," *Astrophysical Journal*, vol. 618, no. 1, pp. 237–246, 2005.

[10] J. Rossa, R. P. Van Der Marel, T. Böker et al., "*Hubble Space Telescope* stis spectra of nuclear star clusters in spiral galaxies: dependence of age and mass on hubble type," *Astronomical Journal*, vol. 132, no. 3, pp. 1074–1099, 2006.

[11] A. C. Seth, J. J. Dalcanton, P. W. Hodge, and V. P. Debattista, "Clues to nuclear star cluster formation from edge-on spirals," *Astronomical Journal*, vol. 132, no. 6, pp. 2539–2555, 2006.

[12] C. J. Walcher, T. Böker, S. Charlot et al., "Stellar populations in the nuclei of late-type spiral galaxies," *Astrophysical Journal*, vol. 649, no. 2, pp. 692–708, 2006.

[13] B. Binggeli, F. Barazza, and H. Jerjen, "Off-center nuclei in dwarf elliptical galaxies," *Astronomy and Astrophysics*, vol. 359, no. 2, pp. 447–456, 2000.

[14] N. Neumayer, C. J. Walcher, D. Andersen, S. F. Sánchez, T. Böker, and H. -W. Rix, "Two-dimensional Hα kinematics of bulgeless disc galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 3, pp. 1875–1888, 2011.

[15] L. Ferrarese, P. Côté, E. Dalla Bontà et al., "A fundamental relation between compact stellar nuclei, supermassive black holes, and their host galaxies," *Astrophysical Journal Letters*, vol. 644, no. 1, pp. L21–L24, 2006.

[16] E. H. Wehner and W. E. Harris, "Supermassive black holes to dwarf elliptical nuclei: a mass continuum," *Astrophysical Journal Letters*, vol. 644, no. 1, pp. L17–L20, 2006.

[17] B. G. Elmegreen, F. Bournaud, and D. M. Elmegreen, "Nuclear black hole formation in clumpy galaxies at high redshift," *Astrophysical Journal Letters*, vol. 684, no. 2, pp. 829–834, 2008.

[18] J. A. Regan and M. G. Haehnelt, "Pathways to massive black holes and compact star clusters in pre-galactic dark matter haloes with virial temperatures $\gtrsim$10 000 K," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 1, pp. 343–353, 2009.

[19] L. Mayer, S. Kazantzidis, A. Escala, and S. Callegari, "Direct formation of supermassive black holes via multi-scale gas inflows in galaxy mergers," *Nature*, vol. 466, no. 7310, pp. 1082–1083, 2010.

[20] V. Bromm and N. Yoshida, "The first galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 49, no. 1, pp. 373–407, 2011.

[21] M. Hartmann, V. P. Debattista, A. Seth, M. Cappellari, and T. R. Quinn, "Constraining the role of star cluster mergers in nuclear cluster formation: simulations confront integral-field data," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 4, pp. 2697–2714, 2011.

[22] T. Ebisuzaki, J. Makino, and T. G. Tsuru, "Missing link found? The "runaway" path to supermassive black holes," *Astrophysical Journal Letters*, vol. 562, no. 1, pp. L19–L22, 2001.

[23] S. F. Portegies Zwart, H. Baumgardt, P. Hut, J. Makino, and S. L.W. McMillan, "Formation of massive black holes through runaway collisions in dense young star clusters," *Nature*, vol. 428, no. 6984, pp. 724–726, 2004.

[24] M. A. Gürkan, M. Freitag, and F. A. Rasio, "Formation of massive black holes in dense star clusters. I. Mass segregation and core collapse," *Astrophysical Journal Letters*, vol. 604, no. 2, pp. 632–652, 2004.

[25] M. Freitag, F. A. Rasio, and H. Baumgardt, "Runaway collisions in young star clusters—I. Methods and tests," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 1, pp. 121–140, 2006.

[26] E. Gaburov, A. Gualandris, and S. Portegies Zwart, "On the onset of runaway stellar collisions in dense star clusters—I. Dynamics of the first collision," *Monthly Notices of the Royal Astronomical Society*, vol. 384, no. 1, pp. 376–385, 2008.

[27] E. Glebbeek, E. Gaburov, S. E. De Mink, O. R. Pols, and S. F.P. Zwart, "The evolution of runaway stellar collision products," *Astronomy and Astrophysics*, vol. 497, no. 1, pp. 255–264, 2009.

[28] A. V. Filippenko and L. C. Ho, "A low-mass central black hole in the bulgeless Seyfert 1 galaxy NGC 4395," *Astrophysical Journal Letters*, vol. 588, no. 1, pp. L13–L16, 2003.

[29] A. Seth, M. Agüeros, D. Lee, and A. Basu-Zych, "The coincidence of nuclear star clusters and active galactic nuclei," *Astrophysical Journal Letters*, vol. 678, no. 1, pp. 116–130, 2008.

[30] R. Schödel, A. Eckart, T. Alexander et al., "The structure of the nuclear stellar cluster of the Milky Way," *Astronomy and Astrophysics*, vol. 469, no. 1, pp. 125–146, 2007.

[31] R. Genzel, F. Eisenhauer, and S. Gillessen, "The Galactic center massive black hole and nuclear star cluster," *Reviews of Modern Physics*, vol. 82, no. 4, pp. 3121–3195, 2010.

[32] M. Freitag, M. Atakan Gürkan, and F. A. Rasio, "Runaway collisions in young star clusters—II. Numerical results," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 1, pp. 141–161, 2006.

[33] A. W. Graham and S. P. Driver, "A log-quadratic relation for predicting supermassive black hole masses from the host bulge sérsic index," *Astrophysical Journal Letters*, vol. 655, no. 1, pp. 77–87, 2007.

[34] J. E. Greene, C. Y. Peng, M. Kim et al., "Precise black hole masses from megamaser disks: black hole-bulge relations at low mass," *Astrophysical Journal Letters*, vol. 721, no. 1, pp. 26–45, 2010.

[35] A. V. Filippenko and W. L. W. Sargent, "Discovery of an extremely low luminosity Seyfert 1 nucleus in the dwarf galaxy NGC 4395," *Astrophysical Journal*, vol. 342, pp. L11–L14, 1989.

[36] J. C. Shields, C. J. Walcher, T. Böker, L. C. Ho, H. -W. Rix, and R. P. Van Der Marel, "An accreting black hole in the nuclear star cluster of the bulgeless galaxy NGC 1042," *Astrophysical Journal Letters*, vol. 682, no. 1, pp. 104–109, 2008.

[37] A. J. Barth, L. E. Strigari, M. C. Bentz, J. E. Greene, and L. C. Ho, "Dynamical constraints on the masses of the nuclear star cluster and black hole in the late-type spiral galaxy NGC 3621," *Astrophysical Journal Letters*, vol. 690, no. 1, pp. 1031–1044, 2009.

[38] M. Gliozzi, S. Satyapal, M. Eracleous, L. Titarchuk, and C. C. Cheung, "A chandra view of NGC 3621: a bulgeless galaxy hosting an agn in its early phase?" *Astrophysical Journal Letters*, vol. 700, no. 2, pp. 1759–1767, 2009.

[39] S. Satyapal, D. Vega, R. P. Dudik, N. P. Abel, and T. Heckman, "*Spitzer* uncovers active galactic nuclei missed by optical surveys in seven late-type galaxies," *Astrophysical Journal Letters*, vol. 677, no. 2, pp. 926–942, 2008.

[40] J. E. Greene and L. C. Ho, "A new sample of low-mass black holes in active galaxies," *Astrophysical Journal Letters*, vol. 670, no. 1, pp. 92–104, 2007.

[41] A. J. Barth, J. E. Greene, and L. C. Ho, "Low-mass seyfert 2 galaxies in the sloan digital sky survey," *Astronomical Journal*, vol. 136, no. 3, pp. 1179–1200, 2008.

[42] W. McAlpine, S. Satyapal, M. Gliozzi, C. C. Cheung, R. M. Sambruna, and M. Eracleous, "Black holes in bulgeless galaxies: an XMM-Newton investigation of NGC 3367 and NGC 4536," *Astrophysical Journal Letters*, vol. 728, article 25, 2011.

[43] K. Gebhardt, T. R. Lauer, J. Kormendy et al., "M33: a galaxy with no supermassive black hole," *Astronomical Journal*, vol. 122, no. 5, pp. 2469–2476, 2001.

[44] D. Merritt, L. Ferrarese, and C. L. Joseph, "No supermassive black hole in M33?" *Science*, vol. 293, no. 5532, pp. 1116–1118, 2001.

[45] P. Erwin and D. Gadotti, "Do nuclear star clusters and supermassive black holes follow the same host-galaxy correlations?" in *Proceedings of the American Institute of Physics Conference*, V. P. Debattista and C. C. Popescu, Eds., vol. 1240, pp. 223–226, 2010.

[46] J. Kormendy, R. Bender, and M. E. Cornell, "Supermassive black holes do not correlate with galaxy disks or pseudobulges," *Nature*, vol. 469, no. 7330, pp. 374–376, 2011.

[47] M. Volonteri, G. Lodato, and P. Natarajan, "The evolution of massive black hole seeds," *Monthly Notices of the Royal Astronomical Society*, vol. 383, no. 3, pp. 1079–1088, 2008.

[48] T. Di Matteo, V. Springel, and L. Hernquist, "Energy input from quasars regulates the growth and activity of black holes and their host galaxies," *Nature*, vol. 433, no. 7026, pp. 604–607, 2005.

[49] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes, and galaxy spheroids," *Astrophysical Journal*, vol. 163, no. 1, pp. 1–49, 2006.

[50] C. Y. Peng, "How mergers may affect the mass scaling relation between geuvitationally bound systems," *Astrophysical Journal Letters*, vol. 671, no. 2, pp. 1098–1107, 2007.

[51] E. Noyola, K. Gebhardt, M. Kissler-Patig et al., "Very large telescope kinematics for omega Centauri: further support for a central black hole," *Astrophysical Journal Letters*, vol. 719, no. 1, pp. L60–L64, 2010.

[52] K. Jahnke and A. Macciò, "The Non-causal Origin of the Black-hole-galaxy Scaling Relations," *The Astrophysical Journal*, vol. 734, no. 2, article 92, 2011.

[53] E. Emsellem, G. Monnet, and R. Bacon, "The multi-gaussian expansion method: a tool for building realistic photometric and kinematical models of stellar systems I. The formalism," *Astronomy and Astrophysics*, vol. 285, pp. 723–738, 1994.

[54] M. Cappellari, "Efficient multi-Gaussian expansion of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 333, no. 2, pp. 400–410, 2002.

[55] J. Krist, *Astronomical Data Analysis Software and Systems IV*, vol. 77 of *Astronomical Society of the Pacific Conference Series*, 1995, edited by R. A. Shaw, H. E. Payne and J. J. E. Hayes.

[56] M. Cappellari, "Measuring the inclination and mass-to-light ratio of axisymmetric galaxies via anisotropic Jeans models of stellar kinematics," *Monthly Notices of the Royal Astronomical Society*, vol. 390, no. 1, pp. 71–86, 2008.

[57] P. Serra and S. C. Trager, "On the interpretation of the age and chemical composition of composite stellar populations determined with line-strength indices," *Monthly Notices of the Royal Astronomical Society*, vol. 374, no. 3, pp. 769–774, 2007.

[58] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *Astrophysical Journal Letters*, vol. 698, no. 1, pp. 198–221, 2009.

[59] B. M. Peterson, M. C. Bentz, L. -B. Desroches et al., "Multiwavelength monitoring of the dwarf seyfert 1 galaxy NGC 4395. I. A reverberation-based measurement of the black hole mass," *Astrophysical Journal*, vol. 632, no. 2, pp. 799–808, 2005.

[60] A. J. Barth, L. C. Ho, R. E. Rutledge, and W. L.W. Sargent, "POX 52: a dwarf Seyfert 1 galaxy with an intermediate-mass black hole," *Astrophysical Journal Letters*, vol. 607, no. 1, pp. 90–102, 2004.

[61] M. Valluri, L. Ferrarese, D. Merritt, and C. L. Joseph, "The low end of the supermassive black hole mass function: constraining the mass of a nuclear black hole in NGC 205 via stellar kinematics," *Astrophysical Journal*, vol. 628, no. 1, pp. 137–152, 2005.

[62] T. Böker, R. P. Van Der Marel, and W. D. Vacca, "CO band head spectroscopy of IC 342: mass and age of the nuclear star cluster 1," *Astronomical Journal*, vol. 118, no. 2, pp. 831–842, 1999.

[63] A. C. Seth, M. Cappellari, N. Neumayer et al., "The NGC 404 nucleus: star cluster and possible intermediate-mass black hole," *Astrophysical Journal Letters*, vol. 714, no. 1, pp. 713–731, 2010.

[64] K. Gebhardt, R. M. Rich, and C. H.O. Luis, "An intermediate-mass black hole in the globular cluster G1: improved significance from New Keck and *Hubble Space Telescope* observations," *Astrophysical Journal*, vol. 634, no. 2, pp. 1093–1102, 2005.

[65] E. Noyola, K. Gebhardt, and M. Bergmann, "Gemini and *Hubble Space Telescope* evidence for an intermediate-mass black hole in ω Centauri," *Astrophysical Journal Letters*, vol. 676, no. 2, pp. 1008–1015, 2008.

[66] B. Jalali, H. Baumgardt, M. Kissler-Patig et al., "A Dynamical N-body Model for the Central Region of ω Centauri," http://adsabs.harvard.edu/abs/2011arXiv1111.5011J .

[67] N. Lützgendorf, M. Kissler-Patig, E. Noyola et al., "Kinematic signature of an intermediate-mass black hole in the globular cluster NGC 6388," *Astronomy & Astrophysics*, vol. 533, artilce A36, 2011.

[68] J. Anderson and R. P. Van Der Marel, "New limits on an intermediate-mass black hole in omega centauri. I. *Hubble Space Telescope* photometry and proper motions," *Astrophysical Journal Letters*, vol. 710, no. 2, pp. 1032–1062, 2010.

[69] R. P. Van Der Marel and J. Anderson, "New limits on an intermediate-mass black hole in omega centauri. II. Dynamical models," *Astrophysical Journal Letters*, vol. 710, no. 2, pp. 1063–1088, 2010.

[70] H. Baumgardt, J. Makino, H. U. T. Piet, S. McMillan, and S. P. Zwart, "A dynamical model for the globular cluster G1," *Astrophysical Journal Letters*, vol. 589, no. 1, pp. L25–L28, 2003.

[71] J. E. Greene, L. C. Ho, and A. J. Barth, "Black holes in pseudobulges and spheroidals: a change in the black hole-bulge scaling relations at low mass," *Astrophysical Journal Letters*, vol. 688, no. 1, pp. 159–179, 2008.

[72] K. Ganda, R. F. Peletier, M. Balcells, and J. Falcón-Barroso, "The nature of late-type spiral galaxies: structural parameters, optical and near-infrared colour profiles and dust extinction," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1669–1694, 2009.

[73] T. Weinzirl, S. Jogee, S. Khochfar, A. Burkert, and J. Kormendy, "Bulge n and B/T in high-mass galaxies: constraints on the origin of bulges in hierarchical models," *Astrophysical Journal Letters*, vol. 696, no. 1, pp. 411–447, 2009.

[74] C. Y. Peng, L. C. Ho, C. D. Impey, and H.-W. Rix, "Detailed structural decomposition of galaxy images," *Astronomical Journal*, vol. 124, no. 1, pp. 266–293, 2002.

[75] A. W. Graham and L. R. Spitler, "Quantifying the coexistence of massive black holes and dense nuclear star clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 4, pp. 2148–2162, 2009.

[76] K. Gültekin, S. Tremaine, A. Loeb, and D. O. Richstone, "Observational Selection Effects and the M-σ Relation," *The Astrophysical Journal*, vol. 738, no. 1, article 17, 2011.

[77] C. Scorza and F. C. Van Den Bosch, "Nuclear stellar discs in early-type galaxies—II. Photometric properties," *Monthly Notices of the Royal Astronomical Society*, vol. 300, no. 2, pp. 469–478, 1998.

[78] K. Gebhardt and J. Thomas, "The black hole mass, stellar mass-to-light ratio, and Dark Halo in m87," *Astrophysical Journal Letters*, vol. 700, no. 2, pp. 1690–1701, 2009.

[79] P. J. Young, J. A. Westphal, J. Kristian, C. P. Wilson, and F. P. Landauer, "Evidence for a supermassive object in the nucleus of the galaxy M87 from SIT and CCD area photometry," *Astrophysical Journal*, vol. 221, pp. 721–730, 1978.

[80] G. A. Bower, R. F. Green, A. C. Quillen et al., "The ionization source in the nucleus of M84," *Astrophysical Journal Letters*, vol. 534, no. 1, pp. 189–200, 2000.

[81] J. L. Walsh, A. J. Barth, and M. Sarzi, "The supermassive black hole in M84 revisited," *Astrophysical Journal Letters*, vol. 721, no. 1, pp. 762–776, 2010.

[82] L. Ferrarese, H. C. Ford, and W. Jaffe, "Evidence for a massive black hole in the active galaxy NGC 4261 from bubble space telescope images and spectra," *Astrophysical Journal Letters*, vol. 470, no. 1, pp. 444–459, 1996.

[83] S. P. Rusli, J. Thomas, P. Erwin, R. P. Saglia, N. Nowak, and R. Bender, "The central black hole mass of the high-σ but low-bulge-luminosity lenticular galaxy NGC 1332," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 2, pp. 1223–1236, 2011.

[84] N. Devereux, H. Ford, Z. Tsvetanov, and G. Jacoby, "STIS spectroscopy of the central 10 parsecs of M81: evidence for a massive black hole," *Astronomical Journal*, vol. 125, no. 3, pp. 1226–1235, 2003.

[85] M. J. Jee, J. P. Blakeslee, M. Sirianni, A. R. Martel, R. L. White, and H. C. Ford, "Principal component analysis of the time- and position-dependent point-spread function of the advanced camera for surveys," *Publications of the Astronomical Society of the Pacific*, vol. 119, no. 862, pp. 1403–1409, 2007.

[86] L. Ferrarese, P. Côté, A. Jordán et al., "The ACS Virgo Cluster survey. VI. Isophotal analysis and the structure of early-type

galaxies," *Astrophysical Journal, Supplement Series*, vol. 164, no. 2, pp. 334–434, 2006.

[87] T. R. Lauer, K. Gebhardt, S. M. Faber et al., "The centers of early-type galaxies with *Hubble Space Telescope*. VI. Bimodal central surface brightness profiles," *Astrophysical Journal*, vol. 664, no. 1, pp. 226–256, 2007.

[88] G. De Francesco, A. Capetti, and A. Marconi, "Measuring supermassive black holes with gas kinematics: the active so galaxy NGC 3998," *Astronomy and Astrophysics*, vol. 460, no. 2, pp. 439–448, 2006.

[89] N. Cretton and F. C. Van Den Bosch, "Evidence for a massive black hole in the S0 galaxy NGC 4342," *Astrophysical Journal Letters*, vol. 514, no. 2, pp. 704–724, 1999.

[90] K. Gebhardt, D. Richstone, J. Kormendy et al., "Axisymmetric, three-integral models of galaxies: a massive black hole in NGC 3379," *Astronomical Journal*, vol. 119, no. 3, pp. 1157–1171, 2000.

[91] D. Merritt and A. Szell, "Dynamical cusp regeneration," *Astrophysical Journal*, vol. 648, no. 2, pp. 890–899, 2006.

[92] M. Koleva, P. Prugniel, S. De Rijcke, and W. W. Zeilinger, "Age and metallicity gradients in early-type galaxies: a dwarf-to-giant sequence," *Monthly Notices of the Royal Astronomical Society*, vol. 417, no. 3, pp. 1643–1671, 2011.

[93] H. Zinnecker, C. J. Keable, J. S. Dunlop, R. D. Cannon, and W. K. Griths, "The harlow-shapley symposium on globular cluster systems in galaxies," in *Proceedings of the 126th Symposium of the International Astronomical Union (IAU '88)*, J. E. Grindlay and A. G. D. Philip, Eds., vol. 126, p. 603, 1988.

[94] K. C. Freeman, *The Globular Cluster-Galaxy Connection*, vol. 48 of *Astronomical Society of the Pacific Conference Series*, 1993, edited by G. H. Smith and J. P. Brodie.

[95] K. Bekki and K. C. Freeman, "Formation of ω Centauri from an ancient nucleated dwarf galaxy in the young Galactic disc," *Monthly Notices of the Royal Astronomical Society*, vol. 346, no. 2, pp. L11–L15, 2003.

[96] T. Böker, "Are globular clusters the remnant nuclei of progenitor disk galaxies?" *Astrophysical Journal Letters*, vol. 672, no. 2, pp. L111–L114, 2008.

[97] M. Milosavljević, "On the origin of nuclear star clusters in late-type spiral galaxies," *Astrophysical Journal Letters*, vol. 605, no. 1, pp. L13–L16, 2004.

[98] K. Bekki, W. J. Couch, and Y. Shioya, "Dissipative transformation of nonnucleated dwarf galaxies into nucleated systems," *Astrophysical Journal Letters*, vol. 642, no. 2, pp. L133–L136, 2006.

[99] J. Pflamm-Altenburg and P. Kroupa, "Recurrent gas accretion by massive star clusters, multiple stellar populations and mass thresholds for spheroidal stellar systems," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 1, pp. 488–494, 2009.

[100] S. D. Tremaine, J. P. Ostriker, and L. Spitzer, "The formation of the nuclei of galaxies. I—M31," *Astrophysical Journal*, vol. 196, pp. 407–411, 1975.

[101] R. Capuzzo-Dolcetta, "The evolution of the globular cluster system in a triaxial galaxy: can a galactic nucleus form by globular cluster capture?" *Astrophysical Journal Letters*, vol. 415, no. 2, pp. 616–630, 1993.

[102] D. E. Mclaughlin, "Was the compact nucleus in M87 formed by destroyed globular clusters?" *Astronomical Journal*, vol. 109, no. 5, pp. 2034–2037, 1995.

[103] K. Bekki, W. J. Couch, M. J. Drinkwater, and Y. Shioya, "Cluster cannibalism and scaling relations of galactic stellar nuclei," *Astrophysical Journal Letters*, vol. 610, no. 1, pp. L13–L16, 2004.

[104] P. Miocchi, R. C. Dolcetta, P. Di Matteo, and A. Vicari, "Merging of globular clusters in inner galactic regions. I. Do they survive the tidal interaction?" *Astrophysical Journal Letters*, vol. 644, no. 2, pp. 940–953, 2006.

[105] M. Agarwal and M. Milosavljević, "Nuclear star clusters from clustered star formation," *Astrophysical Journal Letters*, vol. 729, article 35, 2011.

[106] S. Nayakshin, M. I. Wilkinson, and A. King, "Competitive feedback in galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 1, pp. L54–L57, 2009.

[107] D. Merritt, "Dynamics of galaxy cores and supermassive black holes," *Reports on Progress in Physics*, vol. 69, no. 9, pp. 2513–2579, 2006.

[108] K. Bekki and A. W. Graham, "On the transition from nuclear-cluster- to black-hole-dominated galaxy cores," *Astrophysical Journal Letters*, vol. 714, no. 2, pp. L313–L317, 2010.

[109] K. Gebhardt, J. Adams, D. Richstone et al., "The black hole mass in M87 from Gemini/NIFS adaptive optics observations," *Astrophysical Journal Letters*, vol. 729, article 119, 2011.

[110] G. A. Bower, R. F. Green, A. Danks et al., "Kinematics of the nuclear ionized gas in the radio galaxy M84 (NGC 4374)," *Astrophysical Journal Letters*, vol. 492, no. 2, pp. L111–L114, 1998.

[111] K. Gebhardt, D. Richstone, S. Tremaine et al., "Axisymmetric dynamical models of the central regions of galaxies," *Astrophysical Journal Letters*, vol. 583, no. 1, pp. 92–115, 2003.

[112] M. Sarzi, H.-W. Rix, J. C. Shields et al., "Supermassive black holes in bulges," *Astrophysical Journal Letters*, vol. 550, no. 1, pp. 65–74, 2001.

[113] K. L. Shapiro, M. Cappellari, T. De Zeeuw et al., "The black hole in NGC 3379: a comparison of gas and stellar dynamical mass measurements with HST and integral-field data," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 2, pp. 559–579, 2006.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 946368, 11 pages
doi:10.1155/2012/946368

Research Article

Do Nuclear Star Clusters and Supermassive Black Holes Follow the Same Host-Galaxy Correlations?

Peter Erwin[1,2] and Dimitri Alexei Gadotti[3]

[1] *Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany*
[2] *Universitäts-Sternwarte München, Scheinerstraße 1, 81679 München, Germany*
[3] *European Southern Observatory, Alonso de Cordova 3107, Vitacura, Casilla 19001, Santiago 19, Chile*

Correspondence should be addressed to Peter Erwin, erwin@mpe.mpg.de

Received 6 October 2011; Revised 1 December 2011; Accepted 6 December 2011

Academic Editor: Isabelle Gavignaud

Studies have suggested that there is a strong correlation between the masses of nuclear star clusters (NSCs) and their host galaxies, a correlation which is said to be an extension of the well-known correlations between supermassive black holes (SMBHs) and their host galaxies. But careful analysis of disk galaxies—including 2D bulge/disk/bar decompositions—shows that while SMBHs correlate with the stellar mass of the *bulge* component of galaxies, the masses of NSCs correlate much better with the *total* galaxy stellar mass. In addition, the mass ratio $M_{\rm NSC}/M_{\star,\rm tot}$ for NSCs in spirals (at least those with Hubble types Sc and later) is typically an order of magnitude smaller than the mass ratio $M_{\rm BH}/M_{\star,\rm bul}$ of SMBHs. The absence of a universal "central massive object" correlation argues against common formation and growth mechanisms for both SMBHs and NSCs.We also discuss evidence for a break in the NSC-host galaxy correlation, galaxies with Hubble types earlier than Sbc appear to host systematically more massive NSCs than do types Sc and later.

1. Introduction

As far as we can tell, all massive galaxies in the local universe harbor supermassive black holes (SMBHs, with masses $M_{\rm BH} \sim 10^6$–$10^9 M_\odot$). The masses of these SMBHs correlate strongly with several global properties of the host galaxies, particularly with the central velocity dispersion σ_0 [1, 2] and with the bulge luminosity or mass (e.g., [3, 4]). These correlations imply that the processes which drive galaxy growth and the processes which drive black hole growth are intimately linked—perhaps even the *same* processes.

It is now also clear that many galaxies, particularly later-type spirals, host luminous nuclear star clusters (NSCs; e.g., [5, 6]), with masses in the range 10^5–$10^8 M_\odot$; see the review by Böker [7] for more details. Recently, several authors have argued that NSCs and central SMBHs have the *same* host-galaxy correlations, in particular, that SMBHs and NSCs have the same correlation with bulge luminosity and mass [8–11] (but see Balcells et al. [12]). The suggestion, then, is that NSCs and SMBHs are in a sense members of the same family of "Central Massive Objects" (CMOs), and thus that they may have grown via the same mechanisms (e.g., [13–16]).

We argue, however, that one should be cautious about assuming that NSCs and SMBHs are really part of the same family, with the same host-galaxy relationships. To begin with, the samples of Wehner and Harris [8] and Ferrarese et al. [9], which were used to make the CMO argument, were almost entirely early-type galaxies—mostly ellipticals and dwarf ellipticals. These are galaxies which are, in essence, "pure bulge" systems, so one could just as easily argue for a correlation with *total* galaxy mass. But we know that SMBHs in *disk* galaxies correlate better with just the bulge, and *not* with the total galaxy mass or light (e.g., [17, 18]). Given that there have been previous claims that NSCs in spiral galaxies correlate with the *total* galaxy light (e.g., [19]), we are prompted ask the question: do nuclear clusters *in disk galaxies* correlate with the bulge (like SMBHs), or with the whole galaxy?

2. Samples, Methodology, and Data Sources

Although current studies suggest that the $M_{\rm BH}$-σ_0 relation is tighter and has less intrinsic scatter than the $M_{\rm BH}$-$M_{\star,\rm bul}$ relation (e.g., [20]), velocity dispersion is *not* the ideal host-galaxy measure to use here, for three reasons. First, most of the best-determined NSC masses are based directly on the measured velocity dispersion of the NSC (e.g., [21]), which is often indistinguishable from that of the surrounding bulge; this means a (spurious) correlation between NSC mass, and central velocity dispersion is only to be expected. Second, some NSCs are found in galaxies with *no* detectable bulge at all (see discussion in Section 3). Finally, it is difficult to see how one should discriminate between a velocity dispersion due to the bulge versus one due to the whole galaxy. But discriminating between bulge and whole-galaxy luminosities and masses is much simpler. So we choose instead to compare NSCs and their host galaxies with the $M_{\rm BH}$-$M_{\star,\rm bul}$ relation, which means comparing NSC masses with the *stellar masses* of host galaxies and their bulges.

For NSCs, we emphasize galaxies where the NSC masses have been *dynamically* measured, since this is the most direct analog to well-determined SMBH masses (i.e., those with direct dynamical mass measurements from stellar, gas, or maser kinematics, where the SMBH sphere of influence is resolved). In addition, dynamical measurements avoid possible problems with multiple stellar populations; the latter can potentially bias stellar masses estimated from broadband colors. Spectroscopic studies [11, 22] have shown that NSCs often contain multiple stellar populations; this renders mass estimates based on single stellar population (SSP) models (e.g., those used by [9]) somewhat uncertain. The NSCs we focus on are taken primarily from the sample of Walcher et al. [21], with additional data from Ho and Filippenko [23], Böker et al. [24], Kormendy and Bender [25], Matthews et al. [26] and Gebhardt et al. [27], Barth et al. [28], Seth et al. [29], and Kormendy et al. [30]; we use the estimate of Launhardt et al. [31] for the Milky Way's NSC. This gives us a total of 18 galaxies with dynamically determined NSC masses. These cover Hubble types S0–Sm, but the sample is in fact heavily biased towards later types; over three-quarters are Hubble types Second or later. As an additional, secondary sample, we include 15 galaxies from Rossa et al. [11], where the masses are estimated by fits of multiple SSP models to high-resolution spectroscopy. Most of these galaxies are Sc and later, but a few earlier-type spirals (Sa–Sb) are also included.

Total stellar masses are based on K-band total magnitudes from 2MASS [32] or from Malhotra et al. [33] for M31 and M33 (which are too large for accurate sky subtraction of 2MASS images), combined with color-based mass-to-light (M/L) ratios from Bell et al. [34]. For the latter, we use optical colors from the literature (primarily from HyperLeda (http://leda.univ-lyon1.fr/)) or from direct measurements on Sloan Digital Sky Survey (SDSS, [35]) images. The *bulge* masses are derived using bulge-to-total (B/T) values determined individually for each galaxy by 2D image decomposition, using the BUDDA software package [36, 37], which incorporates bulge and disk components *and* optional bars and central point sources (the latter can be used for both nuclear star clusters and AGN). Note that we explicitly define "bulge" to be the "photometric bulge", that is, the excess light (and stellar mass) which is not part of the disk, bar, or nuclear star cluster. We defer questions of how SMBH (or nuclear cluster) mass relates to so-called "pseudobulges" versus "classical bulges" (e.g., [38, 39]) to a later analysis.

Full 2D decompositions, as described above, were used for all S0 and spiral SMBH host galaxies. For the NSC host galaxies, we follow the same approach, with one simplification. Since we have found that B/T ratios for *unbarred* galaxies do not change dramatically if we use 1D surface-brightness profile decompositions instead of 2D image decompositions, we use the former for genuinely unbarred galaxies; we are careful to exclude (or separately model) the NSC contribution to the surface-brightness profile in these cases. Galaxies which *do* possess bars are subjected to full 2D decompositions; see the following section for details.

2.1. Bulge-Disk Decompositions. As noted previously, we use 2D image decompositions via the BUDDA software package to determine the B/T ratios, and thus the bulge stellar masses, for SMBH host galaxies and for barred NSC host galaxies. For the NSC galaxies, we use *HST* data wherever possible, to enable the NSC itself to be properly modeled as a separate source. However, we have found that when the NSC is sufficiently luminous, and when the bulge is sufficiently low-contrast, we can achieve reasonable decompositions with ground-based images; these are sometimes preferable if they are near-IR (to minimize the effects of dust extinction and recent star formation) and/or large enough to include the entire galaxy (to allow better recovery of the disk component).

We have completed decompositions for the galaxies with dynamically determined NSC masses (we use the published 2D decomposition of [28] for NGC 3621); in the special case of the Milky Way, we assume a bulge mass of $\sim 1.0 \times 10^{10} M_\odot$ and a total stellar mass of $5.5 \times 10^{10} M_\odot$, based on arguments in Dehnen and Binney [40], Klypin et al. [41], and Flynn et al. [42]. Decompositions for the spectroscopic sample are still in progress but are mostly complete; since our primary analysis (e.g., computing the $M_{\rm NSC}$-$M_{\star,\rm bul}$ relation) is based on the dynamical masses, the incompleteness of the spectroscopic sample does not affect our results.

Full details of the individual decompositions will be published elsewhere (Erwin and Gadotti 2012a, in prep.). An example of one of the 2D decompositions is given in Figure 1 for the galaxy NGC 7418, where we fit an H-band image from the Ohio State University Bright Spiral Galaxy Survey [43, OSU BSGS] using an exponential disk (95.6% of the light), a Sérsic bulge (1.6% of the light), a bar (2.4% of the light), and a point source for the NSC (0.5% of the light). (In this particular galaxy, the disk appears to be truncated, but this has little effect on the decomposition; including a broken-exponential profile for the disk changes the B/T ratio from 0.016 to 0.017.) This illustrates the importance of including a separate bar component in the decomposition when the galaxy is barred; the bar has almost twice the luminosity of the bulge, and, without it, the bulge luminosity

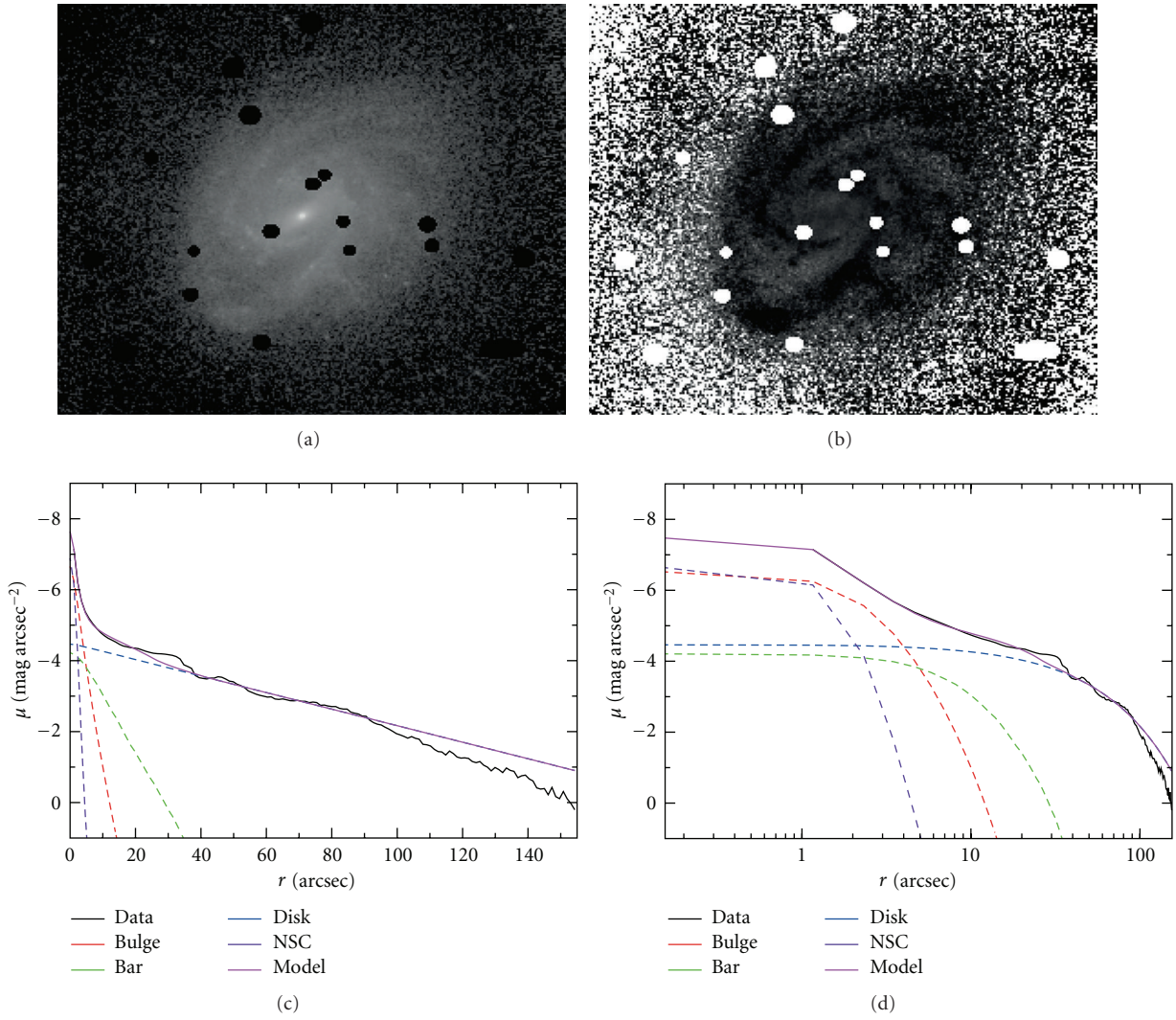


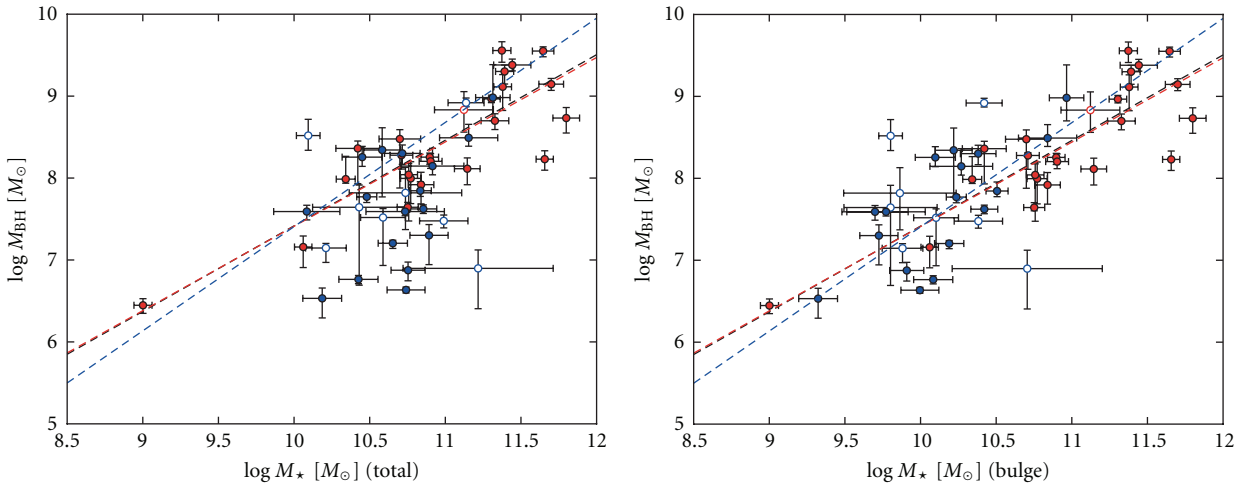
FIGURE 1: An example of one of our 2D decompositions of NSC host galaxies—in this instance, the decomposition of the OSU BSGS H-band image of NGC 7418, using an exponential disk, a Sérsic bulge ($n = 1.5$), a bar, and a point source for the NSC. (a): original H-band image, with masking of bright stars (logarithmic brightness scaling). (b): residual image after subtracting best-fitting model image. (c): major-axis profile (black) along with the components of the model and their sum (purple). (d): same, but plotted with logarithmic major-axis scaling.

(and stellar mass) would certainly be overestimated. In fact, a 1D decomposition for this galaxy gives a B/T value almost twice as large (0.030); similar results were found for four other barred galaxies in the sample, with mean B/T values a factor of 2.1 times larger when the bar was omitted; see also [37].

3. Comparing Black Holes and Nuclear Star Clusters

Although black-hole–bulge correlations are sometimes described as correlations between the black hole mass and the host *galaxy* mass (or luminosity as a proxy for mass), this is really only true for elliptical galaxies, where the entire galaxy is the "bulge." Kormendy and Gebhardt [17] explicitly compared B-band total and bulge luminosities for SMBH hosts and showed that the latter provided a much better correlation. Most recently, Kormendy et al. [18] have show for a larger, updated sample that SMBH masses in disk galaxies correlated much better with (classical) bulge K-band luminosity than with the luminosity of the disk component; this naturally suggests that total-galaxy luminosity is unlikely to correlate well with SMBH mass when the galaxy is disk dominated.

In Figure 2, we compare SMBH masses with total galaxy stellar mass (Figure 2(a)) and with bulge stellar mass (Figure 2(b)), based on our careful bulge/disk/bar decompositions (see Table 1). Error bars include the effects of uncertainties in the distance and in the M/L and B/T ratios. As expected, the correlation between SMBH mass and *bulge* mass is much stronger than any correlation with total galaxy mass; the Spearman correlation coefficients are $r_S = 0.71$ for the $M_{\rm BH}$-$M_{\star,\rm bul}$ relation versus 0.29 for the $M_{\rm BH}$-$M_{\star,\rm tot}$



(a) (b)

FIGURE 2: (a): SMBH mass (red: elliptical galaxies, blue: disk galaxies) versus total galaxy stellar mass. (b): SMBH mass versus bulge stellar mass. (Data and sources in Table 1). The diagonal dashed lines are the best fits to the M_{BH}-$M_{\star,bul}$ relation for the whole sample (black), for the elliptical galaxies (red), and for bulges of the disk galaxies (blue). Open symbols are galaxies without precise distances, which are not used in the fits. It is clear that the SMBH masses of S0 and spiral galaxies (blue) correlate better with the *bulge* stellar mass than with total galaxy mass.

relation, with the latter correlation not being statistically significant.

We also plot linear fits of $\log M_{BH}$ as a function of $\log M_{\star,bul}$; these fits are made using galaxies with well-determined distances (filled points) to minimize distance-based uncertainties, using the Bayesian-based approach of D'Agostini [44], which explicitly incorporates errors in both variables and intrinsic scatter in the black hole mass (see also [45, 46]). By "well-determined distances," we mean those determined using direct methods such as surface-brightness fluctuations and Cepheid stars, or redshift-based distances, where $z > 0.01$ (to avoid large relative uncertainties due to peculiar motions.) The best-fitting relation for the whole sample (black line; the fit to just the elliptical galaxies, shown by the red line, is almost identical) is

$$\log M_{BH} = 8.46 \pm 0.08 + (1.04 \pm 0.12)\log(M_{\star,bul}/10^{11}M_\odot), \quad (1)$$

with intrinsic scatter in SMBH mass of 0.39 ± 0.05 dex; the best-fitting relation for the bulges of disk galaxies only (blue line) is

$$\log M_{BH} = 8.68 \pm 0.20 + (1.27 \pm 0.26)\log(M_{\star,bul}/10^{11}M_\odot), \quad (2)$$

with intrinsic scatter $= 0.41 \pm 0.07$. (The errors are based on bootstrap resampling.)

We apply exactly the same methodology to NSC-host galaxies in Figure 3, plotting NSC mass versus total galaxy stellar mass in the left panel and versus bulge stellar mass in the right panel. Since several of the NSC host galaxies are genuinely *bulgeless* systems (without even a distinct "pseudobulge"), we plot their bulge masses as upper limits ($B/T < 0.001 M_{\star,tot}$). As the figure shows, NSC mass clearly correlates better with *total* stellar mass than it does with bulge mass. (The respective correlation coefficients are $r_S = 0.76$ versus 0.38; the bulge-mass correlation is not statistically significant.) Fitting NSC mass versus total stellar mass, using the same methodology as for the SMBH fits, gives the following relation:

$$\log M_{NSC} = 7.65 \pm 0.23 + (0.90 \pm 0.21)\log(M_{\star,tot}/10^{11}M_\odot), \quad (3)$$

with intrinsic scatter $= 0.43 \pm 0.10$ dex. Note that the slope is formally indistinguishable from unity; that is, the $M_{NSC}/M_{\star,tot}$ ratio does not appear to depend on $M_{\star,tot}$ itself.

It is important to note that the difference in correlation coefficients actually *underestimates* the true difference between the two relations, because the M_{NSC}-$M_{\star,bul}$ correlation was computed assuming that bulgeless spirals still have nominal bulges (using $B/T = 0.001$). In the combined sample of dynamical and spectroscopic NSC masses, we can identify at least three galaxies which have no detectable bulge. In two of these (NGC 1493 and NGC 2139), our 2D decomposition assigned stellar light to a bar in addition to a pure exponential disk; in 1D decompositions (or simple bulge + disk 2D decompositions), light from the bar might be (wrongly, we would argue) interpreted as "bulge" light. For the other galaxy (NGC 300), however, there is no ambiguity; this is an unbarred spiral galaxy with a surface brightness profile consisting of *only* an exponential disk and the NSC (see, e.g., Figure 8 of [47]).

The existence of nuclear star clusters in genuinely bulgeless spirals is simply an additional, direct confirmation

Table 1: Galaxies with well-determined SMBH masses.

Name	T	D (Mpc)	$M_{\rm BH}(+,-)$ ($\log_{10} M_\odot$)	Source	Total $M_\star$ (err) ($\log_{10} M_\odot$)	Bulge $M_\star$ (err) ($\log_{10} M_\odot$)
Milky way	4	0.01	6.63 (+0.03, −0.04)	1	10.74 (0.09)	10.00 (0.13)
M31	3	0.77	8.15 (+0.22, −0.10)	2	10.92 (0.06)	10.27 (0.21)
M32	−5	0.79	6.45 (+0.08, −0.10)	3	9.00 (0.06)	9.00 (0.06)
NGC524	−1	23.3	8.92 (+0.04, −0.02)	4	11.14 (0.09)	10.42 (0.12)
NGC821	−5	23.4	7.92 (+0.15, −0.23)	5	10.84 (0.08)	10.84 (0.08)
NGC1023	−1	11.1	7.62 (+0.04, −0.04)	6	10.85 (0.08)	10.42 (0.09)
NGC1068	3	14.3	6.90 (+0.14, −0.21)	7	11.22 (0.49)	10.71 (0.50)
NGC1300	4	18.9	7.82 (+0.29, −0.29)	8	10.74 (0.37)	9.86 (0.37)
NGC1316	−5	21.3	8.23 (+0.10, −0.13)	9	11.66 (0.06)	11.66 (0.06)
NGC1399	−5	21.1	9.11 (+0.15, −0.29)	10	11.38 (0.06)	11.38 (0.06)
NGC2549	−2	12.3	7.15 (+0.02, −0.16)	4	10.21 (0.12)	9.88 (0.13)
NGC2748	4	23.1	7.64 (+0.25, −0.74)	8	10.43 (0.30)	9.80 (0.31)
NGC2787	−1	7.28	7.59 (+0.04, −0.06)	11	10.08 (0.15)	9.70 (0.22)
NGC3031	1	3.63	7.85 (+0.11, −0.07)	12	10.83 (0.06)	10.51 (0.07)
NGC3227	1	22.9	7.30 (+0.13, −0.35)	13	10.89 (0.11)	9.72 (0.13)
NGC3245	−1	20.3	8.30 (+0.10, −0.12)	14	10.72 (0.09)	10.38 (0.12)
NGC3368	2	10.5	6.88 (+0.09, −0.12)	15	10.75 (0.09)	9.91 (0.11)
NGC3377	−5	10.9	7.99 (+0.28, −0.05)	5	10.34 (0.06)	10.34 (0.06)
NGC3379	−5	10.3	8.00 (+0.20, −0.31)	5	10.77 (0.07)	10.77 (0.07)
NGC3384	−1	11.3	7.20 (+0.03, −0.05)	5	10.65 (0.08)	10.19 (0.10)
NGC3393	1	48.3	7.48 (+0.03, −0.03)	16	10.99 (0.15)	10.38 (0.16)
NGC3489	−1	11.7	6.76 (+0.04, −0.04)	15	10.43 (0.08)	10.08 (0.13)
NGC3585	−3	19.5	8.49 (+0.16, −0.09)	17	11.15 (0.09)	10.84 (0.19)
NGC3607	−5	22.2	8.11 (+0.13, −0.19)	17	11.15 (0.08)	11.15 (0.08)
NGC3608	−5	22.3	8.28 (+0.02, −0.16)	5	10.71 (0.08)	10.71 (0.08)
NGC3998	−2	13.7	8.34 (+0.27, −0.56)	18	10.58 (0.09)	10.22 (0.12)
NGC4026	−3	13.2	8.26 (+0.12, −0.09)	17	10.45 (0.12)	10.10 (0.13)
NGC4151	2	14.5	7.52 (+0.10, −0.56)	19	10.59 (0.13)	10.10 (0.15)
NGC4258	4	7.18	7.59 (+0.04, −0.04)	20	10.73 (0.08)	9.77 (0.26)
NGC4261	−5	30.8	8.70 (+0.08, −0.10)	21	11.33 (0.09)	11.33 (0.09)
NGC4291	−5	25.5	8.48 (+0.10, −0.57)	5	10.70 (0.14)	10.70 (0.14)
NGC4342	−1	16.7	8.52 (+0.20, −0.18)	22	10.09 (0.07)	9.80 (0.08)
NGC4374	−5	18.5	8.97 (+0.04, −0.04)	23	11.30 (0.06)	11.30 (0.06)
NGC4473	−5	15.3	8.04 (+0.13, −0.56)	5	10.76 (0.06)	10.76 (0.06)
NGC4486	−5	16.7	9.56 (+0.11, −0.14)	24	11.37 (0.06)	11.37 (0.06)
NGC4486A	−5	18.4	7.16 (+0.13, −0.25)	25	10.06 (0.06)	10.06 (0.06)
NGC4564	−3	15.9	7.77 (+0.02, −0.07)	5	10.48 (0.06)	10.24 (0.07)
NGC4649	−5	16.4	9.30 (+0.08, −0.15)	26	11.39 (0.06)	11.39 (0.06)
NGC4697	−5	12.5	8.26 (+0.05, −0.09)	5	10.90 (0.06)	10.90 (0.06)
NGC5077	−5	37.5	8.83 (+0.21, −0.23)	27	11.12 (0.19)	11.12 (0.19)
NGC5128	−5	3.42	7.64 (+0.06, −0.03)	28	10.75 (0.07)	10.75 (0.07)
NGC5252	−2	92.9	8.98 (+0.40, −0.27)	29	11.31 (0.09)	10.97 (0.11)
NGC5576	−5	24.8	8.20 (+0.09, −0.08)	17	10.90 (0.08)	10.90 (0.08)
NGC5845	−5	25.2	8.36 (+0.07, −0.41)	5	10.42 (0.15)	10.42 (0.15)
NGC6251	−5	95.9	8.73 (+0.12, −0.18)	30	11.80 (0.09)	11.80 (0.09)
NGC7052	−5	67.9	8.58 (+0.23, −0.22)	31	11.49 (0.12)	11.49 (0.12)
NGC7457	−1	12.9	6.53 (+0.12, −0.23)	5	10.19 (0.10)	9.32 (0.13)
IC1459	−5	28.4	9.38 (+0.05, −0.04)	32	11.44 (0.12)	11.44 (0.12)
IC4296	−5	53.2	9.15 (+0.06, −0.07)	33	11.70 (0.08)	11.70 (0.08)
A1836-BCG	−5	155.6	9.55 (+0.05, −0.06)	33	11.65 (0.07)	11.65 (0.07)

(1) Galaxy name. (2) Hubble type T from RC3. (3) Adopted distance in Mpc. (4) Logarithm of SMBH mass and uncertainties; masses have been rescaled using the distances column 2, if necessary. Uncertainties are 1-σ values. (5) Source of SMBH measurement. (6) Logarithm of total galaxy stellar mass and uncertainty (see text for details). (7) Logarithm of bulge stellar mass and uncertainty, based on 2D decompositions in Erwin and Gadotti (2012b, in prep).

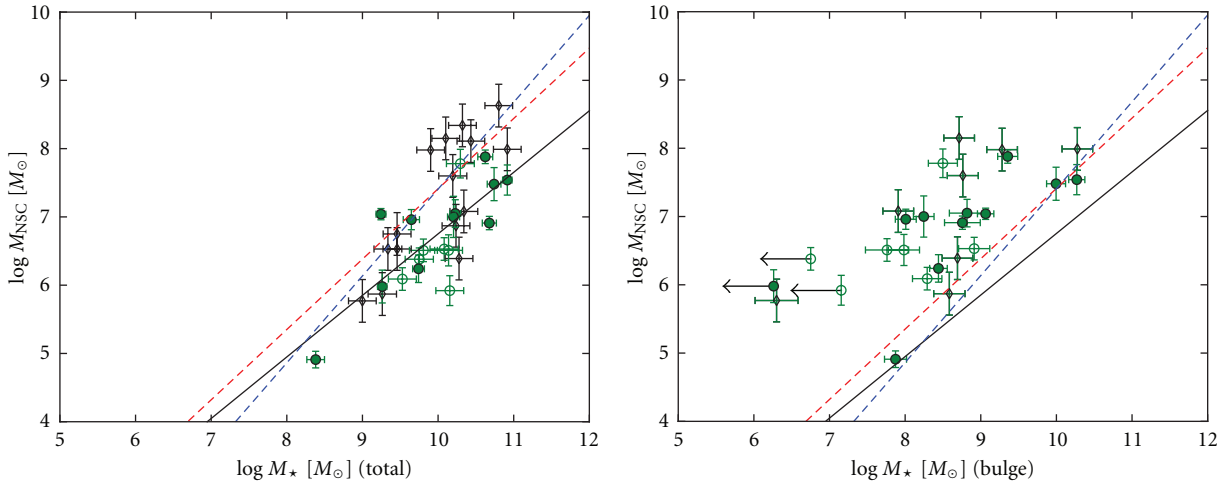


FIGURE 3: As for Figure 2, but now plotting NSC mass versus total stellar mass (a) and bulge stellar mass (b) (Data and sources in Table 2). Green circles are galaxies with dynamical mass estimates for their NSCs; black diamonds are the spectroscopically estimated masses of Rossa et al. [11] (bulge mass estimates are not complete for these galaxies). Filled symbols indicate galaxies with direct distance measurements (e.g., from Cepheid stars). Arrows show nominal upper limits for three *bulgeless* spirals (assuming that $B/T \leq 0.001$). The diagonal black line is a fit of NSC mass to total stellar mass for the dynamical-mass sample (green circles); for comparison, the diagonal dashed red and blue lines are the $M_{\rm BH}$-$M_{\star,\rm bul}$ fits for ellipticals (red) and disk galaxies (blue) from Figure 2. The situation is now the reverse of that for SMBHs: NSC masses clearly correlate better with *total* galaxy mass than they do with bulge mass.

of our basic conclusion: nuclear star cluster masses scale with the total stellar mass of their host galaxies, *not* with the bulge mass. This means that NSCs and SMBHs do not follow a common host-galaxy correlation.

We have also investigated whether other galaxy parameters might correlate with NSC mass, or even with residuals from the $M_{\rm NSC}$-$M_{\star,\rm tot}$ relation. In particular, we have compared NSC mass with rotation velocity and with total *baryonic* mass (stellar mass plus atomic gas from HI measurements). In both cases, correlations exist, but they are not as strong as the correlation with total stellar mass. No particular correlations with residuals of the $M_{\rm NSC}$-$M_{\star,\rm tot}$ relation are seen.

4. Trends with Hubble Type

Closer inspection of the left-hand panel of Figure 3 suggests that the spectroscopic masses (black diamonds) tend to be offset from the NSC-$M_{\star,\rm tot}$ relation, in the sense that they have larger NSC masses for the same total stellar mass. This could, in principle, be evidence of a systematic overestimation of NSC masses in the spectroscopic sample, but of the four galaxies in common between Walcher et al. [21] and Rossa et al. [11] only one has a (slightly) higher spectroscopic mass, while the other three have spectroscopic masses slightly *lower* than the dynamical masses. There is, however, another difference to consider; the spectroscopic sample tends to have earlier Hubble types.

This brings us to something which Seth et al. [48] pointed out several years ago, using a larger dataset of NSCs and host galaxies, with NSC masses based (mostly) on colors or assumed M/L ratios. They noted that NSCs in late-type spirals tended to have lower relative masses ($M_{\rm NSC}/M_{\star,\rm tot}$) than early-type spirals and ellipticals. (Rossa et al. [11] pointed out a similar trend in *absolute* NSC mass for their smaller sample of NSCs in early- and late-type spirals.) Figure 4 makes this explicit by plotting $M_{\rm NSC}/M_{\star,\rm tot}$ versus Hubble type for the galaxies in Seth et al.'s compilation, plus seven galaxies from our updated dynamical-mass sample which were not in their sample. We have also added galaxy stellar-mass estimates for 16 galaxies that did not have masses in Seth et al., using total K-band magnitudes from 2MASS and either B–V colors from HyperLeda or measured g–r colors from SDSS images to derive the K-band M/L via Bell et al. [34].

What is curious about Figure 4 is not just that the $M_{\rm NSC}/M_{\star,\rm tot}$ ratio depends on Hubble type, but that it actually appears to do so in a *bimodal* fashion; Hubble types Sb and earlier have relatively large NSC masses, while Sc and later-type galaxies have significantly smaller relative NSC masses. Plotted on top of the figure are simple fits of a function where the $M_{\rm NSC}/M_{\star,\rm tot}$ ratio can take two constant values, one for Hubble types $T < T_1$ and the other for $T > T_2$, with a simple linear transition between T_1 and T_2. Fits to just the dynamical + spectroscopic masses (red dashed line) and to the entire sample (gray dashed line) are similar, indicating that Sb and earlier Hubble types form one class, with $\langle M_{\rm NSC}/M_{\star,\rm tot}\rangle \sim 0.002$, and Sc and later types form a different group, with $\langle M_{\rm NSC}/M_{\star,\rm tot}\rangle$ almost an order of magnitude smaller ($\sim$0.0003). The corresponding best-fit values of (T_1, T_2) are (3.51, 4.05) for the dynamical + spectroscopic masses and (3.10, 5.01) for the complete sample. As a crude check on whether this split is statistically significant, we performed Kolmogorov-Smirnov tests on the values of $M_{\rm NSC}/M_{\star,\rm tot}$ for galaxies with $T \leq 3$ and galaxies with $T \geq 5$. The K-S test gives a probability $P_{\rm KS} = 0.0038$

Table 2: Galaxies with well-determined NSC masses.

Name	T	D (Mpc)	$M_{\rm NSC}$ (err) ($\log_{10} M_\odot$)	Type	Source	Total $M_\star$ (err) ($\log_{10} M_\odot$)	Bulge $M_\star$ (err) ($\log_{10} M_\odot$)
Milky way	4	0.01	7.48 (0.09)	D	1	10.74 (0.09)	10.00 (0.13)
M31	3	0.77	7.54 (0.06)	D	2	10.92 (0.06)	10.27 (0.11)
M33	6	0.81	6.24 (0.08)	D	3	9.74 (0.08)	8.44 (0.12)
IC342	6	3.37	7.05 (0.07)	D	4	10.23 (0.07)	8.82 (0.23)
NGC300	7	2.02	5.98 (0.06)	D	5	9.26 (0.06)	< 6.26
NGC404	−3	3.18	7.04 (0.06)	D	6	9.24 (0.06)	9.06 (0.11)
NGC428	9	15.5	6.51 (0.18)	D	5	9.81 (0.18)	7.76 (0.29)
NGC1042	6	17.5	6.51 (0.18)	D	5	10.14 (0.18)	7.98 (0.21)
NGC1493	6	11.0	6.38 (0.19)	D	5	9.75 (0.19)	< 6.75
NGC1705	−3	5.11	4.91 (0.12)	D	7	8.38 (0.12)	7.87 (0.15)
NGC2139	6	22.9	5.92 (0.18)	D	5	10.15 (0.18)	< 7.15
NGC3423	6	14.4	6.53 (0.19)	D	5	10.08 (0.19)	8.91 (0.21)
NGC3621	7	6.64	7.00 (0.08)	D	8	10.20 (0.08)	8.25 (0.14)
NGC5457	6	7.05	6.91 (0.09)	D	9	10.68 (0.09)	8.76 (0.24)
NGC6946	6	5.89	7.88 (0.10)	D	9	10.62 (0.10)	9.36 (0.13)
NGC7418	6	17.8	7.78 (0.18)	D	5	10.29 (0.18)	8.50 (0.19)
NGC7424	6	10.5	6.09 (0.18)	D	5	9.53 (0.18)	8.29 (0.19)
NGC7793	7	3.91	6.96 (0.11)	D	5	9.65 (0.11)	8.01 (0.14)
NGC1325	4	19.6	7.08 (0.18)	S	10	10.34 (0.18)	7.91 (0.20)
NGC1385	6	18.1	6.39 (0.18)	S	10	10.28 (0.18)	8.69 (0.20)
NGC2552	9	9.68	5.77 (0.18)	S	10	9.00 (0.18)	6.30 (0.28)
NGC3177	3	19.6	8.15 (0.18)	S	10	10.10 (0.18)	8.71 (0.20)
NGC3277	2	21.4	8.34 (0.18)	S	10	10.32 (0.18)	...
NGC3455	3	16.4	6.75 (0.18)	S	10	9.46 (0.18)	...
NGC4030	4	20.5	7.99 (0.18)	S	10	10.92 (0.18)	10.28 (0.20)
NGC4411B	6	18.6	6.53 (0.19)	S	10	9.46 (0.19)	...
NGC4701	6	10.8	6.53 (0.18)	S	10	9.34 (0.18)	...
NGC4775	7	21.9	7.60 (0.19)	S	10	10.19 (0.19)	8.76 (0.20)
NGC5377	1	28.4	8.63 (0.18)	S	10	10.80 (0.18)	...
NGC5585	7	8.71	5.87 (0.19)	S	10	9.26 (0.19)	8.58 (0.21)
NGC5806	3	20.0	8.11 (0.18)	S	10	10.43 (0.18)	...
NGC7421	4	23.1	6.87 (0.18)	S	10	10.24 (0.18)	...
NGC7690	3	17.7	7.98 (0.18)	S	10	9.90 (0.18)	9.28 (0.20)

(1) Galaxy name. (2) Hubble type T from RC3. (3) Adopted distance in Mpc. (4) Logarithm of NSC mass and uncertainty; masses have been rescaled using the distances column 2, if necessary. Errors are 1-σ values. (5) Type of NSC mass measurement: D: dynamical, S: spectroscopic. (6) Source of NSC measurement: 1 = Launhardt et al. [31]; 2 = Kormendy and Bender [25]; 3 = Matthews et al. [26] + Gebhardt et al. [27]; 4 = Böker et al. [24]; 5 = Walcher et al. [21]; 6 = Seth et al. [29]; 7 = Ho and Filippenko [23]; 8 = Barth et al. [28]; 9 = Kormendy et al. [30]; 10 = Rossa et al. [11]. (7) Logarithm of total galaxy stellar mass and uncertainty (see text for details). (8) Logarithm of bulge stellar mass and uncertainty (or upper limit for bulgeless galaxies), based on decompositions in Erwin and Gadotti (2012a, in prep); galaxies currently missing proper decompositions are indicated by "...".

for the two sets of ratios coming from the same parent population if we use only the dynamical + spectroscopic masses, or $P_{\rm KS} = 3.1 \times 10^{-10}$ if we use the entire set of NSC masses.

Do NSCs in late-type spirals differ from those in early-type spirals, S0s, and ellipticals in any sense other than average mass? The available evidence is ambiguous. Böker [7] notes that NSC *sizes* appear to be independent of Hubble type. On the other hand, Rossa et al. [11] compared stellar populations of NSCs in early- and late-type spirals using fits to their spectroscopy and noted that the NSCs in late-type spirals did tend to have younger stellar populations and (slightly) lower metallicities. (They also argued against any observational effects that might produce systematic overestimates of NSC mass in early spirals.) This does at least suggest that different star-formation histories may lie behind the mass differences in NSCs.

We also plot the $M_{\rm BH}/M_{\star,\rm bul}$ ratio for SMBH host galaxies (thick gray dotted line in Figure 4, based on (1)). What this indicates is that the NSC—host-galaxy relationship for Sb and earlier types *is* consistent with the SMBH relation, *if all of the galaxy mass is in the bulge*. Since most of the galaxies used for the original CMO studies [8, 9] were dwarf and giant ellipticals (or S0 galaxies with high B/T ratios), it is easy to

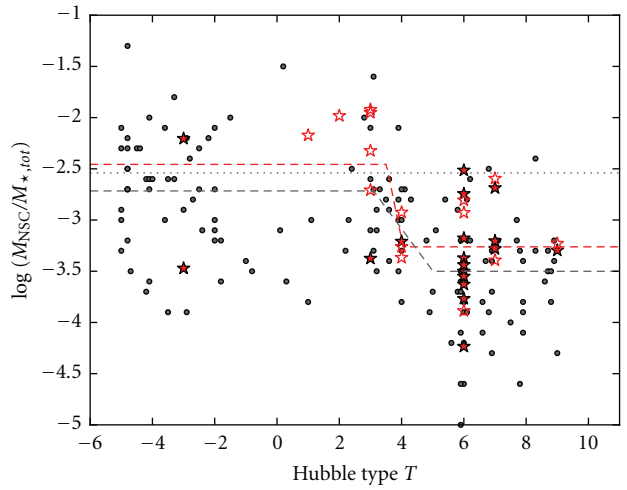


FIGURE 4: Relative masses of NSCs versus Hubble type of host galaxy, based on the compilation of Seth et al. [48]. Filled gray circles are NSC masses estimated from broadband colors or assumed M/L ratios by Seth et al [29]; red stars indicate NSC masses from spectroscopic (hollow) or dynamical (filled) measurements (see Table 2 for references). Also shown are simple fits to the dynamical+spectroscopic masses (dashed red line) and to the entire sample (dashed gray line), along with the mean mass ratio of SMBHs relative to their host *bulges* (dotted gray line).

see why the "NSC = SMBH" connection could be made. But this is clearly true only for very bulge-dominated systems.

Appendix

Data Tables

In Tables 1 and 2, we list the basic data parameters for SMBH and NSC hosts. References for the NSC masses are in the captions for Table 2. For the SMBH masses, the numbers in column 5 of Table 1 translate into the following references: 1 = Gillessen et al. [49]; 2 = Bender et al. [50]; 3 = Verolme et al. [51]; 4 = Krajnović et al. [52]; 5 = Gebhardt et al. [53]; 6 = Bower et al. [54]; 7 = Lodato and Bertin [55]; 8 = Atkinson et al. [56]; 9 = Rusli et al. [57]; 10 = Nowak et al. [58]; 11 = Houghton et al. [59]; 12 = Sarzi et al. [60]; 13 = Devereux et al. [61]; 14 = Davies et al. [62]; 15 = Barth et al. [63]; 16 = Nowak et al. [39]; 17 = Kondratko et al. [64]; 18 = Güeltekin et al. [65]; 19 = de Francesco et al. [66]; 20 = Hicks and Malkan [67]; 21 = Miyoshi et al. [68]; 22 = Ferrarese et al. [69]; 23 = Cretton and van den Bosch [70]; 24 = Walsh et al. [71]; 25 =Macchetto et al. [72]; 26 = Nowak et al. [73]; 27 = Shen and Gebhardt [74]; 28 = de Francesco et al. [75]; 29 = Neumayer et al. [76]; 30 = Capetti et al. [77]; 31 = Ferrarese and Ford [78]; 32 = van der Marel and van der Bosch [79]; 33 = Cappellari et al. [80]; 34 = Bontà et al [81].

Acknowledgments

The authors would like to thank Eva Noyola and Anil Seth for useful and interesting conversations, along with comments from two referees which improved the manuscript. They would also like to thank the organizers of the 2008 "Nuclear Star Clusters across the Hubble Sequence" workshop in Heidelberg and the 2010 "ESO Workshop on Central Massive Objects" in Garching for helping motivate and inspire this research. This work was supported by Priority Programme 1177 ("Witnesses of Cosmic History: Formation and evolution of black holes, galaxies and their environment") of the Deutsche Forschungsgemeinschaft. This work made use of data from the Ohio State University Bright Spiral Galaxy Survey, which was funded by Grants AST-9217716 and AST-9617006 from the United States National Science Foundation, with additional support from Ohio State University. It is also based on observations made with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is http://www.sdss.org/. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. Finally, this research made use of the Lyon-Meudon Extragalactic Database (LEDA; part of HyperLeda at http://leda.univ-lyon1.fr/), and the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

[1] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *The Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.

[2] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *The Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.

[3] A. Marconi and L. K. Hunt, "The relation between black hole mass, bulge mass, and near-Infrared luminosity," *The Astrophysical Journal*, vol. 589, no. 1, pp. L21–L24, 2003.

[4] N. Häring and H.-W. Rix, "On the black hole mass-bulge mass relation," *The Astrophysical Journal*, vol. 604, pp. L89–L92, 2004.

[5] C. M. Carollo, M. Stiavelli, P. T. de Zeeuw, and J. Mack, "Spiral galaxies with WFPC2. I. Nuclear morphology, bulges,

star clusters, and surface brightness profiles," *The Astronomical Journal*, vol. 114, no. 6, pp. 2366–2380, 1997.

[6] T. Böker, S. Laine, R. P. van der Marel et al., "A hubble space telescope census of nuclear star clusters in late-type spiral galaxies. I. Observations and image analysis," *The Astronomical Journal*, vol. 123, no. 3, pp. 1389–1410, 2002.

[7] T. Böker, "Properties of nuclear star clusters," *Journal of Physics Conference Series*, vol. 131, no. 1, Article ID 012043, 2008.

[8] E. H. Wehner and W. E. Harris, "From supermassive black holes to dwarf elliptical nuclei: a mass continuum," *The Astrophysics Journal*, vol. 644, pp. L17–L20, 2006.

[9] L. Ferrarese, P. Côté, E. Dalla Bontà et al., "A fundamental relation between compact stellar nuclei, supermassive black holes, and their host galaxies," *The Astrophysical Journal Letters*, vol. 644, no. 1, pp. L21–L24, 2006.

[10] P. Côté, S. Piatek, L. Ferrarese et al., "The ACS virgo cluster survey. VIII. The nuclei of early-type galaxies," *The Astrophysical Journal Supplement Series*, vol. 165, no. 1, pp. 57–94, 2006.

[11] J. Rossa, R. P. van der Marel, T. Böker et al., "Hubble space telescope stis spectra of nuclear star clusters in spiral galaxies: dependence of age and mass on hubble type," *The Astronomical Journal*, vol. 132, no. 3, pp. 1074–1099, 2006.

[12] M. Balcells, A. W. Graham, and R. F. Peletier, "Galactic bulges from hubble space telescope NICMOS observations: central galaxian objects, and nuclear profile slopes," *The Astrophysical Journal*, vol. 665, no. 2, pp. 1084–1103, 2007.

[13] D. E. McLaughlin, A. R. King, and S. Nayakshin, "The M-σ relation for nucleated galaxies," *The Astrophysical Journal*, vol. 650, no. 1, pp. L37–L40, 2006.

[14] Y. Li, Z. Haiman, and M.-M. Mac Low, "Correlations between central massive objects and their host galaxies: from bulgeless spirals to ellipticals," *The Astrophysical Journal*, vol. 663, no. 1 I, pp. 61–70, 2007.

[15] S. Nayakshin, M. I. Wilkinson, and A. King, "Competitive feedback in galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 398, pp. L54–L57, 2009.

[16] B. Devecchi, M. Volonteri, M. Colpi, and F. Haardt, "High-redshift formation and evolution of central massive objects - I. Model description," *Monthly Notices of the Royal Astronomical Society*, vol. 409, no. 3, pp. 1057–1067, 2010.

[17] J. Kormendy and K. Gebhardt, "Supermassive black holes in galactic nuclei," in *20th Texas Symposium on Relativistic Astrophysics*, J. C. Wheeler and H. Martel, Eds., vol. 586 of *American Institute of Physics Conference Series*, pp. 363–381, 2001.

[18] J. Kormendy, R. Bender, and M. E. Cornell, "Supermassive black holes do not correlate with galaxy disks or pseudobulges," *Nature*, vol. 469, pp. 374–376, 2011.

[19] C. M. Carollo, M. Stiavelli, and J. Mack, "Spiral galaxies with WFPC2. II. The nuclear properties of 40 objects," *The Astronomical Journal*, vol. 116, no. 1, pp. 68–84, 1998.

[20] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter," *The Astrophysical Journal*, vol. 698, no. 1, pp. 198–221, 2009.

[21] C. J. Walcher, R. P. van der Marel, D. Mclaughlin et al., "Masses of star clusters in the nuclei of bulgeless spiral galaxies," *The Astrophysical Journal*, vol. 618, no. 1, pp. 237–246, 2005.

[22] A. C. Seth, J. J. Dalcanton, P. W. Hodge, and V. P. Debattista, "Clues to nuclear star cluster formation from edge-on spirals," *The Astronomical Journal*, vol. 132, no. 6, pp. 2539–2555, 2006.

[23] L. C. Ho and A. V. Filippenko, "High-dispersion spectroscopy of a luminous, young star cluster in NGC 1705: further evidence for present-day formation of globular clusters," *The Astrophysical Journal*, vol. 472, no. 2, pp. 600–610, 1996.

[24] T. Böker, R. P. van der Marel, and W. D. Vacca, "CO band head spectroscopy of IC 342: mass and age of the nuclear star cluster 1," *The Astronomical Journal*, vol. 118, no. 2, pp. 831–842, 1999.

[25] J. Kormendy and R. Bender, "The double nucleus and central black hole of M31," *The Astrophysical Journal*, vol. 522, no. 2, pp. 772–792, 1999.

[26] L. D. Matthews, J. S. Gallagher III, J. E. Krist et al., "WFPC2 observations of compact star cluster nuclei in low-luminosity spiral galaxies," *The Astronomical Journal*, vol. 118, no. 1, pp. 208–235, 1999.

[27] K. Gebhardt, T. R. Lauer, J. Kormendy et al., "M33: a galaxy with no supermassive black hole," *The Astronomical Journal*, vol. 122, no. 5, pp. 2469–2476, 2001.

[28] A. J. Barth, L. E. Strigari, M. C. Bentz, J. E. Greene, and L. C. Ho, "Dynamical constraints on the masses of the nuclear star cluster and black hole in the late-type spiral galaxy NGC 3621," *The Astrophysical Journal*, vol. 690, no. 1, pp. 1031–1044, 2009.

[29] A. C. Seth, M. Cappellari, N. Neumayer et al., "The NGC 404 nucleus: star cluster and possible intermediate-mass black hole," *The Astrophysical Journal*, vol. 714, no. 1, pp. 713–731, 2010.

[30] J. Kormendy, N. Drory, R. Bender, and M. E. Cornell, "Bulgeless giant galaxies challenge our picture of galaxy formation by hierarchical clustering," *The Astrophysical Journal*, vol. 723, no. 1, pp. 54–80, 2010.

[31] R. Launhardt, R. Zylka, and P. G. Mezger, "The nuclear bulge of the galaxy III. Large-scale physical characteristics of stars and interstellar matter," *Astronomy and Astrophysics*, vol. 384, no. 1, pp. 112–139, 2002.

[32] T. H. Jarrett, T. Chester, R. Cutri, S. Schneider, M. Skrutskie, and J. P. Huchra, "2MASS extended source catalog: overview and algorithms," *The Astronomical Journal*, vol. 119, no. 5, pp. 2498–2531, 2000.

[33] S. Malhotra, D. N. Spergel, J. E. Rhoads, and L. I. Jing, "The milky way, local galaxies, and the infrared tully-fisher relation," *The Astrophysical Journal*, vol. 473, no. 2, pp. 687–691, 1996.

[34] E. F. Bell, D. H. McIntosh, N. Katz, and M. D. Weinberg, "The optical and near-infrared properties of galaxies. I. Luminosity and stellar MASS functions," *The Astrophysical Journal Letters*, vol. 149, no. 2, pp. 289–312, 2003.

[35] D. G. York, J. Adelman, J. E. Anderson Jr. et al., "The sloan digital sky survey: technical summary," *The Astronomical Journal*, vol. 120, no. 3, pp. 1579–1587, 2000.

[36] R. E. de Souza, D. A. Gadotti, and S. dos Anjos, "BUDDA: a new two-dimensional Bulge/Disk decomposition code for detailed structural analysis of galaxies," *The Astrophysical Journal Supplement Series*, vol. 153, no. 2, pp. 411–427, 2004.

[37] D. A. Gadotti, "Image decomposition of barred galaxies and AGN hosts," *Monthly Notices of the Royal Astronomical Society*, vol. 384, no. 1, pp. 420–439, 2008.

[38] J. Hu, "The black hole mass-stellar velocity dispersion correlation: bulges versus pseudo-bulges," *Monthly Notices of the Royal Astronomical Society*, vol. 386, no. 4, pp. 2242–2252, 2008.

[39] N. Nowak, J. Thomas, P. Erwin, R. P. Saglia, R. Bender, and R. I. Davies, "Do black hole masses scale with classical bulge luminosities only? The case of the two composite pseudo-bulge galaxies NGC 3368 and NGC 3489," *Monthly Notices of the Royal Astronomical Society*, vol. 403, no. 2, pp. 646–672, 2010.

[40] W. Dehnen and J. Binney, "Mass models of the Milky Way," *Monthly Notices of the Royal Astronomical Society*, vol. 294, no. 3, pp. 429–438, 1998.

[41] A. Klypin, H. Zhao, and R. S. Somerville, "ΛCDM-based models for the Milky Way and M31. I. Dynamical models," *The Astrophysical Journal*, vol. 573, no. 2, pp. 597–613, 2002.

[42] C. Flynn, J. Holmberg, L. Portinari, B. Fuchs, and H. Jahreiß, "On the mass-to-light ratio of the local Galactic disc and the optical luminosity of the Galaxy," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 3, pp. 1149–1160, 2006.

[43] P. B. Eskridge, J. A. Frogel, R. W. Pogge et al., "Near-infrared and optical morphology of spiral galaxies," *The Astrophysical Journal Supplement Series*, vol. 143, no. 1, pp. 73–111, 2002.

[44] G. D'Agostini, "Fits, and especially linear fits, with errors on both axes, extra variance of the data points and other complications," *Physics*, 2005.

[45] C. Guidorzi, F. Frontera, E. Montanari et al., "The slope of the gamma-ray burst variability/peak luminosity correlation," *Monthly Notices of the Royal Astronomical Society*, vol. 371, pp. 843–851, 2006.

[46] O. Y. Gnedin, D. H. Weinberg, J. Pizagno, F. Prada, and H.-W. Rix, "Dark matter halos of disk galaxies: constraints from the Tully-Fisher relation," *The Astrophysical Journal*, vol. 671, no. 2, pp. 1115–1134, 2007.

[47] J. Bland-Hawthorn, M. Vlajić, K. C. Freeman, and B. T. Draine, "NGC 300: an extremely faint, outer stellar disk observed to 10 scale lengths," *The Astrophysical Journal*, vol. 629, no. 1, pp. 239–249, 2005.

[48] A. Seth, M. Agüeros, D. Lee, and A. Basu-Zych, "The coincidence of nuclear star clusters and active galactic nuclei," *The Astrophysical Journal*, vol. 678, no. 1, pp. 116–130, 2008.

[49] S. Gillessen, F. Eisenhauer, S. Trippe et al., "Monitoring stellar orbits around the massive black hole in the galactic center," *The Astrophysical Journal*, vol. 692, pp. 1075–1109, 2009.

[50] R. Bender, J. Kormendy, G. Bower et al., "HST STIS spectroscopy of the triple nucleus of M31: two nested disks in keplerian rotation around a supermassive black hole," *The Astrophysical Journal*, vol. 631, no. 1, pp. 280–300, 2005.

[51] E. K. Verolme, M. Cappellari, Y. Copin et al., "A SAURON study of M32: measuring the intrinsic flattening and the central black hole mass," *Monthly Notices of the Royal Astronomical Society*, vol. 335, pp. 517–525, 2002.

[52] D. Krajnović, R. M. McDermid, M. Cappellari, and R. L. Davies, "Determination of masses of the central black holes in NGC 524 and 2549 using laser guide star adaptive optics," *Monthly Notices of the Royal Astronomical Society*, vol. 399, no. 4, pp. 1839–1857, 2009.

[53] K. Gebhardt, D. Richstone, S. Tremaine et al., "Axisymmetric dynamical models of the central regions of galaxies," *The Astrophysical Journal*, vol. 583, no. 1, pp. 92–115, 2003.

[54] G. A. Bower, R. F. Green, R. Bender et al., "Evidence of a supermassive black hole in the galaxy NGC 1023 from the nuclear stellar dynamics," *The Astrophysical Journal*, vol. 550, no. 1, pp. 75–86, 2001.

[55] G. Lodato and G. Bertin, "Non-Keplerian rotation in the nucleus of NGC 1068: evidence for a massive accretion disk?" *Astronomy and Astrophysics*, vol. 398, no. 2, pp. 517–524, 2003.

[56] J. W. Atkinson, J. L. Collett, A. Marconi et al., "Supermassive black hole mass measurements for NGC 1300 and 2748 based on hubble space telescope emission-line gas kinematics," *Monthly Notices of the Royal Astronomical Society*, vol. 359, pp. 504–520, 2005.

[57] S. P. Rusli, J. Thomas, P. Erwin, R. P. Saglia, N. Nowak, and R. Bender, "The central black hole mass of the high-σ but low-bulge-luminosity lenticular galaxy NGC 1332," *Monthly Notices of the Royal Astronomical Society*, vol. 410, no. 2, pp. 1223–1236, 2011.

[58] N. Nowak, R. P. Saglia, J. Thomas, R. Bender, R. I. Davies, and K. Gebhardt, "The supermassive black hole of FornaxA," *Monthly Notices of the Royal Astronomical Society*, vol. 391, pp. 1629–1649, 2008.

[59] R. C. W. Houghton, J. Magorrian, M. Sarzi, N. Thatte, R. L. Davies, and D. Krajnović, "The central kinematics of NGC 1399 measured with 14 pc resolution," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 1, pp. 2–18, 2006.

[60] M. Sarzi, H.-W. Rix, J. C. Shields et al., "Supermassive black holes in bulges," *The Astrophysical Journal*, vol. 550, no. 1, pp. 65–74, 2001.

[61] N. Devereux, H. Ford, Z. Tsvetanov, and G. Jacoby, "STIS spectroscopy of the central 10 parsecs of M81: evidence for a massive black hole," *The Astronomical Journal*, vol. 125, no. 3, pp. 1226–1235, 2003.

[62] R. I. Davies, J. Thomas, R. Genzel et al., "The star-forming torus and stellar dynamical black hole mass in the Seyfert 1 nucleus of NGC 3227," *The Astrophysical Journal*, vol. 646, no. 2, pp. 754–773, 2006.

[63] A. J. Barth, M. Sarzi, H.-W. Rix, L. C. Ho, A. V. Filippenko, and W. L. W. Sargent, "Evidence for a supermassive black hole in the S0 galaxy NGC 3245," *The Astrophysical Journal*, vol. 555, no. 2, pp. 685–708, 2001.

[64] P. T. Kondratko, L. J. Greenhill, and J. M. Moran, "The parsec-scale accretion disk in NGC 3393," *The Astrophysical Journal*, vol. 678, no. 1, pp. 87–95, 2008.

[65] K. Gültekin, D. O. Richstone, K. Gebhardt et al., "A quintet of black hole mass determinations," *The Astrophysical Journal*, vol. 695, pp. 1577–1590, 2009.

[66] G. de Francesco, A. Capetti, and A. Marconi, "Measuring supermassive black holes with gas kinematics: the active so galaxy NGC 3998," *Astronomy and Astrophysics*, vol. 460, no. 2, pp. 439–448, 2006.

[67] E. K. S. Hicks and M. A. Malkan, "Circumnuclear gas in seyfert 1 galaxies: morphology, kinematics, and direct measurement of black hole masses," *The Astrophysical Journal Supplement Series*, vol. 174, no. 1, pp. 31–73, 2008.

[68] M. Miyoshi, J. Moran, J. Herrnstein et al., "Evidence for a black hole from high rotation velocities in a sub-parsec region of NGC4258," *Nature*, vol. 373, no. 6510, pp. 127–129, 1995.

[69] L. Ferrarese, H. C. Ford, and W. Jaffe, "Evidence for a massive black hole in the active galaxy NGC 4261 from bubble space telescope images and spectra," *The Astrophysical Journal*, vol. 470, no. 1, pp. 444–459, 1996.

[70] N. Cretton and F. C. van den Bosch, "Evidence for a massive black hole in the S0 galaxy NGC 4342," *The Astrophysical Journal*, vol. 514, no. 2, pp. 704–724, 1999.

[71] J. L. Walsh, A. J. Barth, and M. Sarzi, "The supermassive black hole in M84 revisited," *The Astrophysical Journal*, vol. 721, no. 1, pp. 762–776, 2010.

[72] F. Macchetto, A. Marconi, D. J. Axon, A. Capetti, W. Sparks, and P. Crane, "The supermassive black hole of M87 and the kinematics of its associated gaseous disk," *The Astrophysical Journal*, vol. 489, no. 2, pp. 579–600, 1997.

[73] N. Nowak, R. P. Saglia, J. Thomas et al., "The supermassive black hole in NGC 4486a detected with SINFONI at the very large telescope," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 3, pp. 909–914, 2007.

[74] J. Shen and K. Gebhardt, "The supermassive black hole and dark matter halo of NGC 4649 (M60)," *The Astrophysical Journal*, vol. 711, pp. 484–494, 2010.

[75] G. de Francesco, A. Capetti, and A. Marconi, "Measuring supermassive black holes with gas kinematics. II. The LINERs IC 989, NGC 5077, and NGC 6500," *Astronomy & Astrophysics*, vol. 479, pp. 355–363, 2008.

[76] N. Neumayer, M. Cappellari, J. Reunanen et al., "The central parsecs of centaurus A: high-excitation gas, a molecular disk, and the mass of the black hole," *The Astrophysical Journal*, vol. 671, no. 2, pp. 1329–1344, 2007.

[77] A. Capetti, A. Marconi, D. Macchetto, and D. Axon, "The supermassive black hole in the Seyfert 2 galaxy NGC 5252," *Astronomy and Astrophysics*, vol. 431, no. 2, pp. 465–475, 2005.

[78] L. Ferrarese and H. C. Ford, "Nuclear disks of gas and dust in early-type galaxies and the hunt for massive black holes: hubble Space Telescope observations of NGC 6251," *The Astrophysical Journal*, vol. 515, no. 2, pp. 583–602, 1999.

[79] R. P. van der Marel and F. C. van den Bosch, "Evidence for a 3 $\times$ 108 M$\odot$ black hole in NGC 7052 from hubble space telescope observations of the nuclear gas disk," *The Astronomical Journal*, vol. 116, no. 5, pp. 2220–2236, 1998.

[80] M. Cappellari, E. K. Verolme, R. P. van der Marel et al., "The counterrotating core and the black hole mass of IC 1459," *The Astrophysical Journal*, vol. 578, no. 2, pp. 787–805, 2002.

[81] E. Dalla Bontà, L. Ferrarese, E. M. Corsini et al., "The high-mass end of the black hole mass function: mass estimates in brightest cluster galaxies," *The Astrophysical Journal*, vol. 690, no. 1, pp. 537–559, 2009.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 138380, 15 pages
doi:10.1155/2012/138380

Research Article

AGN Triggering in the Infall Regions of Distant X-Ray Luminous Galaxy Clusters at $0.9 < z \lesssim 1.6$

R. Fassbender,[1] R. Šuhada,[2] and A. Nastasi[1]

[1] *Max-Planck-Institut für Extraterrestrische Physik (MPE), Postfach 1312, Giessenbachstraße, 85741 Garching, Germany*
[2] *Department of Physics, Ludwigs-Maximilians University Munich, Scheinerstraße 1, 81679 Munich, Germany*

Correspondence should be addressed to R. Fassbender, rfassben@mpe.mpg.de

Received 14 November 2011; Accepted 20 March 2012

Academic Editor: Angela Bongiorno

Observational constraints on the average radial distribution profile of AGN in distant galaxy clusters can provide important clues on the triggering mechanisms of AGN activity in dense environments and are essential for a completeness evaluation of cluster selection techniques in the X-ray and mm wavebands. The aim of this work is a statistical study with XMM-*Newton* of the presence and distribution of X-ray AGN in the large-scale structure environments of 22 X-ray luminous galaxy clusters in the redshift range $0.9 < z \lesssim 1.6$ compiled by the XMM-*Newton* Distant Cluster Project (XDCP). To this end, the X-ray point source lists from detections in the soft band (0.35–2.4 keV) and full band (0.3–7.5 keV) were stacked in cluster-centric coordinates and compared to average background number counts extracted from three independent control fields in the same observations. A significant full-band (soft-band) excess of ∼78 (67) X-ray point sources is found in the cluster fields within an angular distance of 8′ (4 Mpc) at a statistical confidence level of 4.0σ (4.2σ), corresponding to an average number of detected excess AGN per cluster environment of 3.5 ± 0.9 (3.0 ± 0.7). The data point towards a rising radial profile in the cluster region ($r < 1$ Mpc) of predominantly low-luminosity AGN with an average detected excess of about one point source per system, with a tentative preferred occurrence along the main cluster elongation axis. A second statistically significant overdensity of brighter soft-band-detected AGN is found at cluster-centric distances of 4′–6′ (2-3 Mpc), corresponding to about three times the average cluster radius R_{200} of the systems. If confirmed, these results would support the idea of two different physical triggering mechanisms of X-ray AGN activity in dependence of the radially changing large-scale structure environment of the distant clusters. For high-z cluster studies at lower spatial resolution with the upcoming eROSITA all-sky X-ray survey, the results suggest that cluster-associated X-ray AGN may impose a bias in the spectral analysis of high-z systems, while their detection and flux measurements in the soft band may not be significantly affected.

1. Introduction

Observational studies of the connection of Active Galactic Nuclei (AGN) with the large-scale environment of massive clusters of galaxies and their mutual cosmic evolution can provide important insights into the physical conditions necessary to trigger or suppress AGN activity in galaxies. In this respect, AGN activity can be charted in dependence of the changing environments of galaxy clusters as a function of cluster-centric distance: from the dense cores, to the cluster outskirts, and further out to the matter infall regions and the surrounding cosmic web. Furthermore, the evolution of the occurrence of X-ray and radio AGN in cluster environments as a function of redshift is of key importance for the characterization and completeness evaluation of ongoing and future high-z cluster surveys in the X-ray band and via the Sunyaev-Zeldovich effect (SZE) at mm wavelengths.

Numerous *Chandra* sample studies on the X-ray AGN content of galaxy clusters up to redshifts of about unity, for example, (e.g., [1–6]), have firmly established that the AGN fraction in clusters is significantly rising as a function of redshift. On the other hand, the cluster environment appears to suppress the occurrence of X-ray AGN activity in massive galaxies compared to a field galaxy sample at all probed redshifts so far, for example, (e.g., [3, 5, 7]), and the distribution of X-ray AGN is significantly less concentrated

in terms of cluster-centric distance in comparison to radio AGN, (e.g., [1, 8, 9]). However, owing to the small number of very distant test clusters at $z > 0.9$ available to these studies, the data basis at redshifts beyond unity is still very sparse, and at $z > 1.3$ X-ray cluster studies are limited to a single system [10, 11]. So far, the only cluster AGN study with a sizable sample of systems at the epoch $1 < z \lesssim 1.5$ was presented in [12] based on infrared selected clusters from the IRAC Shallow Cluster Survey [13], which found a continuing trend of increasing AGN fractions with redshifts in systems with an average halo mass of $\sim 10^{14}\,M_{\odot}$.

The aim of this paper is to extend the accessible redshift regime for a sample of X-ray selected clusters to $0.9 < z \lesssim 1.6$ in order to perform a statistical study of the X-ray point source excess and its radial dependence in distant X-ray luminous systems. In contrast to the previous studies based on targeted follow-up observations with *Chandra*, this work is built upon archival XMM-*Newton* observations, in which the clusters have been serendipitously detected. This distant cluster sample and the performed X-ray point source stacking analysis are introduced in Section 2. The results are presented in Section 3, followed by the discussion in Section 4, and conclusions in Section 5. A standard ΛCDM cosmological model with parameters $(H_0, \Omega_{\rm m}, \Omega_{\rm DE}, w) = (70\,{\rm km\,s^{-1}Mpc^{-1}}, 0.3, 0.7, -1)$ is assumed throughout this paper, which yields a median projected angular scale of 8.2 kpc/″ (corresponding to 2.0′/Mpc) with <5% variation in the probed redshift interval $z = 0.9$–1.6.

2. Distant Cluster Sample and X-Ray Stacking Analysis

2.1. Cluster Sample. This work uses the largest published sample of distant X-ray luminous galaxy clusters at redshifts $z > 0.9$ to date as presented in Fassbender et al. [14]. The sample comprises 22 X-ray selected clusters in the redshift range $0.9 < z \lesssim 1.6$ with a median system mass of $M_{200} \simeq 2\times10^{14}\,M_{\odot}$. All clusters are part of the XMM-*Newton* Distant Cluster Project (XDCP, e.g., [14–16]), a serendipitous X-ray survey focussing on the detection and study of galaxy clusters in the first half of cosmic time. Table 1 provides an overview of the considered cluster sample, including the XMM-*Newton* observations used for this study, their effective clean exposure time (ECT) after flare removal, and original relevant source publications for the individual clusters. (We define the ECT as the period during which all three instruments in imaging operation would collect the equivalent number of soft science photons for the particular observation.)

All clusters in this sample have by construction medium to deep XMM-*Newton* observations available, whereas the *Chandra* archive coverage of this newly constructed sample is currently less than 30%. The complete coverage and high sensitivity is the main advantage to perform a first X-ray point source study around distant X-ray clusters with XMM-*Newton*. A second advantage is the larger 30′-diameter field of view (FoV), which allows direct measurements of background counts in the same observation. XMM-*Newton*'s spatial resolution of 5–15″ (FWHM), on the other hand, is significantly lower than for *Chandra*, which implies that the central cluster core regions of up to $\sim 15'' \simeq 120$ kpc cannot be properly probed for point sources in addition to the underlying detected extended X-ray emission of the cluster.

In any case, the statistical detection of a point source excess around distant clusters at $z > 0.9$ is a challenging task since the average background density of X-ray point sources in the considered XMM-*Newton* fields is more than 20 within a 6′-radius (∼3 Mpc) compared to an expected cluster excess of a few sources (∼1.5 for $z < 0.9$ systems [4]). On the other hand, any measurable point source excess associated with $z > 0.9$ clusters can be directly attributed to AGN activity with X-ray luminosities of $L_X > 10^{43}\,{\rm erg\,s^{-1}}$ owing to the average soft-band point source detection limit in the XMM-*Newton* fields of $\gtrsim 10^{-15}\,{\rm erg\,s^{-1}\,cm^{-2}}$.

2.2. X-Ray Data. The XMM-*Newton* observations used for this work are in general the original discovery fields of the clusters as listed in Table 1 (column 6), with the exception of C08 for which the deeper follow-up observation was used. The clean effective exposure time of these fields after flare removal (column 7) ranges from 8.5 to 82 ksec with a median integration time of 19 ksec. Three of the fields contain 2-3 confirmed distant clusters inside their FoV, which are treated here as independent observations for each cluster source.

Most of the identified high-z clusters are located at off-axis angles $\Theta_{\rm off}$ between 5 and 12 arcmin from the optical axis, which was defined as the aimpoint of the PN (the most sensitive instrument) labelled as center in Figure 1. The 30′-diameter FoV of XMM-*Newton* allows to define three quasi independent background control regions inside the same XMM-*Newton* observation at the same off-axis angle with centers rotated by 90, 180, and 270 degrees about the optical axis as is shown in Figure 1. These fixed rotation angles to the control field centers ensure maximally spaced distances at a given off-axis angle between control and clusters fields in order to maximize the non overlapping regions between them. When comparing the cluster environment and background field point source counts, this approach (to first order) automatically accounts for (i) changes in the effective exposure time across the FoV due to vignetting, (ii) the changing PSF as a function of off-axis angle, and (iii) possible incomplete coverage beyond some cluster-centric distance in the outer radial direction due to the edge of the instrumental FoV. For the observations in which the cluster location is at small off-axis angles of $\Theta_{\rm off} < 8'$ the control field locations were shifted back to $\Theta_{\rm off} = 8'$ at their respective rotation angles of 90, 180, and 270 degrees. This ensures that cluster and background fields are not overlapping within their minimum separations of >8′, while only slightly changing the average effective sensitivity in the control fields relative to the cluster environment.

The X-ray source detection was performed as part of the XDCP distant cluster survey as detailed in Fassbender et al [14]. Here the focus of the analysis lies on the detected point sources in the individual cluster fields, which are displayed by green circles in Figures 1 and 2. The X-ray source lists

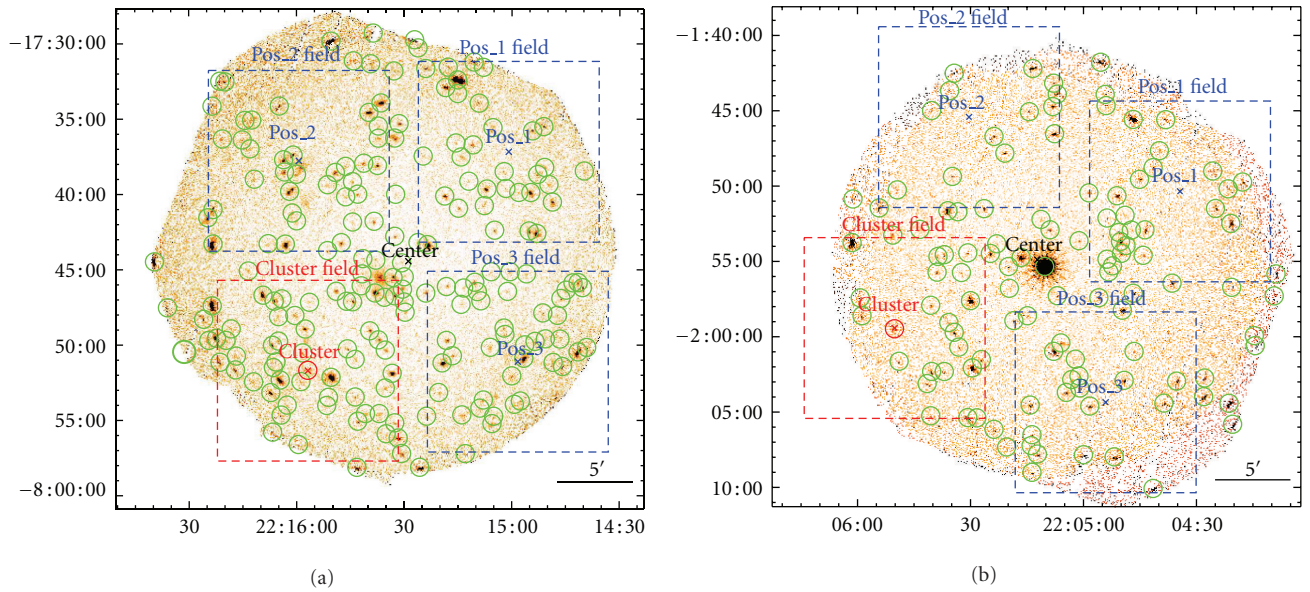


Figure 1: Examples of the extraction and stacking process for the 82.2 ksec field of cluster XDCP J2215.9 − 1751 (C09) in the left panel and the 24.9 ksec observation of the system XDCP J2205.8 − 0159 (C12) in the right panel. The cluster positions and the $12' \times 12'$ region around them are indicated in red. The three control fields at equivalent off-axis angles from the center position (black) are marked in blue. Green circles mark the positions of all detected soft-band X-ray point sources in the fields.

Table 1: List of the 22 distant galaxy clusters at $z > 0.9$ from Fassbender et al. [14] used for the stacking analysis in this work. The table lists a cluster identification number (column 1), the system redshift (2), the official cluster name (3), X-ray centroid coordinates (4 + 5), the observation identification number of the XMM-*Newton* field used (6), the corresponding effective clean time (ECT) of the field in (7), and relevant literature references to the cluster in (8).

ID	z	Official name	RA J2000	DEC J2000	OBSID	ECT ksec	References
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C01	1.579	XDCP J0044.0 − 2033	00 : 44 : 05.2	−20 : 33 : 59.7	0042340201	8.5	[26]
C02	1.555	XDCP J1007.3 + 1237	10 : 07 : 21.6	+12 : 37 : 54.3	0140550601	19.4	[27]
C03	1.490	XDCP J0338.8 + 0021	03 : 38 : 49.5	+00 : 21 : 08.1	0036540101	18.0	[28]
C04	1.457	XMMXCS J2215.9 − 1738	22 : 15 : 58.5	−17 : 38 : 05.8	0106660101	51.7	[10, 29–31]
C05	1.396	XDCP J2235.3 − 2557	22 : 35 : 20.4	−25 : 57 : 43.2	0111790101	13.6	[21, 32]
C06	1.358	XDCP J1532.2 − 0837	15 : 32 : 13.2	−08 : 37 : 01.4	0100240801	22.4	[33]
C07	1.335	SpARCS J0035.8 − 4312	00 : 35 : 50.1	−43 : 12 : 10.3	0148960101	47.2	[14, 34]
C08	1.237	RDCS J1252.9 − 2927	12 : 52 : 54.5	−29 : 27 : 18.0	0057740401	62.0	[22]
C09	1.227	XDCP J2215.9 − 1751	22 : 15 : 56.9	−17 : 51 : 40.9	0106660601	82.2	[35]
C10	1.185	XDCP J0302.1 − 0001	03 : 02 : 11.9	−00 : 01 : 34.3	0041170101	40.9	[33]
C11	1.122	XDCP J2217.3 + 1417	22 : 17 : 20.8	+14 : 17 : 54.6	0103660301	10.3	[14]
C12	1.117	XDCP J2205.8 − 0159	22 : 05 : 50.3	−01 : 59 : 27.4	0012440301	24.9	[14, 36]
C13	1.097	XDCP J0338.7 + 0030	03 : 38 : 44.2	+00 : 30 : 01.8	0036540101	18.0	[37]
C14	1.082	XDCP J1007.8 + 1258	10 : 07 : 50.5	+12 : 58 : 18.1	0140550601	19.4	[38]
C15	1.053	XLSS J0227.1 − 0418	02 : 27 : 09.2	−04 : 18 : 00.9	0112680101	22.7	[39, 40]
C16	1.050	XLSS J0224.0 − 0413	02 : 24 : 04.1	−04 : 13 : 31.7	0112680301	19.2	[23, 39, 40]
C17	1.000	XDCP J2215.9 − 1740	22 : 15 : 57.5	−17 : 40 : 25.6	0106660101	51.7	[35]
C18	0.975	XDCP J1229.4 + 0151	12 : 29 : 29.2	+01 : 51 : 31.6	0126700201	8.7	[41]
C19	0.975	XDCP J1230.2 + 1339	12 : 30 : 16.9	+13 : 39 : 04.3	0112552101	10.3	[18]
C20	0.959	XDCP J0027.2 + 1714	00 : 27 : 14.3	+17 : 14 : 36.3	0050140201	41.8	[14]
C21	0.947	XDCP J0104.3 − 0630	01 : 04 : 22.3	−06 : 30 : 03.1	0112650401	18.4	[42]
C22	0.916	XDCP J0338.5 + 0029	03 : 38 : 30.5	+00 : 29 : 20.2	0036540101	18.0	[14]

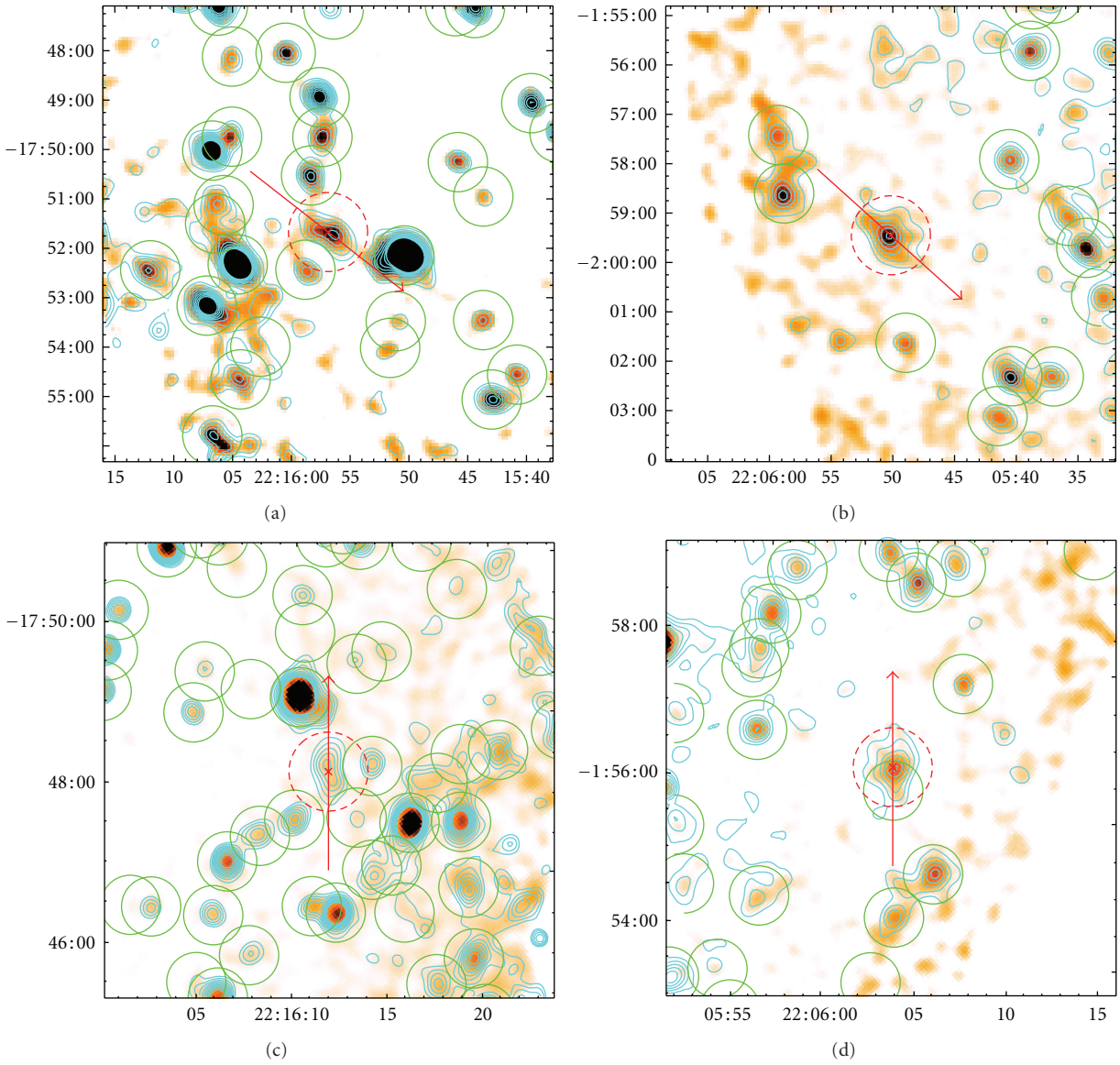


Figure 2: Close-up examples of the stacking process for the $9.2' \times 9.2'$ cluster environments of XDCP J2215.9 − 1751 (C09) in the left panels and XDCP J2205.8 − 0159 (C12) in the right panels. Cluster positions are marked by red crosses and dashed circles, and the main elongation axis of the ICM emission is indicated by the red arrow with $2'$ length in both directions from the cluster centers. The top panels show the smoothed soft-band image and the 0.35–2.4 keV detected sources (green circles) in their original orientation (North up, East to the left). The bottom panels display the rotated images and detected sources in the full 0.3–7.5 keV energy range with rotation angles that align the cluster elongation axis in the North-South direction. All panels show overlaid the soft-band X-ray surface brightness contours in cyan.

were produced with the XMM Science Analysis Software (http://xmm.esac.esa.int/sas/) (SAS) tasks `eboxdetect` for a first sliding box detection of candidate sources followed by a maximum likelihood fitting procedure with `emldetect` for the final source parameter determination. For this work, the source lists of two different detection procedures are used: (1) a *soft-band* source detection in the energy band 0.35–2.4 keV down to point source significances of about 3σ and (2) a *full-band* detection in multiple bands covering the energy range 0.3–7.5 keV down to point source significances of about 2σ. Although originally developed and optimized for distant cluster detections, the two schemes are able to distinguish and probe different aspects of the AGN-cluster connection. The investigation in the *soft band* is mostly sensitive to low-absorption type-I AGN and can probe the effects of cluster-AGN on the detection efficiency of distant clusters at lower spatial resolution with the upcoming all-sky survey eROSITA [17] as discussed in Section 4.5. The *full-band* detection, on the other hand, is sensitive to the full type-I and type-II AGN population and can hence provide a more complete census of AGN activity in the vicinity of distant clusters.

In order to allow a robust determination of the pure point source excess, all 105 detected *extended* X-ray sources in the considered fields were removed from the master source

lists, including all of the 22 distant cluster targets listed in Table 1. Additionally, the *soft-* and *full-band* master lists were manually cleaned from obvious spurious detections ($\simeq$5.5%) in the immediate vicinity of very bright X-ray sources in the FoV, such as the central source displayed in the right panel of Figure 1. The final combined clean X-ray point source lists contain 2770 objects for the *soft-band* and 4228 sources for the *full-band* detection. With an effective total X-ray coverage of 4.246 deg^2 for the considered fields, the average point source surface densities amount to 0.182 srcs/arcmin^2 in the *soft band* and 0.277 srcs/arcmin^2 in the *full band*.

2.3. Point Source Stacking Analysis. In order to keep the analysis simple, the focus is placed on a possible *detectable* AGN point source excess in cluster environments with respect to control fields, irrespective of the flux and luminosity distribution of such sources. Owing to the varying effective exposure time of the different XMM-*Newton* fields in Table 1 (columns 6 + 7), the point source detection sensitivities scale with roughly the inverse square root of the exposure time and vary across the FoV as a function of off-axis angle due to increased vignetting effects at larger $\Theta_{\rm off}$. The typical 0.5–2.0 keV soft-band point source sensitivities of $\simeq$(1-2) $\times$ 10^{-15} erg s^{-1} cm^{-2} ensure, however, that any detected point source excess associated with clusters in the redshift regime at $0.9 < z \lesssim 1.6$ can be safely attributed to X-ray AGN (under the well-justified assumption that a statistically significant, background-subtracted population of excess counts is physically associated with the LSS cluster environments at the respective redshifts), rather than, for example, star-forming galaxies at lower X-ray luminosities.

For the point source stacking analysis the following approach is adapted. (a) For each of the 22 distant clusters in the considered sample a subimage is extracted from the corresponding XMM-*Newton* observation that places the cluster at the central image coordinates ($X_{\rm cen,cl}$, $Y_{\rm cen,cl}$) as displayed in red in Figure 1. The same procedure is repeated at the three control field positions at rotation angles of 90, 180, and 270 degrees about the optical axis shown in blue in the same figure. (b) In order to ensure a homogeneous effective exposure time distribution in the final stacked cluster and control fields, the identical sub-images at the four positions as in the first step are extracted from the associated exposure maps. (c) The X-ray source lists for the *soft-band* and *full-band* detections are loaded to each of the four associated extracted subframes (one cluster plus three control positions) in the world coordinate system (WCS) and resaved in the image coordinate system of each frame. (d) For each of the four positions per cluster, the 22 extracted sub-images and exposure maps around the cluster position and control fields are coadded to result in the final deep image and exposure map stacks for each position. (e) Similarly, the 22 source lists in image coordinates at each position are concatenated into single-master-source files at each cluster and control field position for the *soft-band* and *full-band* detections separately. The source distribution of these stacked master files can now be further analyzed as a function of distance to the central extraction position with image coordinates ($X_{\rm cen,cl}$, $Y_{\rm cen,cl}$).

The extraction procedure with the cluster (or control field) position in the center of the sub-image is depicted in the top panels of Figure 2 for the two systems C09 (left) and C12 (right). The detected extended X-ray emission of the cluster (red dashed circles) does by definition not contribute to the point source statistics (green circles). In order to also investigate potentially preferred directions of the AGN excess in cluster environments, a second variant of the previously described stacking analysis is performed. This time the visually determined main elongation axis of the detected cluster emission (red arrows in Figure 2) is assumed to be a proxy for the main matter infall and assembly axis of the cluster (e.g., see [18]). All extracted sub-images and source lists of steps (a–c) are then rotated about the central position with coordinates ($X_{\rm cen,cl}$, $Y_{\rm cen,cl}$) in a way to align the cluster elongation axis for each of the 22 cluster fields in the Northsouth direction as shown in the lower panels of Figure 2. Subsequently, step (e) is repeated with these rotated source lists to produce a stacked master catalog of X-ray sources in a coordinate system where the cluster elongation axes are coaligned in the vertical direction.

A robust determination of the background number counts in the three control fields is of critical importance for the results of the stacking analysis. As discussed, some control field positions had to be placed at slightly larger off-axis angles compared to the cluster in order to avoid a significant overlap of the source lists at relatively small off-center distances (i.e., <8′). Such a shift was applied to three of the 22 fields with up to 2′, to four fields with 2′–4′, and to two fields with >4′, which inevitably results in slight differences of the stacked effective clean exposure time in the cluster and control fields. Figure 3 shows these stacked exposure time profiles (left panel), from which a median fractional exposure difference for the cluster field of $dt_{\rm exp}/t_{\rm exp} = 0.113 = 11.3\%$ is determined as displayed by the blue dashed line in the right panel. As an important cross-check, it can be confirmed that no significant systematic radial trend is present in the control field exposure time ratios (black line) with respect to the median.

For the deep, background-limited XMM-*Newton* observations this translates into an average fractional difference of the flux limit of $df_{\rm lim}/f_{\rm lim} \simeq -(dt_{\rm exp}/t_{\rm exp})^{0.5} \simeq -3.4\%$. With the knowledge of the log$N$-log$S$ distribution from point source number counts in deep fields, the effect of the slightly increased sensitivity in the cluster field can be quantified. Using the measured faint end slope of the log$N$-log$S$ distribution of $\alpha = (1.65 \pm 0.05)$ [19], the fractional effect on the number counts can be determined as $dN/N \simeq -(\alpha - 1)(dS/S) \simeq 0.022 \simeq 2.2\%$. For a fair evaluation of excess counts in the cluster environment, the control field counts have thus to be scaled up by the small correction factor $b_{\rm cor} \simeq 1.022$ that accounts for the slight average sensitivity differences.

The sensitivity-corrected radial distribution of the background counts in the three control fields is displayed in Figure 4 (dotted lines) for the *soft-band* (top panel) and

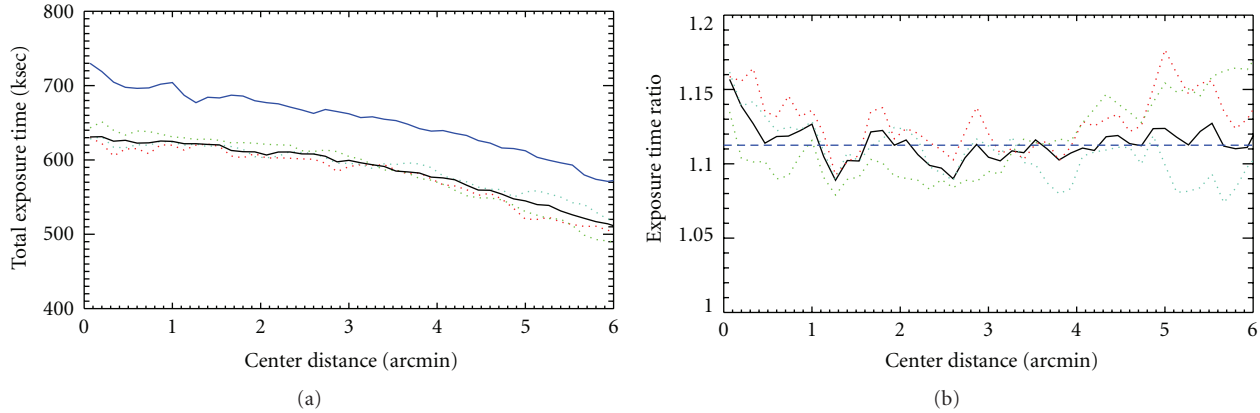


FIGURE 3: (a) Comparison of the stacked effective exposure time radial profile with respect to the defined image center positions in the cluster fields (blue solid line), the average control field (black solid line), and the three individual control fields (dotted lines). (b) Fractional difference of the stacked effective exposure time profile of the cluster field and the average control fields as a function of radial distance from the field centers (black solid line). The blue dashed line indicates the median exposure time difference of 11.3%, dotted lines depict the individual control fields as before.

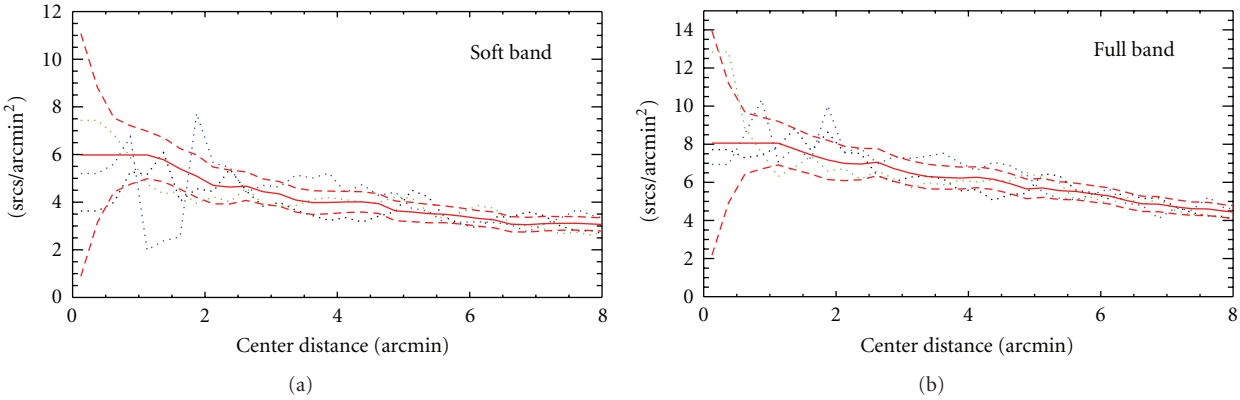


FIGURE 4: Background model as a function of center distance for the detected sources in the *soft band* (a) and the *full band* (b). The red solid line shows the average smoothed background model with 1σ uncertainties indicated by the red dashed lines, whereas the colored dotted lines show the extracted counts in the three control fields. The gradual decline as a function of off-center angle reflects the radial change in effective exposure time as shown in Figure 3.

full-band detection (bottom panel). The final robust background model for the further analysis should on one hand reflect the gradual radial change in effective exposure time and on the other hand be a smooth function in order to suppress Poisson fluctuations in the individual radial bins, in particular close to the central position where the counts per bin drop to $\lesssim 5$. This is achieved by applying a boxcar filter with a radius of four bins (each $15''$) and inverse variance weighting to the averaged radial count distribution of the control fields. The resulting final average background models for the radial distribution in the *soft band* and *full band* are shown by the solid red lines in Figure 4, whereas the dashed red lines illustrate the estimated 1σ Poisson uncertainties.

The determined background model may still be subject to residual systematic uncertainties not fully captured by the adopted approach with the three independent control fields and the applied average correction factor. Such residual systematic uncertainties include (i) slight effective exposure time offsets dependent on the exact locations of the optical axes of the different detectors, (ii) the effects of the removed extended and spurious sources, and (iii) slight geometric area mismatches at large off-center positions due to the non-axis-symmetric field-of-view edge. The magnitude of these potential systematic effects needs to be estimated for the combined stacked fields, which are the sums of the 22 random center positions across the XMM-*Newton* FoV. The potential exposure time offsets (i) can be shown to amount to a negligible sub-per-cent effect. Similarly, the effect of the area covered by the removed extended sources (ii) is insignificant and stacking analysis tests without any sources removed showed the qualitative same results. The largest residual systematic arises from geometric edge effects (iii), which can be estimated analogous to the statistical spatial uncertainty of a one-dimensional random walk with 22 steps, since the edge effects are randomly positive or negative relative to the cluster field. In this case, the step size for

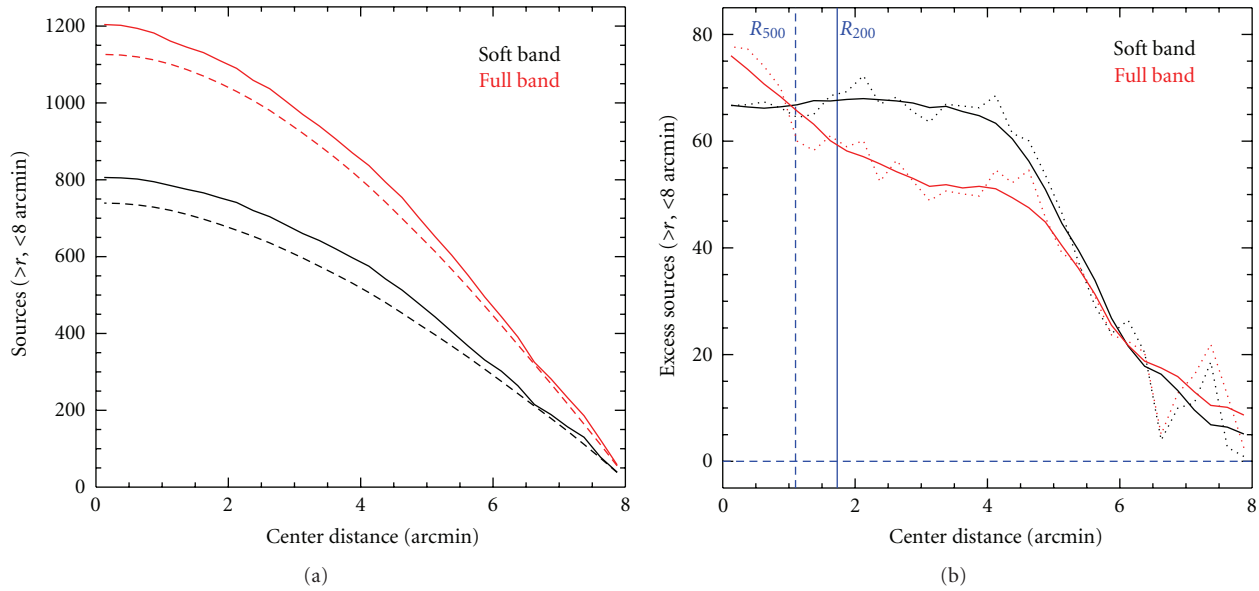


FIGURE 5: Cumulative number count distributions starting from an off-center angle of 8′ inwards. (a) Total cumulative counts for cluster fields (solid lines) in comparison to the median background models (dashed lines) for the *soft-band* (black) and *full-band* detection (red). (b) Cumulative distribution of excess counts after subtraction of the background model for the *soft band* (black) and *full-band* (red). Dotted lines show the original measured distribution, whereas solid lines depict the smoothed trends after boxcar filtering. Average cluster radii R_{500} (dashed) and R_{200} *(solid)* for the distant cluster sample are marked by the blue vertical lines.

the random walk is given by the average geometric mismatch of the analysis area per field of 8-9 arcmin^2, resulting in a total statistical geometric mismatch after 22 steps of $\pm\sqrt{22}\times 9 \simeq \pm 42$ arcmin^2. Using the average source densities yields approximate upper limits for the total impact of this effect of $\sigma_{\rm sys}(\textit{full-band}) \lesssim 12$ sources and $\sigma_{\rm sys}(\textit{soft-band}) \lesssim 7$ point sources. Consequently, the total systematic error budget is estimated to be $\lesssim$60% of the total Poisson uncertainties derived in the next section and is hence still subdominant compared to the statistical errors.

3. Results

In the following, the results of the X-ray stacking analysis are presented for the cumulative number counts, the radial distribution of excess counts, and an evaluation of source counts along the principal cluster elongation axis. All results are given in stacked units, that is, the sum of all 22 cluster environments is combined into a single radial X-ray source count profile. For all analyses a radial bin size for each independent ring segment of 15″ is used and the maximal off-center radius is limited to 8′ (∼4 Mpc) in order to avoid significant field-of-view edge effects.

3.1. Cumulative Source Counts. The left panel of Figure 5 displays the cumulative number counts for the cluster (red and black) and average background field counts (blue) starting at a radial distance of 8′ and moving inwards towards the center position. A significant excess of X-ray sources in the cluster field is apparent starting at off-center distances of about 6′ in both the *soft-band* (lower dashed lines) and *full-band* counts (upper solid lines). For the *full-band* source list the number of excess X-ray counts in the cluster field within the 8′ analysis radius is about 78, while the *soft band* shows approximately 67 excess counts. This translates into a fractional source excess in the distant cluster environments of +9.0% in the *soft band* and +6.9% in the *full band*.

Interestingly, the radial distribution of the excess counts for the two different band schemes is different, which is shown in the right panel of Figure 5. The smoothed trends (solid lines) indicate that the main excess counts in the *soft-band* (black) originate at cluster-centric distance of 4–6′, corresponding to 2-3 Mpc, while excess source counts in the *full band* rise sharply all the way to the center.

Applying Poisson statistics to the total cumulative background counts, which are based on three independent control fields, the statistical 1σ uncertainties can be estimated as $\sqrt{N_{\rm tot}/3}$, which amounts to $\sigma^{\rm soft}_{\rm tot,stat} \simeq 15.7$ sources and $\sigma^{\rm full}_{\rm tot,stat} \simeq 19.4$ counts for the two different band schemes. The significance of the total detected X-ray source excess within a cluster-centric distance of 8′ is hence $4.2\sigma_{\rm stat}$ in the *soft-band* and $4.0\sigma_{\rm stat}$ for the *full-band* detection. The average number of detected excess X-ray AGN per cluster environment amounts to 3.0 ± 0.7 *softband* and 3.5 ± 0.9 *fullband* AGN, respectively, which are to be interpreted as lower limits since not the full geometric area out to 8′ was covered in the detector FoV for each cluster environment.

3.2. Radial Distribution of Excess AGN. The background-subtracted radial distribution of excess X-ray sources in Figure 6 shows the data points for each individual bin with Poisson errors (black) and the smoothed radial trend

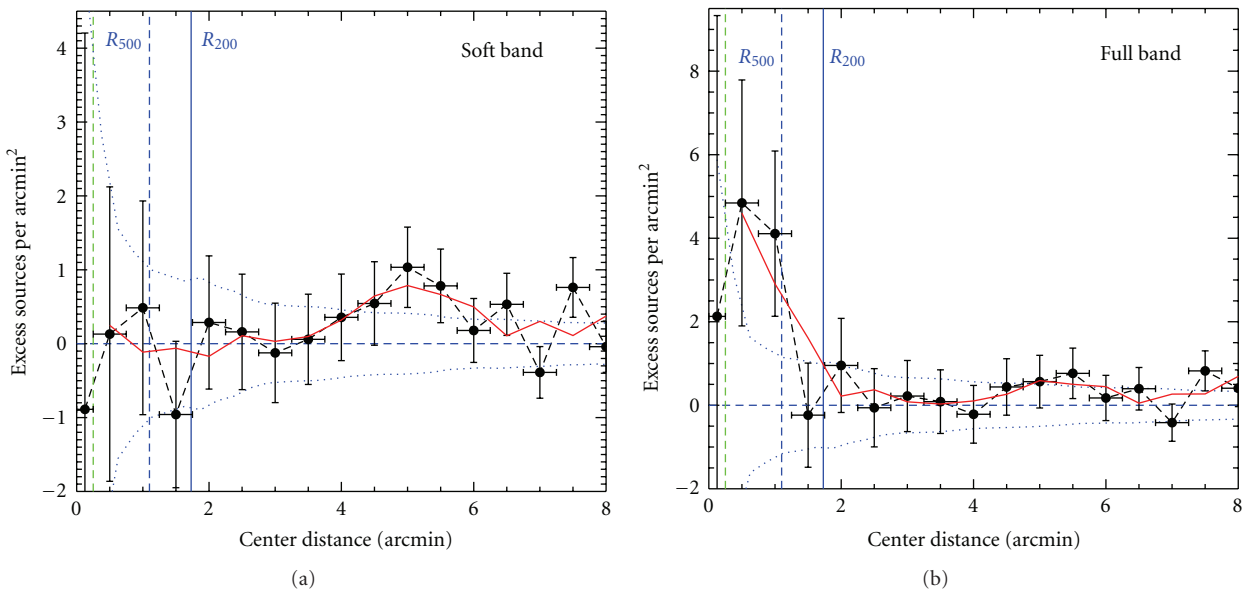


FIGURE 6: Background-subtracted radial distributions of excess X-ray sources for the *soft band* (a) and *full-band* (b). Dotted blue lines indicate the 1σ uncertainties about the mean background counts (blue dashed). Data points and the black dashed line display the measurements in each of the independent radial bins, whereas the red solid line shows the boxcar smoothed radial trend. The green dashed vertical line on the left shows the 15″ radius below which the measurements are biased due to the comparable spatial resolution of XMM-*Newton*. The blue vertical lines mark the average cluster radii R_{500} (dashed) and R_{200} (solid).

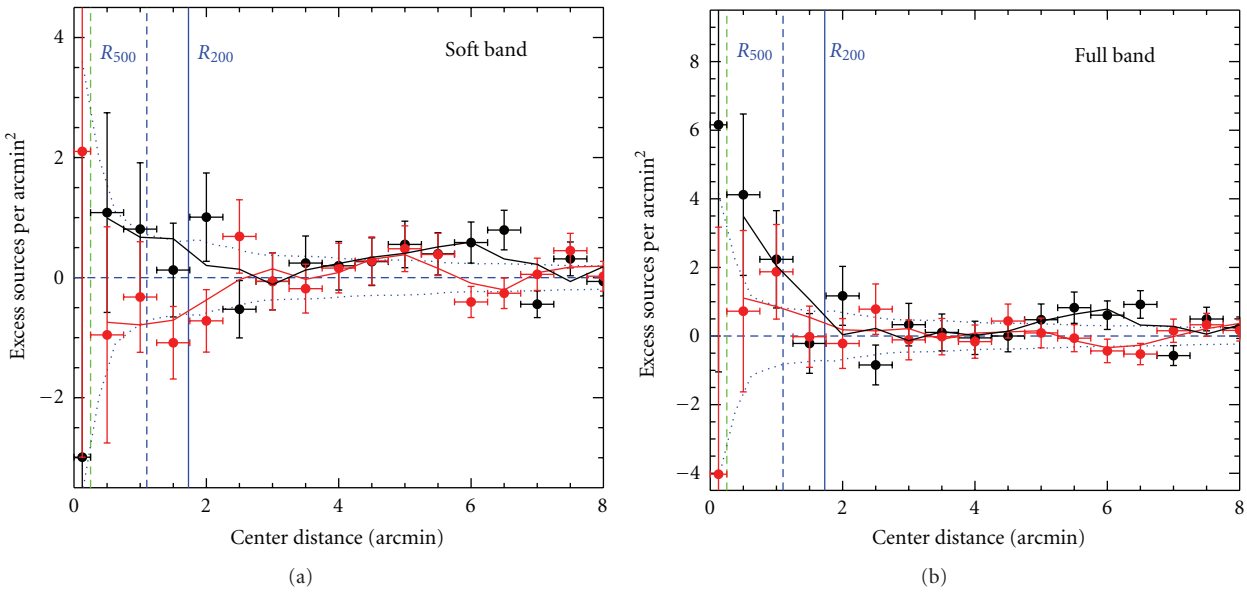


FIGURE 7: Same as Figure 6 but now distinguishing the sectors aligned with the cluster elongations (black points and solid black line) and the perpendicular direction (red points and solid red line).

(solid red line) for the *soft-band* (upper panel) and *full-band* detection (lower panel). The observed trends in the cumulative distribution of the previous section can now be evaluated for the independent data points with associated error bars and the given uncertainty in the subtracted background counts (blue dotted lines). For representation purposes (Figures 6 and 7) and better statistics in each bin, the displayed radial bins were increased to 30″ each outside the biased central 30″ region (green dashed vertical line, (see Section 2.1)), which is also excluded for the smoothed solid trend lines.

In the *soft band* (upper panel), the most significant feature in the radial distribution is the hump between cluster-centric distance 4′–6′as depicted by the red trend line. The covered solid angle at cluster-centric distances $4' \leq r \leq 6'$ is 63 square arcmin implying that significant excess counts

normalized to a unit area are hard to measure in comparison to the background counts. However, in this $4' \lesssim r \lesssim 6'$ range six independent adjacent radial bins show an excess, of which four have more than 1σ significance each, yielding a high combined confidence that this feature is indeed real. The hump is also discernible in the *full-band* detection (lower panel), although at lower statistical significance.

The most prominent feature in the *full-band* detection is the strong rise of excess X-ray sources at radial distances below $2'$, which correspond approximately to the radius inside the fiducial R_{200} (vertical blue line) of the distant cluster sample. (R_{200} (R_{500}) are the radii for which the mean enclosed total mass density of the cluster is 200 (500) times the critical energy density of the Universe $\rho_{\rm cr}(z)$ at the given redshift z. The considered distant cluster sample has an average R_{200} (R_{500}) of about $100''(64'')$ with a standard deviation of $\pm 20\%$). Owing to the relatively small enclosed central area with a corresponding low number of total counts, the statistical uncertainties inside the fiducial cluster regions are quite significant. Nevertheless, the two innermost bins outside the core region indicate a clear and significant trend of a steep inner radial profile of *full-band* detected excess sources. The inner profile of the *soft-band* sources in the upper panel, on the other hand, is fully consistent with a null excess of sources inside radii $r < 2'$.

3.3. X-Ray Counts along the Principal Cluster Elongation Axis. In order to investigate the possibility that AGN activity occurs along a preferential direction in the cluster-frame system, the radial profile analysis for excess X-ray counts was repeated with the rotated stacked source lists that aligned the principal elongation axis of the cluster X-ray emission in the vertical direction (see Section 2.3 and bottom panels of Figure 2). The radial distribution was then split into two disjoint sectors, the *aligned* sector along the cluster elongation direction covering the range within an angle of $\pm 45^\circ$ to the vertical axis with the cluster center as origin and the *perpendicular* sector within $\pm 45^\circ$ to the horizontal coordinate axis. The average background models (Figure 4) and the associated uncertainties were adopted to account for the factor of two smaller total area covered by each of the two sectors. The background subtracted radial counts along the *aligned* (black) and *perpendicular* sectors (red) are displayed in Figure 7 for the *soft-band* (upper panel) and *full-band* detection (lower panel).

For the *full-band* counts in the lower panel the profiles at $r < 2'$ indicate that the X-ray source excess inside the fiducial cluster radius R_{200} is dominated by AGN along the cluster-*aligned* direction (black line). Additionally, the secondary hump at $4.5'$–$6.5'$ is now more significant in the *aligned* (black) direction, while consistent with a null excess for the *perpendicular* sector (red). However, owing to the lower number of excess counts in the individual split sector bins with a corresponding decreased statistical significance these results are to be considered as tentative.

For the *soft-band* detection in the upper panel, the $4'$–$6'$ hump appears also to be enhanced in the *aligned* (black) direction around $r \sim 6'$. Below a radius of $2'$ the sector counts seem to split into an excess in the *aligned* and a deficit in the *perpendicular* sector, but both are on the level of the background uncertainty.

3.4. Robustness of Results. The tentative nature of the latter results in Section 3.3 with the separated sectors along two directions illustrates the limits of the statistical detectability of radial features in the X-ray source distribution for the currently available sample size of distant X-ray clusters. However, the results based on the overall excess counts in the radial distributions in Sections 3.1 and 3.2 are detected on the 4σ level (for $r < 8'$) and can hence be considered as robust.

In particular the central rise of the *full-band* counts and the *soft-band* $4'$–$6'$ hump seem to be real features in the radial distribution of distant cluster environments with an average of 1-2 excess sources per cluster in each feature. A central rise in the *full band* while flat in the *soft band* could well be physically explained by a dominance of faint and spectrally hard cluster AGN. With on average about one *full-band* excess source per cluster in the inner $\sim 2'$, the central part of the background model would have to be underestimated by almost a factor of two to attribute the observed cumulative excess to a background systematic.

The cumulative counts for the $4'$–$6'$ hump in the *soft band* exceed the *full-band* counts by about 15 sources around $r \simeq 4'$ (see right panel of Figure 5). On the other hand, all *soft-band* sources should also be part of the *full-band* catalogs, that is, the cumulative *full-band* excess would be expected to be at least as high for the *soft-band*. However, this slight discrepancy at intermediate off-center angles between $2'$ and $5'$ in the cumulative excess distribution can likely be attributed to Poisson noise in the *full-band* background counts, which was estimated to be slightly higher in magnitude compared to the measured difference of excess *soft-band* sources at the hump location. The *full-band* signal at these intermediate off-center distances hence seems to be diluted by the larger Poisson noise at a level within the statistical expectations.

The presence of the hump in the *full band* itself, although at lower confidence level, provides some extra confidence in its real nature. Other cross-checks with the exposure time profile (Figure 3) and the background (Figure 4) do not reveal any particular radial feature in the $4'$–$6'$ range that could attribute the observed hump to a systematic. The combined systematic error budget was estimated in Section 2.3 to be at most 60% of the total statistical Poisson error ($\lesssim$12 sources), which is far below the observed signal. The cumulative excess magnitude in the hump of more than three dozen sources also makes the possibility of individual systematic structures at $\sim 5'$ in the cluster fields an unlikely explanation.

4. Discussion

4.1. Spectroscopically Confirmed AGN. The ultimate proof to quantify the statistical properties of AGN in high-z clusters from the centers to the large-scale environments would

Table 2: List of spectroscopically confirmed X-ray AGN within a search radius of 12′(≃6 Mpc) from the cluster centers and with a spectroscopic redshift that is within ±3000 km/s of the cluster restframe velocity. The table lists the cluster environment ID of Table 1 in column (1), an AGN ID in (2), the angular distance to the cluster center in (3), the AGN redshift in (4), the RA and DEC coordinates in (5-6), the source name as listed in NED in (7), and a redshift reference in (8).

Cl. env.	AGN ID	Distance arcmin	z	RA J2000	DEC J2000	NED name	Reference
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C03	A01	9.52	1.486	03 : 39 : 08.3	+00 : 29 : 25	SDSS J033908.27 + 002924.9	[43]
C03	A02	11.3	1.483	03 : 39 : 12.3	+00 : 30 : 56	SDSS J033912.34 + 003055.5	[43]
C04	A03	1.31	1.462	22 : 16 : 03.7	−17 : 38 : 31	XMMUJ221603.6 − 173830	[44]
C04	A04	0.24	1.462	22 : 15 : 59.1	−17 : 37 : 54	PS1	[10]
C04	A05	0.12	1.453	22 : 15 : 58.9	−17 : 38 : 10	PS2	[10]
C06	A06	6.36	1.373	15 : 32 : 06.0	−08 : 30 : 55	XMS J153206.0 − 083055	[45]
C07	A07	5.32	1.330	00 : 36 : 18.6	−43 : 13 : 20	IRAC 109477	[46]
C07	A08	6.52	1.318	00 : 35 : 23.9	−43 : 16 : 38	IRAC 229193	[47]
C07	A09	9.02	1.334	00 : 36 : 22.1	−43 : 19 : 03	IRAC 111580	[47]
C08	A10	7.81	1.234	12 : 53 : 08.3	−29 : 20 : 10	GMOS-F2 08	[48]
C10	A11	3.08	1.179	03 : 02 : 14.8	+00 : 01 : 25	SDSS J030214.82 + 000125.3	[43]
C12	A12	11.6	1.110	22 : 06 : 26.0	−01 : 52 : 01	FBQS J2206 − 0152	[49]
C13	A13	10.8	1.120	03 : 38 : 10.2	+00 : 23 : 25	SDSS J033810.16 + 002325.1	[43]
C15	A14	7.97	1.053	02 : 26 : 38.0	−04 : 19 : 45	VVDS 020465089	[50]
C21	A15	7.66	0.932	01 : 04 : 02.8	−06 : 36 : 00	XMMUJ0104.1 − 0635 000	[51]

require a combination of high-resolution X-ray imaging plus an optical spectroscopic identification to identify all X-ray point sources. Such a data set is currently not yet available for sizable samples of $z > 0.9$ X-ray luminous systems. However, as a starting point the observed X-ray point source catalogs can be crossmatched with public redshift data from the NASA Extragalactic Data Base (http://nedwww.ipac.caltech.edu) (NED) complemented by individual cluster publications [10]. The resulting list of 15 spectroscopically confirmed and X-ray identified AGN within 12′ of the cluster positions and spectroscopic restframe velocity offsets of less than ±3000 km/s is shown in Table 2.

Five of the spectroscopic AGN are located at cluster-centric angular separations $8' < r < 12'$, corresponding to projected distances of 4–6 Mpc. Ten spectroscopic members are found at the probed radial distance range of ≤8′, four of them at <4′ (<2 Mpc) and six in the range $4' < r < 8'$ (2–4 Mpc). The two objects A04 and A05 at the center of XMMXCS $J2215.9 - 1738$ at $z = 1.46$ are not spatially resolved with XMM-*Newton* [10]; all others are identified X-ray point sources (except object A12, which is outside the XMM-*Newton* FoV, but is classified in NED as QSO).

Out of the measured *full-band* X-ray point source excess of about 78 (±25%) in the cluster environments at $r \leq 8'$, eight or about 10% are spectroscopically confirmed X-ray AGN at the cluster redshift. As can be expected, the serendipitous spectroscopic identification rate of cluster AGN in this high-z sample is still fairly low, but sufficiently high to confirm the presence of X-ray AGN in the large-scale structure environment of the probed systems.

4.2. Selection Effects and Comparison to Chandra Observations. The considered distant galaxy cluster sample is X-ray selected based on XMM-*Newton* data as discussed in Sections 2.1 and 2.2, implying that potential selection effects could have an impact on some of the results presented in Section 3. All clusters were originally detected as *extended* X-ray sources, with most of the detection weight originating from the soft band owing to the expected spectral properties of high-z thermal ICM emission. In principle, cluster-embedded X-ray point sources could have both an (negative) antibias or (positive) bias effect on the XMM-*Newton* selection. An anti-bias, that is, a missed fraction of proper cluster sources, arises when a bright central AGN is dominating over the extended thermal emission in the soft band, resulting in a potential misclassification of the cluster as an X-ray point source. On the other hand, an opposite bias effect could arise in the cases where a weak unresolved cluster AGN adds a subdominant fraction of flux to the underlying extended cluster emission or where the superposition of multiple embedded X-ray point sources even mimics an X-ray extended at the given XMM-*Newton* resolution limit.

Neither of the described possible anti-bias or bias effects are currently accessible to a robust quantitative evaluation since this would require large well-defined $z > 0.9$ cluster samples based on different selection techniques (e.g., IR, SZE, and X-ray) and followed up with high-resolution *Chandra* observations and extensive optical spectroscopy that are not yet available for such a study. The two most distant spectroscopically confirmed infrared selected systems in the galaxy group mass regime, CL J1449 + 0856 at $z = 2.07$ [20] and SXDF-XCL $J0218 - 0510$ at $z = 1.62$ [11],

feature both a central X-ray point source that dominates the X-ray emission and could hint at a common occurrence of central AGN in low-mass systems at $z \gtrsim 1.6$. However, given the different system masses, redshift regime, and selection techniques such first hints based on low-number statistics may not be representative for the considered X-ray selected high-z cluster sample.

A meaningful cross-check on the influence of potentially unresolved embedded cluster AGN for the cluster sample of this work can currently only be obtained from the available published *Chandra* observations of a subsample of the five systems C04 [10], C05 [21], C08 [22], C16 [23], and C19 [18] listed in Table 1. For the four systems at $z \lesssim 1.4$ (C05, C08, C16, C19) *Chandra* did not reveal any additional central point sources embedded in the extended ICM emission. Only for the system C04 at $z = 1.46$ the high-resolution observations disclosed the two additional sub-dominant clusters AGN A04 and A05 (see Table 2 and Section 4.1) at $r \lesssim 15''$ that are unresolved with XMM-*Newton* both for the soft- and full-band detection. The cross-check with this *Chandra* subsample is hence consistent with a fair XMM-*Newton* accounting of X-ray point sources in the distant cluster environments outside the biased central $15''$ radius, which was excluded from the trend analysis of the radial profiles in Sections 3.2 and 3.3. In particular, there is no indication that the soft-band radial point source profile could be antibiased in the range $0.25' < r \lesssim 2'$ due to the presence of extended ICM emission. Similarly, the *Chandra* sub-sample does not reveal any cases of systems that entered the distant X-ray cluster sample solely based on a bias effect of embedded point sources.

In summary, although some sample selection effects inherit to the XMM-*Newton* discovery method for the distant clusters may apply compared to an ideal (nonexistent) sample, a qualitative cross-check with the available *Chandra* observations for five systems did not reveal any evident biases concerning the radial profile of point sources in the main targeted cluster-centric distance range of $0.25' < r \lesssim 8'$.

4.3. Comparison to Previous Studies. The high-z results of this work are consistent with numerous previous *Chandra* studies that find a general trend of increasing AGN activity in cluster environments with redshift for X-ray selected samples (e.g., [1–3, 5]) and infrared-selected clusters [12]. Although the used sample is X-ray selected with XMM-*Newton*, implying the discussed natural anti-bias against potential systems with bright central soft-band AGN, a centrally rising radial distribution of lower luminosity AGN was found with about one detected excess source per system within a projected radius of 1 Mpc. However, because of the limited spatial resolution of XMM-*Newton*, the recovered excess sources are incomplete in the very core ($r \lesssim 15$–$20''$), where individual systems at the highest accessible redshifts have revealed embedded cluster AGN as discussed in Section 4.2.

Concerning the detected secondary $4'$–$6'$ hump (2-3 Mpc), Ruderman and Ebeling [6] reported on the same general radial profile shape of the X-ray point source excess around very massive MACS clusters at intermediate redshifts ($0.3 < z < 0.7$) with a central spike plus an additional broad excess at projected cluster-centric distances of 2-3 Mpc. They interpreted this finding as evidence for distinct triggering mechanisms of nuclear activity in the center through close encounters of infalling galaxies, and the cluster-field interface, where galaxy mergers fuel the central supermassive black holes. On the other hand, Gilmour et al. [4] did not recover the secondary radial excess feature in their larger sample of 148 clusters at $z < 0.9$. Their tests with cluster subsamples revealed that a point source excess at $r > 2$ Mpc may be systematically boosted by structures related to nonassociated foreground clusters in the FoV [4].

With respect to such a test, the current sample size of 22 high-z cluster fields is still too small to discard all fields with a potential contamination from foreground structures, which are ubiquitous at the exposure depth of the considered observations. For the nature of the $4'$–$6'$ hump, a systematic bias related to foreground structures can hence not be fully ruled out at this point, although the measured excess in this region would require multiple such structures not captured by the background model. A test with the two most crowded fields removed from the analysis (excluding clusters C03, C04, C09, C13, C17, and C22) showed no qualitative difference concerning the shape and the presence of the hump.

On the other hand, a cluster-associated excess of X-ray AGN at projected cluster-centric radii of $r > 4'$ (>2 Mpc) must be present as revealed by the 11 spectroscopically confirmed objects in Table 2 in the range $4' < r < 12'$. Moreover, Rumbaugh et al. [9] showed in their spectroscopic study of cluster environments in the redshift range 0.7–0.9 that half of the AGN host galaxies are located at projected distances of >1.5 Mpc away from the nearest cluster or group. In particular, they also confirmed a significant difference in AGN X-ray luminosity between objects in the dense inner cluster regions and the low-density large-scale structure environments in the sense that the central objects show low luminosities, while the brightest objects are all found at large cluster centric distances.

4.4. The Distant Cluster-AGN Connection. Taking the presented results at face value, the following scenario for AGN activity in distant X-ray cluster environments at $0.9 < z \lesssim 1.6$ emerges. The detected significant (4σ) excess of X-ray AGN in the distant cluster environments at angular radii of up to $8'$ is split into two distinct populations that reflect different triggering mechanisms of nuclear activity.

At cluster-centric radii less than the fiducial average cluster size of the sample of $R_{200} \simeq 830\,\mathrm{kpc} \simeq 1.7'$ a centrally peaked population of low-luminosity X-ray AGN exists. The conclusion of low X-ray luminosities (or a heavily absorbed type-II spectrum) for this population originates from the fact that the central excess is not observed in the *soft-band* detection with its higher source significance threshold. The tentative result of Section 3.3 suggests that the AGN activity preferentially occurs along the main matter infall axis of the cluster, as indicated by the principal elongation axis of the extended cluster emission. This nuclear activity is

likely triggered by close encounters of infalling objects and is mostly found in red or green transitional galaxies as reported in [9]. Overall, the detected (incomplete) excess of about one AGN source per cluster field inside its fiducial radius R_{200} is still very moderate when considering the expected (e.g., [24]) average enclosed stellar mass of $\simeq 4 \times 10^{12}\,M_\odot$ for the cluster systems. Comparing this value to the average AGN abundance in the COSMOS field in the same $0.9 \leq z \leq 1.6$ redshift range of one X-ray AGN per $2 \times 10^{12}\,M_\odot$ in stellar mass (Bongiorno et al., in prep.) suggests that the AGN activity in high-z cluster environments is still suppressed by a factor of two compared to the field abundance.

The second AGN population is preferentially located at cluster-centric distances of 2-3 Mpc as part of the observed 4′–6′ hump of excess sources. The *soft-band* detection of this feature indicates significantly higher soft-band X-ray luminosities compared to the central AGN. The projected distance of the hump location corresponds to about $3 \times R_{200}$ of the average cluster radius. At this distance, the relative galaxy velocities are still small while the object density is already significantly enhanced, resulting in an environment where major merging processes are expected to be efficient. The observed AGN excess in the larger-scale environment of distant X-ray clusters is hence likely to be attributed to merger-induced nuclear activity in "quasar mode" [25] occurring at a potential sweet spot for merging events in the infall regions at cluster-centric projected distances of around $3 \times R_{200}$.

4.5. Implications for eROSITA. The next generation all-sky X-ray survey will be conducted by the upcoming eROSITA mission [17], which has an expected average survey PSF of 25″–30″ and hence provides a lower spatial resolution than *Chandra* and XMM-*Newton*. In order to evaluate the performance for distant cluster applications, the high-z AGN excess close to the cluster position is of critical importance.

The measured average excess of about one *full-band*-detected AGN within R_{200} with a centrally peaked radial profile may pose a severe potential bias for X-ray spectroscopic applications at $z > 0.9$, most notably high-z ICM temperature measurements. For faint and compact high-z systems, the full detected X-ray emission within the extraction aperture has to be considered to obtain enough signal, without the possibility to subtract embedded point sources. Hidden AGN can hence harden the spectrum and bias the T_X determination high (see, e.g., [10]).

Although the statistical uncertainties in the central region are still large, the measured *soft-band* radial profile did not reveal a significant central excess of AGN. The situation for the detection of extended high-z Xray sources and accurate *soft-band* flux determinations hence still looks very promising. These tasks can be performed in the soft-band only, most commonly in the 0.5–2 keV range, where the cluster AGN contribution seems to be low for the presented X-ray selected sample and sensitivities, which correspond to the planned eROSITA deep fields. The detectability of distant clusters with eROSITA may hence not be significantly influenced by the present findings. The measured soft-band X-ray luminosity for high-z eROSITA clusters will hence likely be the preferable low-bias mass proxy.

5. Summary and Conclusions

This work investigated the X-ray point source excess in the environment of distant X-ray luminous galaxy clusters in the redshift range $0.9 < z \lesssim 1.6$ based on a sample of 22 systems with an average mass of $2 \times 10^{14}\,M_\odot$ compiled by the XMM-*Newton* Distant Cluster Project. The X-ray source detection in the available XMM-*Newton* observations for each cluster with a median clean field exposure time of 19 ksec was performed in two energy ranges, a 0.35–2.4 keV *soft band* and a 0.3–7.5 keV *full band*. The X-ray source counts were stacked in cluster-centric coordinates and compared to the average background counts extracted from three independent control fields in the XMM-*Newton* field of view. The main findings of this X-ray stacking analysis can be summarized as follows.

(1) The cumulative radial X-ray source counts within a cluster-centric region of 8′ reveal a significant excess of ∼67 (*softband*) and ∼78 (*soft-band*) sources with a statistical confidence of 4.2σ in the *soft band* and 4.0σ in the *full band*. The resulting average detected point source excess of 3.0 ± 0.7 (*soft-band*) and 3.5 ± 0.9 (*full-band*) sources per cluster environment is to be interpreted as lower limit due to an incomplete coverage of the full geometric area out to 8′.

(2) The radial cumulative distributions of the detected excess counts show a different radial behavior for the *soft band* and *full band*. *Soft-band* detected excess sources are mainly located at cluster-centric projected distances of 2-3 Mpc (the 4′–6′-hump), while the *full-band* counts show an additional steep rise at small distances of $r < 2'$.

(3) The radial profiles of the background-subtracted excess counts confirm the statistical significance of both radial features. The centrally peaked point source excess within the average cluster radius R_{200} is consistent with various previous studies, most of them conducted at lower redshifts. The detected (incomplete) average excess per cluster is found to be about one point source per system with evidence for lower X-ray luminosities below the *soft-band* detection limit.

(4) A second analysis was performed with all source lists rotated about the cluster center in a way that the principal elongation axes of the cluster emission are aligned for all systems. With the interpretation that the cluster elongation points in the direction of the main cluster assembly axis the tentative conclusion can be derived that the observed AGN activity is mostly occurring along these identified matter infall directions.

(5) The second outer radial feature at cluster-centric distances of 2-3 Mpc was previously reported also for

an intermediate-redshift cluster sample by Ruderman and Ebeling [6]. However, a systematic boosting of this feature by foreground structures cannot be fully ruled out with the statistics of the present sample. The cross-correlation of public spectroscopic redshift information with the X-ray data, on the other hand, confirms the presence of a significant population of X-ray AGN in the cluster environments beyond a projected distance of 2 Mpc. The observed 4′–6′ hump feature is consistent with a population of bright *soft-band*-detected AGN triggered at a project distance of about $3 \times R_{200}$.

(6) Taking all results at face value lends support to the idea of two different AGN populations and triggering mechanisms of nuclear activity in the distant cluster environments. In this picture, the fiducial cluster regions inside R_{200} of high-z X-ray luminous systems harbor low-luminosity AGN triggered by close galaxy encounters in the infall regions, while the excess in the outer cluster environment at distances of 2-3 Mpc is due to major-merger induced AGN activity.

(7) With respect to distant cluster applications with the upcoming all-sky survey eROSITA, the results suggest that the spectroscopic temperature analysis of samples of distant cluster sources may be biased highly due to embedded point sources, while the detection and flux measurements in the soft band may not be significantly influenced in the general case.

The presented results are based on the currently largest homogeneously selected sample of X-ray luminous clusters at $z > 0.9$. Statistical improvements and tests of the discussed results will require significantly larger, well-defined samples of distant X-ray clusters of about 50 objects, which should soon be available. The ultimate experiment of X-ray AGN in distant cluster environments at $r > 1$ Mpc will be made possible by eROSITA in the near future.

Acknowledgments

The authors thank the anonymous referee for insightful comments that helped to improve the clarity of the paper. They thank the XDCP team members for carrying out the distant cluster survey. R. Fassbender would like to thank Gabriel Pratt for his comments, Andrea Merloni for helpful discussions, and Angela Bongiorno for the invitation to the special issue. This research was supported by the DFG cluster of excellence "Origin and Structure of the Universe" (http://www.universe-cluster.de/) and by the DFG under Grants BO 702/16-3. The XMM-*Newton* project is an ESA Science Mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

[1] M. Branches, I. M. Gioia, C. Fanti, R. Fanti, and N. Cappelluti, "Chandra point-source counts in distant galaxy clusters," *Astronomy and Astrophysics*, vol. 462, no. 2, pp. 449–458, 2007.

[2] N. Cappelluti, M. Cappi, M. Dadina et al., "X-ray source overdensities in Chandra distant cluster fields: a new probe to map the cosmic tapestry?" *Astronomy and Astrophysics*, vol. 430, no. 1, pp. 39–45, 2005.

[3] J. Eastman, P. Martini, G. Sivakoff, D. D. Kelson, J. S. Mulchaey, and K. V. Tran, "First measurement of a rapid increase in the AGN fraction in high-redshift clusters of galaxies," *Astrophysical Journal*, vol. 664, no. 1, pp. L9–L12, 2007.

[4] R. Gilmour, P. Best, and O. Almaini, "The distribution of active galactic nuclei in a large sample of galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 392, no. 4, pp. 1509–1531, 2009.

[5] P. Martini, G. R. Sivakoff, and J. S. Mulchaey, "The evolution of active galactic nuclei in clusters of galaxies to redshift 1.3," *Astrophysical Journal Letters*, vol. 701, no. 1, pp. 66–85, 2009.

[6] J. T. Ruderman and H. Ebeling, "The origin of the spatial distribution of X-ray-luminous active galactic nuclei in massive galaxy clusters," *Astrophysical Journal*, vol. 623, no. 2, pp. L81–L84, 2005.

[7] E. Koulouridis and M. Plionis, "Luminous X-ray active galactic nuclei in clusters of galaxies," *Astrophysical Journal Letters*, vol. 714, no. 2, pp. L181–L184, 2010.

[8] O. Johnson, P. N. Best, and O. Almaini, "The content of active galactic nuclei in the $z = 0.83$ cluster MS 1054-0321," *Monthly Notices of the Royal Astronomical Society*, vol. 343, no. 3, pp. 924–932, 2003.

[9] N. Rumbaugh, D. D. Kocevski, R. R. Gal et al., "The evolution and environments of X-ray emitting active galactic nuclei in high-redshift large-scale structures," *The Astrophysical Journal*, vol. 746, p. 155, 2012.

[10] M. Hilton, E. Lloyd-Davies, S. A. Stanford et al., "The *Xmm* cluster survey: active galactic nuclei and starburst galaxies in xmmxcs J2215.9−1738 at $z = 1.46$," *The Astrophysical Journal*, vol. 718, no. 1, p. 133, 2010.

[11] M. Pierre, N. Clerc, B. Maughan et al., "A Chandra view of the $z = 1.62$ galaxy cluster IRC-0218A," *Astronomy & Astrophysics*, vol. 540, article A4, 2012.

[12] A. Galametz, D. Stern, P. R. M. Eisenhardt et al., "The cosmic evolution of active galactic nuclei in galaxy clusters," *Astrophysical Journal Letters*, vol. 694, no. 2, pp. 1309–1316, 2009.

[13] P. R. M. Eisenhardt, M. Brodwin, A. H. Gonzalez et al., "Clusters of galaxies in the first half of the universe from the IRAC shallow survey," *Astrophysical Journal Letters*, vol. 684, no. 2, pp. 905–932, 2008.

[14] R. Fassbender, H. Bohringer, A. Nastasi et al., "The x-ray luminous galaxy cluster population at $0.9 < z \lesssim 1.6$ as revealed by the XMM-*Newton* distant cluster project," *New Journal of Physics*, vol. 13, Article ID 125014, 2011.

[15] R. Fassbender, *Studying cosmic evolution with the XMM-Newton distant cluster project: x-ray luminous galaxy clusters at $z \sim > 1$ and their galaxy populations*, Ph.D. thesis, Ludwig Maximilians University Munich, 2007, astro-ph/0806.0861.

[16] H. Bohringer, C. R. Mullis, P. Rosati et al., "Galaxy cluster archaeology," in *Proceedings of the ESO Messenger*, vol. 120, pp. 33–36, 2005.

[17] P. Predehl, R. Andritschke, H. Bohringer et al., "eROSITA on SRG," in *Proceedings of the Society of PhotoOptical Instrumentation Engineers SPIE Conference Series*, vol. 7732, Society of Photo-Optical Instrumentation Engineers (SPIE), 2010.

[18] R. Fassbender, H. Böhringer, J. S. Santos et al., "A panchromatic view of the galaxy cluster XMMU J1230.3 + 1339 at $z = 0.975$: observing the assembly of a massive system," *Astronomy and Astrophysics*, vol. 527, no. 10, article A78, 2011.

[19] N. Cappelluti, G. Hasinger, M. Brusa et al., "The XMM-Newton wide-field survey in the COSMOS field. II. X-ray data and the log N-log S relations," *Astrophysical Journal*, vol. 172, no. 1, pp. 341–352, 2007.

[20] R. Gobat, E. Daddi, M. Onodera et al., "A mature cluster with X-ray emission at $z = 2.07$," *Astronomy and Astrophysics*, vol. 526, no. 16, article A133, 2011.

[21] P. Rosati, P. Tozzi, R. Gobat et al., "Multi-wavelength study of XMMU J2235.3 – 2557: the most massive galaxy cluster at $z > 1$," *Astronomy & Astrophysics*, vol. 508, no. 2, pp. 583–591, 2009.

[22] P. Rosati, P. Tozzi, S. Ettori et al., "Chandra and XMM-Newton observations of RDCS 1252.9 − 2927, a massive cluster at $z = 1.241$," *Astronomical Journal*, vol. 127, no. 1, pp. 230–238, 2004.

[23] B. J. Maughan, L. R. Jones, M. Pierre et al., "Testing the galaxy cluster mass-observable relations at $z = 1$ with XMM-Newton and Chandra observations of XLSSJ022403.9-041328," *Monthly Notices of the Royal Astronomical Society*, vol. 387, no. 3, pp. 998–1006, 2008.

[24] S. Giodini, D. Pierini, A. Finoguenov et al., "Stellar and total baryon mass fractions in groups and clusters since redshift," *Astrophysical Journal Letters*, vol. 703, no. 1, pp. 982–993, 2009.

[25] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[26] J. S. Santos, R. Fassbender, A. Nastasi et al., "Discovery of a massive X-ray luminous galaxy cluster at $z = 1.579$," *Astronomy and Astrophysics*, vol. 531, article L15, 2011.

[27] R. Fassbender, A. Nastasi, H. Bohringer et al., "The X-ray luminous galaxy cluster XMMU J1007.4 + 1237 at $z = 1.56$. The dawn of starburst activity in cluster cores," *Astronomy and Astrophysics*, vol. 527, article L10, 2011.

[28] A. Nastasi, R. Fassbender, H. Böhringer et al., "Discovery of the X-ray selected galaxy cluster XMMU J0338.8 + 0021 at $z = 1.49$: indications of a young system with a brightest galaxy in formation," *Astronomy and Astrophysics*, vol. 532, article L6, 2011.

[29] S. A. Stanford, A. K. Romer, K. Sabirli et al., "The XMM cluster survey: a massive galaxy cluster at $z = 1.45$," *Astrophysical Journal*, vol. 646, no. II, pp. L13–L16, 2006.

[30] M. Hilton, C. A. Collins, S. A. Stanford et al., "The XMM cluster survey: the dynamical state of XMMXCS J2215.9 − 1738 at $z = 1.457$," *Astrophysical Journal Letters*, vol. 670, no. 2, pp. 1000–1009, 2007.

[31] M. Hilton, S. A. Stanford, J. P. Stott et al., "The xmm cluster survey: galaxy morphologies and the color-magnitude relation in XMMXCS J2215.9 − 1738 at $z = 1.46$," *Astrophysical Journal Letters*, vol. 697, no. 1, pp. 436–451, 2009.

[32] C. R. Mullis, P. Rosati, G. Lamer et al., "Discovery of an x-ray-luminous galaxy cluster at $z = 1.4$," *Astrophysical Journal*, vol. 623, no. 2, pp. L85–L88, 2005.

[33] R. Šuhada, R. Fassbender, A. Nastasi et al., "Exploring the galaxy cluster-group transition regime at high redshifts: physical properties of two newly detected $z > 1$ systems," *Astronomy and Astrophysics*, vol. 530, article A110, 2011.

[34] G. Wilson, A. Muzzin, H. K. C. Yee et al., "Spectroscopic confirmation of a massive red-sequence-selected galaxy cluster at $z = 1.34$ in the sparcs-south cluster Survey," *The Astrophysical Journal*, vol. 698, p. 1943, 2009.

[35] A. de Hoon et al., submitted to. *Astronomyand Astrophysics*.

[36] K. S. Dawson, G. Aldering, R. Amanullah et al., "An intensive hubble space telescope survey for $z > 1$ type Ia supernovae by targeting galaxy clusters," *Astronomical Journal*, vol. 138, no. 5, pp. 1271–1283, 2009.

[37] D. Pierini, R. Suhada, R. Fassbender et al., "First simultaneous optical/near-infrared imaging of an X-ray selected, high-redshift cluster of galaxies with GROND: the galaxy population of XMMU J0338.7 + 0030 at $z = 1.1$," *Astronomy & Astrophysics*, vol. 540, article A45, 2012.

[38] A. D. Schwope, G. Lamer, A. de Hoon et al., "XMMU J100750.5 + 125818: a strong lensing cluster at $z = 1.082$," *Astronomy and Astrophysics*, vol. 513, no. 11, article L10, 2010.

[39] F. Pacaud, M. Pierre, C. Adami et al., "The XMM-LSS survey: the Class 1 cluster sample over the initial 5 deg 2 and its cosmological modelling," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 3, pp. 1289–1308, 2007.

[40] C. Adami, A. Mazure, M. Pierre et al., "The XMM-LSS survey: optical assessment and properties of different X-ray selected cluster classes," *Astronomy and Astrophysics*, vol. 526, no. 2, article A18, 2010.

[41] J. S. Santos, P. Rosati, R. Gobat et al., "Multiwavelength observations of a rich galaxy cluster at $z \sim 1$: the HST/ACS colour—magnitude diagram," *Astronomy and Astrophysics*, vol. 501, no. 1, pp. 49–60, 2009.

[42] R. Fassbender, H. Böhringer, G. Lamer et al., "Indications for 3 Mpc-scale large-scale structure associated with an X-ray luminous cluster of galaxies at $z = 0.95$," *Astronomy and Astrophysics*, vol. 481, no. 2, pp. L73–L77, 2008.

[43] D. P. Schneider, P. B. Hall, G. T. Richards et al., "The sloan digital sky survey quasar catalog. IV. Fifth data release," *Astronomical Journal*, vol. 134, no. 1, pp. 102–117, 2007.

[44] F. J. Carrera, J. Ebrero, S. Mateos et al., "The XMM-Newton serendipitous survey III. the AXIS X-ray source counts and angular clustering," *Astronomy and Astrophysics*, vol. 469, no. 1, pp. 27–46, 2007.

[45] X. Barcons, F. J. Carrera, M. T. Ceballos et al., "The XMM-Newton serendipitous survey," *Astronomy and Astrophysics*, vol. 476, no. 3, pp. 1191–1203, 2007.

[46] C. Feruglio, F. Fiore, F. La Franca et al., "The XMM-Newton survey of the ELAIS-S1 field - II. Optical identifications and multiwavelength catalogue of X-ray sources," *Astronomy and Astrophysics*, vol. 488, no. 1, pp. 417–428, 2008.

[47] N. Sacchi, F. La Franca, C. Feruglio et al., "Spectroscopic identifications of spitzer sources in the swire/xmm-newton/elais-s1 field: a large fraction of active galactic nuclei with high f(24 μm)/f(r) ratio," *Astrophysical Journal Letters*, vol. 703, no. 2, pp. 1778–1790, 2009.

[48] M. Tanaka, C. Lidman, R. G. Bower et al., "Star formation activities of galaxies in the large-scale structures at $z = 1.2$," *Astronomy and Astrophysics*, vol. 507, no. 2, pp. 671–682, 2009.

[49] M.-P. Veron-Cetty and P. Veron, "A catalogue of quasars and active nuclei: 10th edition," *Astronomy and Astrophysics*, vol. 374, pp. 92–94, 2001.

[50] O. Le Fèvre, G. Vettolani, B. Garilli et al., "The VIMOS VLT deep survey: first epoch VVDS-deep survey: 11 564 spectra

with $17.5 \leq \text{iAB} \leq 24$, and the redshift distribution over $0 \leq z \leq 5$," *Astronomy and Astrophysics*, vol. 439, no. 3, pp. 845–862, 2005.

[51] X. Barcons, F. J. Carrera, M. G. Watson et al., "The XMM-Newton serendipitous survey: II. First results from the AXIS high galactic latitude medium sensitivity survey," *Astronomy and Astrophysics*, vol. 382, no. 2, pp. 522–536, 2002.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 950641, 24 pages
doi:10.1155/2012/950641

Review Article

Evidence for AGN Feedback in Galaxy Clusters and Groups

Myriam Gitti,[1,2,3] Fabrizio Brighenti,[4] and Brian R. McNamara[2,5,6]

[1] *INAF-Osservatorio Astronomico di Bologna, Via Ranzani 1, 40127 Bologna, Italy*
[2] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge MA 02138, USA*
[3] *INAF-Istituto di Radioastronomia di Bologna, Via Gobetti 101, 40129 Bologna, Italy*
[4] *Dipartimento di Astronomia, Università di Bologna, Via Ranzani 1, 40127 Bologna, Italy*
[5] *Physics and Astronomy Department, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 2G1*
[6] *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON, Canada N2L 2Y5*

Correspondence should be addressed to Myriam Gitti, myriam.gitti@oabo.inaf.it

Received 29 July 2011; Accepted 14 September 2011

Academic Editor: Angela Bongiorno

The current generation of flagship X-ray missions, *Chandra* and *XMM-Newton*, has changed our understanding of the so-called "cool-core" galaxy clusters and groups. Instead of the initial idea that the thermal gas is cooling and flowing toward the center, the new picture envisages a complex dynamical evolution of the intracluster medium (ICM) regulated by the radiative cooling and the nongravitational heating from the active galactic nucleus (AGN). Understanding the physics of the hot gas and its interplay with the relativistic plasma ejected by the AGN is key for understanding the growth and evolution of galaxies and their central black holes, the history of star formation, and the formation of large-scale structures. It has thus become clear that the feedback from the central black hole must be taken into account in any model of galaxy evolution. In this paper, we draw a qualitative picture of the current knowledge of the effects of the AGN feedback on the ICM by summarizing the recent results in this field.

1. Introduction

The physics of the intracluster medium (ICM) of clusters and groups of galaxies is complex. The current generation of X-ray satellites, *Chandra* and *XMM-Newton*, has shown indeed that it is regulated by yet poorly understood non-gravitational processes beyond simple gravity, gas dynamics, and radiative cooling usually considered in the standard cold dark matter cosmological scenario [1]. In particular, an important discovery from high-resolution X-ray observations was that the amount of thermal gas radiatively cooling to low temperatures is much less than what is predicted by the standard "cooling flow" model (see [2–5] and references therein), thus radically changing our understanding of the so-called "cool-core" systems. The implication is that the central gas must experience some kind of heating due to a feedback mechanism that prevents cool cores from establishing cooling flows at the rates predicted by earlier, low-resolution X-ray observations. Establishing the source of this heating, and understanding when and how it takes place, has become a major topic of study in extragalactic astrophysics.

Based on observational evidence and theoretical modelling, the primary source of feedback has been identified in the outbursts and accompanying energy injection, likely intermittent, from the active galactic nucleus (AGN) of the dominant cD galaxies (e.g., [6] and references therein), which host the most massive black holes in the local Universe. AGNs manifest as central radio sources, which are commonly observed in cool-core clusters [7]. Most of the cool-core systems have highly disturbed X-ray morphologies, and radio observations clearly show that AGN jets are the cause of many of the structures revealed by the X-ray telescopes. Such surface brightness features, including apparent depressions or "cavities" in the X-ray images and sharp density discontinuities interpreted as shocks, indicate a strong interaction between the central AGN and the intracluster medium (ICM). The incidence and variety of cavities, shocks, and ripples observed both in the radio and in X-rays in the hot ICM provides direct evidence of the widespread presence of AGN-driven phenomena (e.g., [8, 9] for sample studies of clusters and groups, resp.).

Such AGN feedback has a wide range of impacts, from the formation of galaxies, through to the explanation of the observed relation between the black hole mass and the bulge velocity dispersion (which indicates a causal connection or feedback mechanism between the formation of bulges and their central black holes, e.g., [10]), to the regulation of cool cores which explains why cooling and star formation still proceed at a reduced rate. On the other hand, the details of how the feedback loop operates are still unknown. Feedback is also required to suppress the overproduction of massive galaxies predicted by dark-matter-only simulations and to break the self-similarity of clusters (e.g., [11–13]). The nature of this feedback is therefore vital to our understanding of galaxy evolution ([14] and references therein).

After a brief discussion of the importance of galaxy clusters and their scaling relations for the study of cosmic evolution (Section 2), we overview the role of AGN feedback in structure formation (Section 3) and the basic properties of clusters in X-rays emphasizing the hot intracluster medium (Section 4). We then focus on the observational evidence of AGN feedback in action in galaxy clusters and groups (Section 5) and finally give our conclusions (Section 6). *The present paper does not intend to be a comprehensive review, but aims instead at drawing a qualitative picture of the impact of the AGN feedback on the ICM addressed particularly to the novices in this field.* General reviews of clusters from an X-ray perspective were given by Sarazin [15] and more recently by Mushotzky [16] and Arnaud [17]. A review of clusters as cosmological probes was given by Voit [18], and cold fronts and shocks associated with cluster mergers were reviewed by Markevitch and Vikhlinin [19]. An exhaustive review of the issues of AGN heating in the hot atmospheres was recently given by McNamara and Nulsen [6].

Throughout the paper we assume a cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_M = 1 - \Omega_\Lambda = 0.3$, where not specified otherwise.

2. Clusters of Galaxies as Cosmological Probe

The existence of clusters of galaxies and of other cosmic structures demonstrates that the Universe is not perfectly homogeneous. The matter density of the primordial Universe must have been slightly inhomogeneous, with overdense perturbations which deviate from the mean density. In the so-called "Concordance Model" largely accepted today as the standard cold dark matter cosmological scenario, cosmic structures like galaxies and clusters of galaxies originate from the gravitational instability of these primordial density fluctuations. The formation of structures from perturbations in the density distribution of cold dark matter is a hierarchical process. Small subclumps of matter are the first to deviate from the Hubble flow, to collapse and to experience gravitational relaxation because the density perturbations have larger amplitudes on smaller mass scales. These small objects then undergo a merging process to form larger and larger structures, up to the clusters of galaxies [1].

Galaxy clusters trace the high-density tail of the primordial field of dark matter density perturbations, and their numerical density as a function of redshift, z, is highly sensitive to the specific cosmological scenario (e.g., [20] and references therein). Therefore, if one builds the so-called "cluster mass function" $n_M(M)$, that is, the number density of clusters with mass greater than M in a comoving volume element and determines its evolution with redshift, it is possible to constrain the main cosmological parameters from the comparison between the observations and the theory predictions. A complete understanding of the details of the process of hierarchical merging would require accurate numerical simulations. However, many fundamental aspects can be illustrated by spherically symmetric, simplified models of cluster formation. In particular, the combination of spherical top-hat collapse models with the growth function for linear perturbations (assumed to be gaussian) has led to a variety of semianalytical methods to express the cluster mass function in terms of cosmological parameters (the seminal work in this field is by [21]).

Therefore, the comparison between the theoretical mass function and the mass function determined from observations allows one to constrain the main cosmological parameters, although with a degeneracy between the matter density parameter, Ω_M, and the power spectrum normalization of the perturbations within a comoving sphere of radius $8\,\mathrm{h^{-1}\,Mpc}$, σ_8. Such a degeneracy can be broken by studying the redshift evolution of the mass function, which is highly sensitive to Ω_M, by taking into account the evolution of the observables (see the reviews by [18, 20, 22] and references therein). The accuracy of the cosmological parameter measurements is currently limited by uncertainties in the relations between cluster masses and the observable properties that trace these masses, such as luminosity or temperature. In order to measure the mass function from a large sample of clusters is indeed necessary to link the mass to these quantities which are easily observable.

In this context are very useful the so-called "*self-similar scaling relations*," derived naturally by considering that the cosmological structures originate from scale-free density perturbations and that the thermodynamical properties of the ICM are determined by scale-free gravity only [23]. Under these assumptions, galaxy clusters of different masses may be considered as a scaled version of each other. The density of each dark matter halo, ρ_{DM}, is proportional to the critical density of the Universe at the cluster's redshift, $\rho_{c,z}$, through the so-called "*overdensity*" $\Delta = \rho_{\mathrm{DM}}/\rho_{c,z}$, where $\rho_{c,z} = 3H_z^2/8\pi G$, and the expression for the Hubble constant at redshift z in a flat ΛCDM Universe is (e.g., [24]): $H_z = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \equiv H_0 E(z)$. Therefore, all clusters should have the same properties when rescaled by Δ.

If we define the mass M as the mass M_Δ inside the radius R_Δ at a given overdensity Δ, we obtain: $M_\Delta \propto \rho_{c,z} \cdot \Delta \cdot R_\Delta^3 \propto \rho_{c,0} \cdot E(z)^2 \cdot \Delta \cdot R_\Delta^3$, and thus we get to the M-R relation in the form

$$R \propto M^{1/3} \cdot E(z)^{-2/3}. \tag{1}$$

During cluster formation, the gravitational collapse of the diffuse gas in the potential well of the dark matter halo heats the gas itself at the virial temperature of the potential well that confines it: $T_{\mathrm{vir}} \sim GM\mu m_p/kR \sim 10^8\,\mathrm{K}$, where M

is the total mass, k is the Boltzmann constant, $\mu \sim 0.6$ is the mean molecular weight and R is the virial radius. The gas is thus heated to X-ray emitting temperatures and becomes a plasma in hydrostatic equilibrium whose emissivity is proportional to the square of its density (see Section 4). The virial temperature of an isothermal sphere of mass M is $kT \propto M/R \propto M^{2/3} \cdot E(z)^{2/3}$, leading to the M-T relation in the form:

$$M \propto T^{3/2} \cdot E(z)^{-1}. \quad (2)$$

From these relations it is possible to derive the relation between temperature and luminosity emitted by the hot gas through bremsstrahlung emission: $L_X \propto \rho^2 \cdot \Lambda \cdot V$, where ρ is the average gas density and Λ is the cooling function, that in the bremsstrahlung regime is $\propto T^{1/2}$ (see Section 4). Assuming that the gas distribution traces the dark matter distribution, $\rho \propto \rho_{\mathrm{DM}} \propto \rho_{c,z}$, we obtain: $L_X \propto \rho \cdot T^{1/2} \cdot M \propto \rho_o \cdot E(z)^2 \cdot T^{1/2} \cdot M \propto E(z)^2 \cdot T^{1/2} \cdot T^{3/2} \cdot E(z)^{-1}$, thus deriving the L-T relation in the form:

$$L_X \propto T^2 \cdot E(z). \quad (3)$$

By combining the M-T relation (2) with the L-T relation (3), we can finally derive the M-L relation that links the mass directly to the observable luminosity: $L_X \propto [M^{2/3} E(z)^{2/3}]^2 E(z)$, finding

$$M \propto L_X^{3/4} \cdot E(z)^{-7/4}. \quad (4)$$

In principle, once calibrated with simulations and/or observations, these scaling relations provide a method to link the mass of clusters to observables under the assumption that the process of structure formation is guided by gravity alone. On the other hand, deviations from these relations testify the presence of physical processes more complex than gravitational dynamics only, which modify the thermodynamical properties of the diffuse baryons and therefore the relations between observables and cluster masses. In particular, a number of observational measurements seems to indicate that the L-T relation is steeper than that predicted by self-similar models and is in the form $L \propto T^{2.5-3}$ (e.g., [25–30], see also Section 5.6). This observed breaking of the scaling relation has been ascribed to the presence of some excess entropy in the gas due to primordial nongravitational heating before the cluster virialization [31, 32] and is one of the strongest evidence for nongravitational processes acting in the ICM.

The main source of uncertainty in the determination of cosmological parameters from studies of cluster samples arises then from the uncertainty in the normalization, shape, and evolution of the relationships that relate the cluster masses to the observables. In order to understand better such relations, it is essential to investigate how the structure formation and AGN feedback affect the evolution of what we can observe, that is, the baryons in clusters.

3. Role of AGN Feedback in Galaxy Evolution

One of the main problems of the current cosmological model is why so few baryons have formed stars [33, 34]. Numerical simulations of cosmological structure formation that include the hydrodynamics of baryons and the radiative cooling processes predict that $\gtrsim$20% of the baryons should have condensed into galaxies, but only $\lesssim$10% have been observed in the form of stars (e.g., [35]). In particular, simulations that include only gravitational heating predict an excessive cooling of baryons that results in a population of galaxies which are too massive and too bright with respect to the ones observed, thus failing to reproduce the truncation of the high-luminosity end of the galaxy luminosity function [11, 36].

Instead of residing in the cD galaxies as predicted by simulations, most baryons are observed in the hot ICM. This problem may find a solution in the nongravitational heating supplied by supernovae (SN) and active galactic nuclei (AGN). Supernovae are essential in the process of enrichment of the ICM to the metallicity level observed [37, 38], and from the heavy-element abundances in clusters it is estimated that during a cluster history they supply a total amount of energy of the order of 0.3–1 keV/particle (e.g., [39, 40]). This is not enough energy to quench cooling in massive galaxies [38], as the energy input required to explain the mass-observable relations is ∼1-2 keV/particle [41–43]. Energetically, AGN heating appears to be the most likely mechanism to severely reduce the supply of gas from the hot ICM in massive galaxies and to explain the observed entropy profiles [11, 18, 44–46]. AGNs are powered by accretion of material onto a black hole (BH), which is located at the center of each stellar spheroid (both bulges within spirals and ellipticals). Matter falling onto a black hole releases an energy of the order of $E_{\mathrm{BH}} = \epsilon M c^2$, where $\epsilon \approx 0.1$ is the efficiency. For supermassive black holes (SMBHs) of masses $\sim 10^9 M_\odot$, the amount of energy released during their formation and growth is of the order of $E_{\mathrm{BH}} \sim 2 \times 10^{62}\,\mathrm{erg\,s^{-1}}$. Even a tiny fraction ($\lesssim$1%) of the energy released within the bulge could heat and blow away its entire gas content in small systems and prevent cooling, thus explaining the lack of star formation in bulges. An extraordinary discovery obtained recently in astrophysics is the correlation between the mass of the central black hole (M_{BH}) and the velocity dispersion (σ) of the galaxy's bulge used to estimate the mass of the bulge itself [10]. This "Magorrian relation" M_{BH}-σ suggests that the large-scale properties of the galaxy and the small-scale properties of the black hole are related. In particular, each massive galaxy seems to host a central black hole, whose mass is ∼0.1–0.2% of the bulge stellar mass [47–49]. Such a correlation may arise from the fact that the central black hole is able to regulate the amount of gas available for star formation in the galaxy. The formation of black holes and the formation of bulges are closely linked. Therefore supermassive black holes can have a profound influence on the formation and evolution of galaxies. The physical process regulating these phenomena has been called *"feedback,"* and the understanding of how it acts in detail is one of the main open issues in extragalactic astrophysics.

Clusters of galaxies are the only locations in the Universe where we can find an almost complete census of the intergalactic baryons and a very good description of their thermodynamical status and of their enrichment in heavy elements.

Therefore, X-ray observations of the ICM can provide us with new important insights into the processes of cooling and feedback which regulate galaxy formation.

4. Clusters of Galaxies in X-Rays and Thermal ICM

Clusters of galaxies are the largest virialized structures in the Universe, with typical sizes of ∼2–4 Mpc and total gravitational masses of $\sim 10^{14}$–$10^{15} M_\odot$. They are luminous X-ray sources, with typical luminosities ranging from a few $\times 10^{43}$–10^{46} erg s^{-1}. As first suggested by Felten et al. [50], the X-rays from clusters are primarily thermal bremsstrahlung emission from the diffuse ICM which fills the deep potential wells and is heated to temperatures of $\sim 10^8$ K (where $kT = 1$ keV for $T = 1.16 \times 10^7$ K) during the process of cluster formation.

4.1. Physical Properties of Hot Diffuse Plasma. The simple assumptions which are generally made in the study of the ionization state and X-ray line and continuum emission from a low-density hot plasma are briefly reviewed below [15].

(1) The time scale for elastic Coulomb collisions between particles in the plasma is much shorter than the age or cooling time of the plasma, therefore the free particles are assumed to have a Maxwell-Boltzmann distribution at the temperature T. This follows from considerations on the mean free paths of particles in a plasma without a magnetic field. The mean free path λ_e for an electron to suffer an energy exchange with another electron via Coulomb collisions is given by [51]:

$$\lambda_e = \frac{3^{3/2}(kT_e)^2}{4\pi^{1/2} n_e e^4 \ln\Lambda}, \tag{5}$$

where T_e is the electron temperature, n_e is the electron number density, and Λ is the ratio of largest to smallest impact parameters for the collisions ($\ln\Lambda \approx 38$). Equation (5) assumes that the electrons have a Maxwellian velocity distribution at the electron temperature T_e. However, it can be demonstrated that if a homogeneous plasma is created in a state in which the particle distribution is non-Maxwellian, elastic collisions will cause it to relax to a Maxwellian distribution on a time scale determined by the mean free paths [51, 52]. Electrons will achieve this equilibrium on a time scale given roughly by $t_{\rm eq}(e,e) \equiv \lambda_e/\langle v_e\rangle_{\rm rms}$, where $\langle v_e\rangle_{\rm rms}$ is the rms electron velocity $= (3kT_e/m_e)^{1/2}$:

$$t_{\rm eq}(e,e) \approx 3.3\times 10^5\,{\rm yr}\left(\frac{T_e}{10^8\,{\rm K}}\right)^{3/2}\left(\frac{n_e}{10^{-3}\,{\rm cm}^{-3}}\right)^{-1}. \tag{6}$$

The time scale for Coulomb collisions between protons to bring them into kinetic equilibrium is about $t_{\rm eq}(p,p) \approx (m_p/m_e)^{1/2} t_{\rm eq}(e,e)$, roughly 43 times longer than that for electrons. After this time, the electrons and ions (generally assumed to be protons) would each have Maxwellian distribution, but generally at different temperatures, respectively, T_e and T_i. The time scale for the electrons and ions to reach equipartition $T_e = T_i$ is $t_{\rm eq}(p,e) \approx (m_p/m_e) t_{\rm eq}(e,e)$, and for typical values of the ICM temperature and density is $t_{\rm eq}(p,e) \lesssim 6\times 10^8$ yr. Since this is shorter than the age of the clusters or their large-scale cooling time (although it is comparable to or longer than the cooling time in the cores of clusters, see (21) below), the intracluster plasma can generally be characterized by a single kinetic temperature $T = T_e = T_i$, which determines the rates of all excitation and ionization processes. It is important to note that the mean free paths, that is,

$$\lambda_e = \lambda_i \approx 23\,{\rm kpc}\left(\frac{T_e}{10^8\,{\rm K}}\right)^2\left(\frac{n_e}{10^{-3}\,{\rm cm}^{-3}}\right)^{-1}, \tag{7}$$

are generally much shorter than the length scales of interest in clusters (≈1 Mpc), and therefore the ICM can be treated as a collisional fluid, satisfying the hydrodynamic equations.

(2) At these low densities, collisional excitation and deexcitation processes are much slower than radiative decays, therefore any ionization or excitation process is assumed to be initiated from the ground state of an ion. Three (or more) body collisional processes are ignored because of the low density.

(3) Stimulated radiative transitions are not important, since the radiation field in the ICM is sufficiently dilute.

(4) At these low densities, the gas is optically thin and the transport of the radiation field can therefore be ignored.

Under these conditions, ionization and emission result primarily from collisions of ions with electrons, and collisions with other ions can be ignored. The time scales for ionization and recombination are generally considerably shorter than the age of the cluster or any relevant hydrodynamic time scale, therefore the plasma is assumed to be in ionization equilibrium (e.g., [53]). The equilibrium ionization state of a diffuse plasma depends only on the electron temperature: since in nearly all astrophysical plasmas most of the electrons originate in hydrogen and helium atoms, and these are fully ionized under the conditions considered here, the ICM is generally treated as a fully ionized plasma.

By indicating with X, Y, Z the mass fraction of hydrogen, helium, and heavier elements, respectively, the corresponding atom number densities can be written in the form: $n_{\rm H} = n_p \equiv \rho X/m_p$, $n_{\rm He} = \rho Y/4m_p = n_p Y/4X$, $n_z = \rho Z/Am_p = n_p Z/AX$, where ρ is the gas density, m_p the proton mass and A the mean atomic mass number (i.e., the number of nucleons) of heavier elements. Assuming that the gas pressure $p = nkT$ is contributed only by electrons and protons, thus neglecting nuclei ($n = n_e + n_p$), it is possible to derive the electron density n_e in terms of the proton density

n_p. From the expression of the number of particles contributing to the pressure, $n = 2n_{\rm H} + 2n_{\rm He} + (1/2)An_z$, one obtains

$$n = \left(2 + \frac{1}{2}\frac{Y}{X} + \frac{1}{2}\frac{Z}{X}\right)n_p, \tag{8}$$

which for solar abundances ($X = 0.71$, $Y = 0.265$, $Z = 0.025$) leads to $n_e \sim 1.2n_p$. It is also possible to calculate the mean molecular weight in amu, μ, such that the total number density of particles (electrons, protons and ions) is $\overline{n} = \rho/\mu m_p$. From the expression $\overline{n} = 2n_{\rm H} + 3n_{\rm He} + ((1/2)A + 1)n_z$, in the approximation $A \gg 1$, one obtains

$$\mu = \left(2X + \frac{3}{4}Y + \frac{1}{2}Z\right)^{-1}, \tag{9}$$

which for solar abundances leads to $\mu \sim 0.6$.

4.2. X-Ray Emission from the ICM. The X-ray continuum emission from a hot diffuse plasma, such as the ICM, is due primarily to two processes: thermal bremsstrahlung (free-free emission) and recombination (free-bound) emission. Processes that contribute to X-ray line emission (bound-bound radiation) from a diffuse plasma include collisional excitation of valence or inner shell electrons, radiative and dielectric recombination, inner shell collisional ionization, and radiative cascades following any of these processes.

At the high temperatures, typical of clusters (in particular at $kT \gtrsim 2.5\,\text{keV}$), thermal bremsstrahlung is the predominant X-ray emission process. The bremsstrahlung emissivity at a frequency ν (defined as the emitted energy per unit time, frequency, and volume) of a plasma with temperature T, electron density n_e and ion density n_i is given by (e.g., [54]):

$$J_{\rm br}(\nu, T) = 6.8 \times 10^{-38} Z^2 n_e n_{\rm i} T^{-1/2} e^{-h\nu/kT} \overline{g}(\nu, T), \tag{10}$$

where the Gaunt factor $\overline{g}(T)$, which corrects for quantum mechanical effects and for the effect of distant collisions, is a slowly varying function of the parameters [55, 56]. If the ICM is mainly at a single temperature, then (10) indicates that the X-ray spectrum should be close to an exponential of the frequency, as is generally observed.

The total power per unit volume emitted by thermal bremsstrahlung is obtained by integrating (10) over frequency, obtaining:

$$J_{\rm br}(T) = 1.4 \times 10^{-27} n_e n_i T^{1/2} Z^2 \overline{g}(T), \tag{11}$$

where $\overline{g}(T)$ is a frequency average of $\overline{g}(\nu, T)$ and is in the range 1.1 to 1.5 (choosing a value of 1.2 will give an accuracy in the estimate of $J_{\rm br}(T)$ to within about 20%, [54]). For solar abundances, the emission is primarily from hydrogen and helium.

Compilations of the different emissivities for X-ray lines and continua can be found in the literature (e.g., [57, 58]). Early detailed calculations of the X-ray spectra predicted by different models of the ICM have been given by Sarazin and Bahcall [59] and Bahcall and Sarazin [60, 61]. In these models most of the X-ray emission is thermal bremsstrahlung continuum, and the strongest lines (highest equivalent width) are in the 7 keV iron line complex.

The frequency-integrated total emissivity at a temperature T can be written as:

$$J_X(T) = \Lambda(T) n_e n_p \,\text{erg}\,\text{s}^{-1}\,\text{cm}^{-3}, \tag{12}$$

where $\Lambda(T)$ is the *cooling function*, which depends upon the mechanism of the emission and can be represented as:

$$\Lambda(T) = lT^{\alpha}, \tag{13}$$

where $-0.6 \lesssim \alpha \lesssim 0.55$; for thermal bremsstrahlung it is $l \sim 2.5\times10^{-27}$ and $\alpha = 1/2$ [62]. The general behaviour of the cooling function was calculated and discussed by Sutherland and Dopita [63].

The projection on the sky of the plasma emissivity gives the X-ray surface brightness: in order to constrain the physical parameters of the ICM, the observed surface brightness can be either geometrically deprojected or, more simply, fitted with a model obtained from an assumed distribution of the gas density.

4.3. Hydrostatic Models for ICM Distribution. From the expression for the sound speed $c_s^2 = \gamma kT/\mu m_p$, where $\gamma = 5/3$ for a monatomic gas, the time required for a sound wave in the ICM to cross a cluster is given by:

$$t_s \approx 6.6 \times 10^8 \left(\frac{T}{10^8\,\text{K}}\right)^{-1/2} \left(\frac{D}{1\,\text{Mpc}}\right)\text{yr}, \tag{14}$$

where D is the cluster diameter. Since this time is short compared to the likely age of a cluster (in first approximation assumed to be $\sim 10^{10}$ yr), the gas is generally thought to be in hydrostatic equilibrium in the gravitational potential of the cluster: $\nabla p = -\rho\nabla\phi$, where $p = \rho kT/\mu m_p$ is the gas pressure, ρ is the gas density, and ϕ is the gravitational potential of the cluster. Under the assumptions that the ICM is locally homogeneous and the cluster is spherically symmetric, the hydrostatic equilibrium equation reduces to

$$\frac{1}{\rho}\frac{dp}{dr} = -\frac{d\phi}{dr} = -\frac{GM(r)}{r^2}, \tag{15}$$

where r is the radius from the cluster center and $M(r)$ is the total cluster mass within r. If the gas self-gravity is ignored, then the distribution of the ICM is determined by the cluster potential $\phi(r)$ and the temperature distribution of the gas $T(r)$, and (15) is a linear equation for the logarithm of the gas density. Under these assumptions, the gravitational mass $M_{\rm tot}$ of a galaxy cluster can be written as:

$$M_{\rm tot}(< r) = -\frac{kTr}{G\mu m_p}\left[\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right]. \tag{16}$$

This expression is commonly used to estimate the gravitational mass of galaxy clusters and groups from X-ray observations, through the measurements of the radial profiles of temperature and density (e.g., [64–67]). However, we note that (16) neglects the contribution of possible additional, nonthermal pressure that, if present, should be included in the estimate of the total mass. In particular, recent results

from numerical simulations indicate that the total mass of simulated clusters estimated through the X-ray approach is lower that the true one due to gas bulk motions (i.e., deviation from the hydrostatic equilibrium) and the complex thermal structure of the gas [68, 69]. Possible observational biases in the derivation of X-ray masses are also discussed in Piffaretti and Valdarnini [70].

4.3.1. The β-Model. Cavaliere and Fusco-Femiano [71] studied the X-ray emission by the hot plasma in galaxy clusters and developed a hydrostatic model based on the assumption that the gas and the galaxies are in equilibrium in the same gravitational potential ϕ (see, e.g., [15, 72] for a recent commentary on this model). By further assuming that the galaxy distribution is well described by King's approximation to the isothermal sphere [73], the expression for the ICM distribution may be written as [71]:

$$\rho(r) = \rho_0 \left[1 + \left(\frac{r}{r_{\rm core}}\right)^2\right]^{-(3/2)\beta}, \tag{17}$$

and the surface brightness profile observed at a projected radius b, $I(b)$, is in the form [71]:

$$I(b) = I_0 \left[1 + \left(\frac{b}{r_{\rm core}}\right)^2\right]^{1/2-3\beta}. \tag{18}$$

The parameter β is defined as

$$\beta = \frac{{\sigma_r}^2}{kT/\mu m_p}, \tag{19}$$

where σ_r is the line-of-sight velocity dispersion and represents the ratio of specific kinetic energies of galaxies and gas.

This self-consistent isothermal model, called the "β-model," is widely used in the X-ray astronomy to parametrize the gas density profile in clusters of galaxies by fitting their surface brightness profile. One of the advantages of using a β-model to parameterize the observed X-ray surface brightness is that the total mass profiles can be recovered analytically and expressed by a simple formula:

$$M_{\rm tot}(<r) = \frac{kr^2}{G\mu m_p}\left[\frac{3\beta r T}{r^2 + r_c^2} - \frac{dT}{dr}\right]. \tag{20}$$

Equation (17) states that the gas density rises towards the cluster center. Since the bremsstrahlung and line emission depend on the square of the gas density (12), in the central regions of clusters the loss of energy by X-ray emission represents an important process for the thermal particles in the ICM. In particular, if the gas density reaches high enough values, large amounts of gas cool and flow into the centers of clusters, forming the so-called *cooling flows*. In cooling flow clusters, the single β-model is found to be a poor description of the entire surface brightness profile: a fit to the outer regions shows a strong excess in the center as compared to the model (see Section 4.4). Conversely, the centrally peaked emission is a strong indication of a cooling flow in relaxed cluster.

4.4. Cool Cores and Cooling Flow Problem. The X-rays emitted from clusters of galaxies represent a loss of energy of the ICM. The resultant cooling time is calculated as the time taken for the gas to radiate its enthalpy per unit volume H_v using the instantaneous cooling rate at any temperature:

$$t_{\rm cool} \approx \frac{H_v}{n_e n_{\rm H} \Lambda(T)} = \frac{\gamma}{\gamma - 1}\frac{kT}{\mu X n_e \Lambda(T)}, \tag{21}$$

where $\gamma = 5/3$ is the adiabatic index; $\mu \approx 0.6$ is the molecular weight; $X \approx 0.71$ is the hydrogen mass fraction; $\Lambda(T)$ is the cooling function. In the central region, the cooling rate of the ICM is sufficiently high that the particles lose their energy via radiation, as inferred from X-ray images of the cores of many clusters which show strongly peaked surface brightness distributions. The density of the gas then rises to maintain the pressure required to support the weight of the overlying gas in the rest of the cluster, causing a slow subsonic inflow of material towards the cluster center. This qualitative picture describes the physics of the process known as a *cooling flow* (see [2] for a review of the standard model, and [74] for a quantitative description of the evolution of cooling flows). The steady cooling flow is confined within the region in which $t_{\rm cool}$ is less than the time for which the system has been relaxed (thus allowing the establishment of a cooling flow). This *cooling region* is delimited by the so-called *cooling radius* $r_{\rm cool}$, which is usually defined as the radius at which $t_{\rm cool}$ is equal to the look-back time to $z = 1$, that is, $\sim 7.7 \times 10^9$ yr in the concordance cosmology. The fraction of clusters with a central surface excess with respect to a β-model, the so-called *cool cores*, is large. Cool core clusters are about 90% of X-ray-selected clusters with total mass $M_{\rm tot} \lesssim 10^{14} M_\odot$, and about 50% of X-ray-selected clusters with total mass $M_{\rm tot} \gtrsim 10^{14} M_\odot$ [75]. Cool cores are also characterized by strong enhancements in the central abundance (e.g., [76, 77]) and declining temperature profiles toward the central region (e.g., [78–80]).

In the standard model, the "magnitude" of a cooling flow is measured from the amount of matter which crosses $r_{\rm cool}$ and flows towards the center, that is, $\dot{M}$, the *mass inflow rate*. The mass inflow rate, due to cooling, can be estimated from the X-ray images by using the luminosity $L_{\rm cool}$ associated with the cooling region and assuming that it is all due to the radiation of the total thermal energy of the gas plus the pdV work done on the gas as it enters r_c: $L_{\rm cool} = dE/dt$, where $dE = dE_{\rm th} + pdV = (\gamma/\gamma - 1)pdV$, and $pdV = (\rho kTdV)/(\mu m_p) = (dMkT)/(\mu m_p)$, with $\gamma = 5/3$. By substituting one obtains the expression for $L_{\rm cool}$:

$$L_{\rm cool} = \frac{5}{2}\frac{\dot{M}}{\mu m_p} kT, \tag{22}$$

where T is the temperature of the gas at $r_{\rm cool}$. $L_{\rm cool}$ ranges from $\sim 10^{42}$ to $> 10^{44}$ erg s^{-1} and generally represents $\sim 10\%$ of the total cluster luminosity [2]. Value of $\dot{M} \sim 100 M_\odot$ yr^{-1} are fairly typical for cluster cooling flows.

However, the current generation of X-ray satellites, *Chandra* and *XMM-Newton*, has radically changed our understanding of cooling flow systems. Albeit confirming the existence of short cooling times, high densities and low

temperatures in the cluster cores, the arrival of high-resolution X-ray spectral data has shown the absence or weakness of the soft X-ray line Fe XVII, indicating that the amount of gas cooling radiatively below about one-third of its original temperature is ten times less than expected (e.g., [3, 4, 81]). The lack of evidence for central gas cooling to very low temperatures at the rates predicted in the hot atmospheres of galaxy clusters and groups represents an open question which is often referred to as the so called "*cooling flow problem*" (see [5, 6, 82] for reviews).

Historically, two main approaches were adopted to solve this problem. As the gas radiates but does not appear to cool, either the normal signatures of radiative cooling below 1-2 keV are somehow suppressed, or there must be an energy-injection mechanism into the ICM which compensates cooling. Different possibilities considered in the former hypothesis include absorption [3, 83], inhomogeneous metallicity [83, 84], and the emerging of the missing X-ray luminosity in other bands, like ultraviolet, optical and infrared due to mixing with cooler gas/dust [83, 85, 86]. Proposed heating mechanisms in the context of the latter approach include, for example, processes associated with relativistic AGN outflows [87–93], electron thermal conduction from the outer regions of clusters [94–97], continuous subcluster merging [98], contribution of the gravitational potential of the cluster core [99], and feedback from intracluster supernovae [100]. Among all these, feedback by the central AGN appears to be the most promising solution.

5. X-Ray Cavities and Shocks: AGN Feedback in Action

It was already known in the early 90s that central dominant (cD) galaxies of cool core clusters have a high incidence of radio activity, showing the presence of central FR-I radio galaxies in 70% of the cases [7, 101, 102]. Their behaviour differs from that of quasar: in many low-accretion-rate AGNs almost all the released energy is channeled into jets because the density of the gas surrounding the black hole is not high enough for an efficient radiation (e.g., [103]). In fact, the importance of these objects has been underestimated for a long time due to their poor optical luminosity. The importance of radio galaxies in cool cores began to emerge after the discovery, with the X-ray satellite *ROSAT*, of deficits in the X-ray emission of the Perseus and Cygnus A clusters which are spatially coincident with regions of enhanced synchrotron emission [104, 105]. With the advent of the new high-resolution X-ray observations performed with the current generations of X-ray telescopes, *Chandra* and *XMM-Newton*, it became clear that the central radio sources have a profound, persistent effect on the ICM. In particular, *Chandra* images, which are obtained at the superb angular resolution of $\sim 1''$, showed that the Perseus and Cygnus A clusters are not isolated cases—indeed the central hot gas in many cool core systems is not smoothly distributed, but is instead cavitated on scales often approximately coincident with lobes of extended radio emission. These observations also reveal highly disturbed structures in the cores of many clusters, including shocks, ripples, and sharp density discontinuities. The comparison with radio images obtained at similar angular resolution has revealed that AGN jets are the cause of many of these disturbances. The most typical configuration is for jets from the central dominant elliptical of a cluster to extend outwards in a bipolar flow, inflating lobes of radio-emitting plasma. These lobes push aside the X-ray emitting gas of the cluster atmosphere, thus excavating depressions in the ICM which are detectable as apparent "cavities" in the X-ray images. The brightness depressions observed in X-rays, which are mostly found in pairs, are ~20–40% below the level of the surrounding gas, consistently with the expected decrement along an empty bubble immersed in a β-model atmosphere [106–110]. The cavities are often surrounded by bright shells, or rims, which are typically found to be cooler than the ambient medium [106–109, 111]. This is likely due to the compression of the central, low-entropy gas into the bright shell during the cavity rising and expansion into the hot atmosphere.

X-ray cavities are present in $\gtrsim$70% of cool-core clusters [112], but this fraction could be underestimated due to the limitation of cavity detectability [113]. Identifying radio galaxies as a primary source of feedback in the hot atmospheres of galaxy clusters and groups has been one of the major achievements of the current generation of X-ray observatories (for a comprehensive review see [6] and references therein). Well-studied examples of cavity clusters are Perseus (e.g., [104, 106, 114, 115]), Hydra A (e.g., [4, 107, 116–120]), M 87 (e.g., [89, 121–125]), A 2052 (e.g., [108, 126–128]), RBS 797 (e.g., [111, 129–131]), A 133 (e.g., [132, 133]), A 262 (e.g., [126, 134]), and MS 0735 + 7421 (e.g., [135, 136]). In-depth analyses of individual objects, which are now numerous in the literature, combined with studies of cavity samples [8, 112, 137–140] have enabled us to identify the global properties which are common among the cavities, thus shedding light of the feedback mechanism. The emerging picture is that bipolar outflows emanating from the BCG core inflate large bubbles while driving weak shocks, heat the ICM and induce a circulation of gas and metals on scales of several 100 s kpc. Weak shocks have been observed as ripple-like features in the ICM in the deepest X-ray images of Perseus and A 2052 [115, 128].

However, the differences between groups and clusters imply that the existing studies on cavities in clusters tell us little about how feedback operates in groups. With respect to rich galaxy clusters, the observation of cavities in galaxy groups and ellipticals is complicated by the lower X-ray surface brightness, which limit their detection in shallow X-ray images. On the other hand, there are several examples of AGN-ICM interaction just a few tens Mpc away from us, which allow us to probe regions closer to the central black hole. In particular, low mass systems with cavities which now have deep *Chandra* images are M 84 [141, 142], NGC 4636 [143–145], NGC 5044 [146–149], HCG 62 [150–153], and NGC 5846 [154]. Performing in-depth individual studies and sample studies of the lower-energy outbursts in these smaller systems is of major interest because the relationship between AGNs and hot gas can significantly influence galaxy evolution in the group environment, which

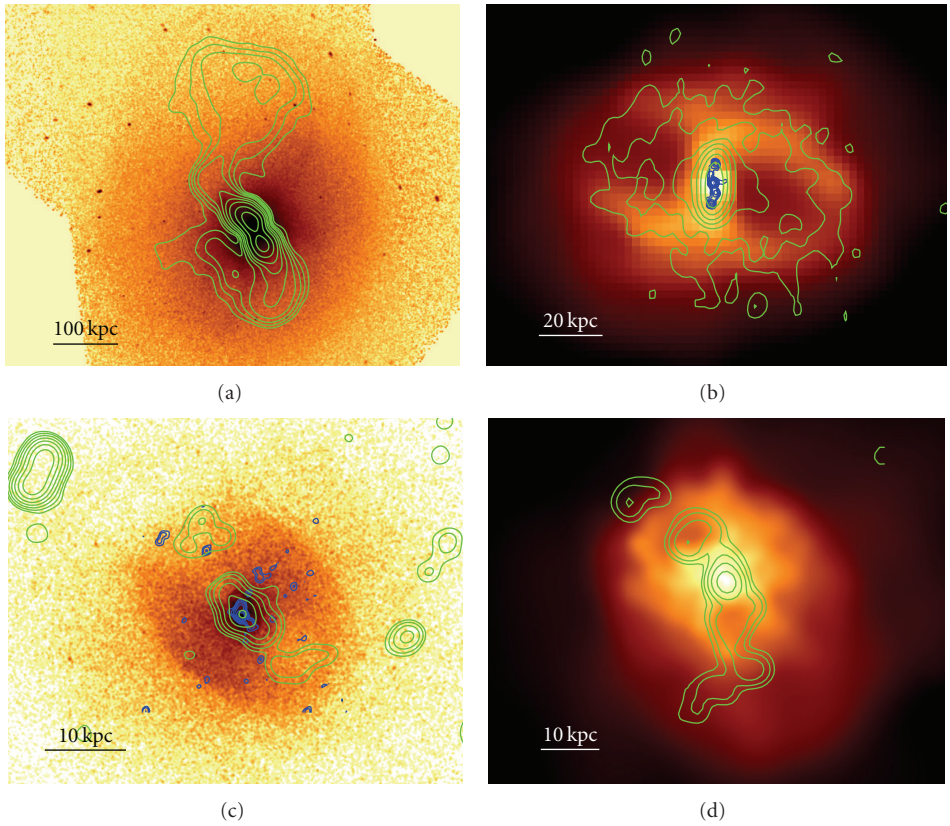


FIGURE 1: (a) The green contours outlining the 330 MHz radio emission from Lane et al. [164] are overlaid onto the 0.5–7.5 keV *Chandra* image of the galaxy cluster **Hydra A** ($z = 0.0538$). The extended radio lobes fill a large-scale system of X-ray cavities and are surrounded by a "cocoon" shock. See also Sections 5.6 and 5.8. adapted from Nulsen et al. [116]. (b) Very Large Array (VLA) radio contours overlaid onto the 0.5–7.0 *Chandra* image of the galaxy cluster **RBS 797** ($z = 0.35$). The subarsec resolution radio image shows the details of the innermost 4.8 GHz radio jets (blue contours), which clearly point in a north-south direction. Remarkably, these inner jets are almost perpendicular to the axis of the 1.4 GHz emission observed at 1″ resolution (green contours), which is elongated in the northeast-southwest direction filling the X-ray cavities. Adapted from Gitti et al. [130]. (c) 0.3–2 keV *Chandra* image of the galaxy group **NGC 5813** ($z = 0.0066$) with 1.4 GHz VLA (blue) and 235 MHz Giant Metrewave Radio Telescope (green) radio contours overlaid. The image shows two pairs of cavities, plus an outer cavity to the northeast, two sharp edges to the northwest and southeast, and bright rims around the pair of inner cavities. Adapted from Randall et al. [157]. (d) 235 MHz GMRT contours overlaid on the smoothed 0.5–2.0 keV *Chandra* image of the compact group **HCG 62** ($z = 0.0137$). The radio source shows a disturbed morphology with inner lobes clearly filling the well-defined X-ray cavities, but with outer lobes having no associated X-ray cavities (see also Section 5.4). Adapted from Gitti et al. [153].

is the locus of the majority of galaxies in the Universe [155]. Due to the shallower group potential, the AGN outburst can have a large impact on the intragroup medium in terms of altering the thermal history and spatial distribution of the intragroup medium, as the mechanical output by radio AGN is of the same order of magnitude as the binding energy of groups [156]. Such investigations have been undertaken only recently for individual objects (e.g., NGC 5813: [157], AWM 4: [158, 159]) and for group samples [9, 160–162], and are rapidly improving our understanding of these systems. However, this observational effort is still awaiting detailed theoretical work in order to corroborate the observational findings. Recent detailed simulations indicate that groups are not simply a scaled-down version of clusters, as there may be remarkable differences between how AGN feedback operates in galaxy group and in galaxy cluster environments [163]. In particular, AGN heating in groups seems to act through persistent, gentle injection of mechanical energy. On the other hand, in clusters there must be also the action of rare, powerful outbursts [163], although more extensive theoretical work is required to reach firm conclusions.

Examples of well-studied cavity systems in clusters and groups are shown in Figure 1.

5.1. Cavity Heating. The heating is thought to occur through the dissipation of the cavity enthalpy and through shocks driven by the AGN outburst. The energy required to create a cavity with pressure p and volume V is the sum of the pV work done by the jet to displace the X-ray emitting gas while it inflates the radio bubble, and the internal energy of the lobes, that is, the enthalpy given by:

$$E_{\rm cav} \equiv H = E_{\rm int} + pV = \frac{\gamma}{\gamma - 1} pV, \tag{23}$$

where γ is the ratio of the specific heats of the cavity content. If the lobes are dominated by the magnetic field, by

nonrelativistic gas, or by relativistic plasma, H can vary from $2pV$ and $4pV$. In particular, typically it is assumed that the internal composition of the cavity is dominated by relativistic plasma, therefore $\gamma = 4/3$ and $H = 4pV$. The product of pressure and volume can be estimated directly by X-ray observations through measurements of the cavity size and of the temperature and density of the surrounding ICM. A potential issue is represented by the uncertainties in the measurement of the cavity volume. The cavity size is usually estimated through a visual inspection of the X-ray images. This method is therefore dependent on the quality of the X-ray data and also subject to the arbitrariness of the observer. The cavity size and geometry measured by different observers may vary significantly depending on the approach adopted, leading to differences between estimates of up to a factor of a few in pV (e.g., [131, 153, 165]).

Systematic observational studies of samples of X-ray cavities show that their enthalpies measured from (23) lie between $\sim 10^{55}$ erg (in ellipticals, groups, and poor clusters) and $\gtrsim 10^{61}$ erg (in rich clusters). On the other hand, simulations indicate that pV varies with time during the cavity evolution and may be an inaccurate measure of the total energy released [166, 167]. Cavity power estimates within a factor of two of the simulated values seem possible provided the inclination angle of the jets is known accurately [168]. Bearing this caveat in mind, when divided by the cavity age, $t_{\rm cav}$, the observational measurements give an estimate of the so-called "cavity power," $P_{\rm cav}$. Since shocks are very difficult to detect and are currently known only in a few systems (e.g., Hydra A [116], MS 0735 + 7421 [135], HCG 62 [153], NGC 5813 [157]), for consistency, the usual approach in sample studies is that of considering only the cavity power. $P_{\rm cav}$ thus provides a lower limit (and best-available gauge) to the true total mechanical power of the AGN, that is, the jet power: $P_{\rm jet} \gtrsim P_{\rm cav} = E_{\rm cav}/t_{\rm cav}$.

As proposed by Bîrzan et al. [137], the cavity age can be estimated in three ways: (1) by assuming that the cavity rises the hot gas atmosphere at the sound speed $c_s = \sqrt{\gamma kT/\mu m_p}$ in this case the cavity reaches the projected distance R from the cluster center in the sound crossing time $t_s = R/c_s = R/\sqrt{\gamma kT/\mu m_p}$; (2) by assuming that the cavity is buoyant and move outwards at the terminal velocity $v_t = \sqrt{2gV/SC}$, where $g = GM_{<R}/R^2$ is the gravitational acceleration at the cavity position R, V is the volume of the cavity, S is the cross-section of the cavity, and $C = 0.75$ is the drag coefficient [89]: in this case the cavity age is the buoyancy-time $t_{\rm buoy} \sim R/\sqrt{2gV/SC}$; (3) by considering the time required for gas to refill the displaced volume of the cavity as it rises: in this case the cavity age is estimated as $t_{\rm ref} \sim 2\sqrt{r/g}$, where r is the radius of the cavity. Typically, the age estimates agree to within a factor of 2, with the buoyancy times lying in between the sound crossing time and the refill times. Most sample studies adopt the buoyancy time, which for typical values gives cavity ages of the order of few 10^7 yr (e.g., [139]).

5.2. The Relationship between Jet Power and $L_{\rm cool}$. Once a cavity is detected, it is relatively simple to estimate its power from the measurements of $E_{\rm cav}$ and $t_{\rm cav}$ by applying (23). The cavity power, $P_{\rm cav}$, which is a measure of the energy injected into the hot gas by the AGN outburst, can then be compared directly with the gas luminosity inside the cooling radius, $L_{\rm cool}$, which represents the luminosity that must be compensated for by heating to prevent cooling. The gas luminosity inside the cooling radius is estimated as the bolometric X-ray luminosity derived from a deprojection spectral analysis. In Figure 2(a) is shown a quantitative comparison between $P_{\rm cav} = 4pV/t_{\rm buoy}$ and $L_{\rm cool}$ calculated for the extended sample discussed in O'Sullivan et al. [165], who combined new data of 9 galaxy groups with the cluster sample of Bîrzan et al. [8] and with the elliptical sample of Cavagnolo et al. [169].

This plot follows those presented in Figure 2 of Bîrzan et al. [137] and in Figure 6 of Rafferty et al. [139]. As already noted by these authors, it is evident that the cavity power scales in proportion to the cooling X-ray luminosity, although with a big scatter. In general, it appears that the high mass (corresponding to high $L_{\rm cool}$) systems need an average of $4pV$ per cavity to counter cooling. On the other hand, if we recalculate $P_{\rm cav}$ as $1pV/t_{\rm buoy}$ all the points in the plot will shift down by a factor 4, and only the lower mass systems will still lie around the line $P_{\rm cav} = L_{\rm cool}$. These systems require $1pV$ per cavity to offset cooling at the present time. A few low mass systems will even still be above the equality line, thus indicating that the total mechanical power of the AGN far exceeds the radiative losses and their atmospheres are being heated. Although this extended sample is not a complete sample and therefore is not representative of the whole population of clusters, groups, and ellipticals, it is interesting to consider the mean values of heating and cooling to see how they compare. We estimated a mean cooling power of $4.09 \times 10^{44}\,{\rm erg\,s^{-1}}$, and a mean cavity power of $6.18 \times 10^{44}\,{\rm erg\,s^{-1}}$. In order to quantify properly the contribution of AGN feedback, over the system lifetime, in the energetics of cooling flow, it is important to determine the "duty-cycle" of AGN. Many studies have attempted such calculation by adopting different approaches, for example, by considering the luminosity function of radio galaxies [170], the fraction of clusters that contain bubbles and cavities [112, 138, 171], the frequency of bubble required to produce sufficient heating [172], the prevalence of radio-loud AGN [101, 173]. In particular, by considering the cavities as tracers of the feedback mechanism, that is, by assuming that the feedback is active and efficient as long as the cavities are visible, we can correct the mean cavity power by the fraction of cool core clusters with cavities, estimated by Dunn and Fabian [112] of the order of at least 70%. The ratio of mean cavity power to cooling flow power is thus very close to unity. The mean values for the whole sample are only indicative and do not reflect the different behaviour of groups and ellipticals with respect to clusters. In particular, such mean powers miss the point that in order to counter cooling the low mass systems require outbursts with relatively less total energy, lower powers, and repeating more rapidly than high mass systems. This is supported by recent numerical simulations of galaxy groups which show that, in contrast to galaxy clusters, the AGN self-regulated feedback must act through a quasi-continuous gentle injection with subrelativistic outflows, rather

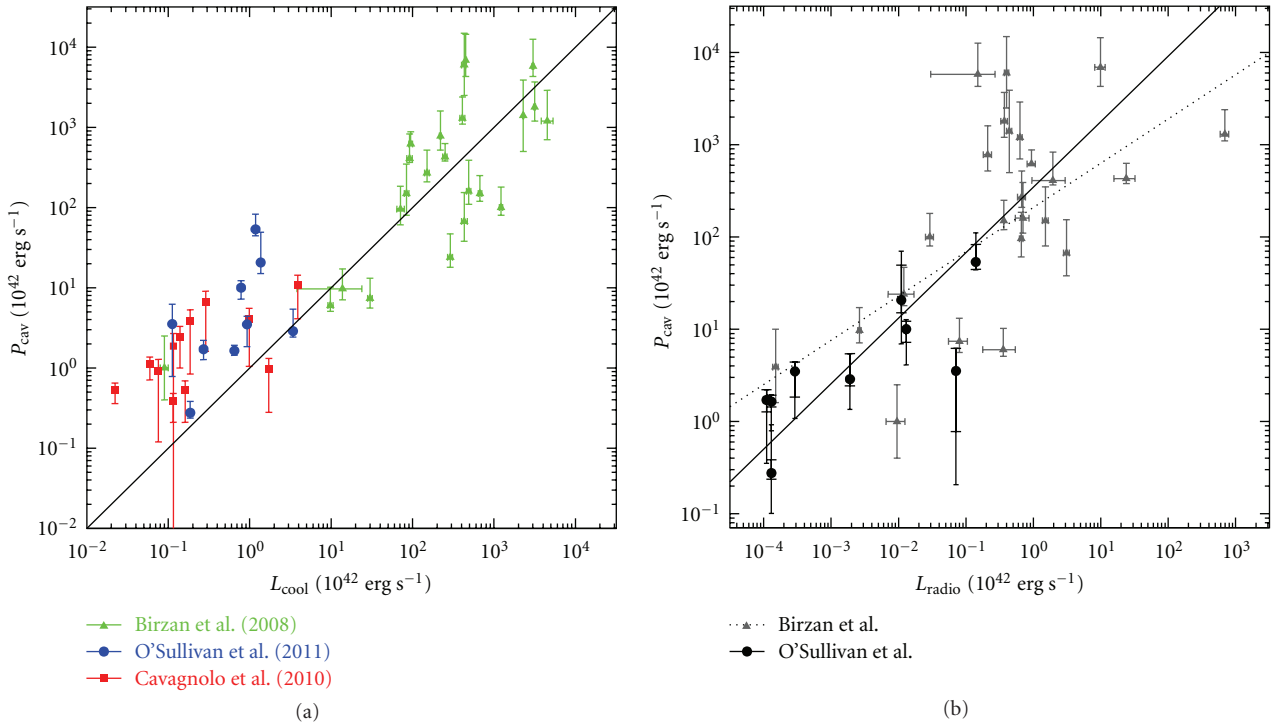


FIGURE 2: (a) Cavity power of the central AGN, P_{cav}, versus the X-ray luminosity of the ICM inside the cooling region, L_{cool}. The cavity power is calculated assuming $4pV$ of energy per cavity and the buoyancy timescale. Different symbols denote systems in different samples presented in the literature: *green triangles*: Bîrzan et al. [8], *red squares*: Cavagnolo et al. [169], *blue circles*: O'Sullivan et al. [165]. 1σ uncertainties on cavity power are indicated by error bars (see [165] for details). The diagonal line denotes $P_{cav} = L_{cool}$. Credit O'Sullivan (private communication). (b) Cavity power of the central AGN, P_{cav}, versus integrated 10 MHz–10 GHz radio power, L_{radio}, for the systems in the sample of Bîrzan et al. (grey triangles) and the groups in the sample of O'Sullivan et al. (black circles). The solid fit line indicates the regression fit to the data points calculated by O'Sullivan et al. [165]. The dotted line indicates the relation found by Bîrzan et al. [8]. Adapted from O'Sullivan et al. [165].

than through rare and powerful episodes [163]. An attempt to produce more meaningful averages could be that of dividing the sample plotted in Figure 2(a) in two subsets. In fact, by doing this we find that the ratio of mean cavity power to cooling power for the groups and ellipticals is 7.94 (samples of Cavagnolo et al. [169] and O'Sullivan et al. [165]), compared to a ratio of 1.51 calculated for clusters only (sample of Bîrzan et al. [8]). If the duty cycle of low mass systems is the same as (or not lower than) high mass systems [112, 173], the relative ratio of heating to cooling appears to be a factor $\gtrsim$5 higher in low mass systems. In other words, groups and ellipticals seem to have five times as much power available to counter cooling than rich clusters.

A study of a complete unbiased sample including both cool-core and non cool-core systems is necessary to derive definite constraints on the balance between heating and cooling. However, it seems plausible that the time-averaged AGN feedback balances radiation losses of the ICM. Therefore, the general picture emerging from the observed trend between X-ray luminosity and bubble mechanical luminosity, together with the existence of short central cooling time, is that the AGN is fueled by a cooling flow that is itself regulated by feedback from the AGN. The basic idea of this AGN-cooling flow scenario is that a self-regulated equilibrium may be achieved, in which the radiative losses from the thermal ICM are balanced by mechanical heating from the central AGN over the system lifetime. Although this scenario is no longer in doubt, it is still not clear how heating can act preserving at the same time the observed temperature gradient and the cool core (e.g., [174]).

5.3. The Relationship between Jet Power and Radio Power. Studies of cavity samples allow to derive the relationship between the mechanical power and radio emission of AGN jets and lobes. Such a relationship is of great interest because it helps to understand the physics of AGN jets (e.g., [165, 175]), and because it provides an estimate of the energy available from AGN based directly on the radio data (e.g., [101]), thus avoiding the problem of cavity detectability in shallow X-ray images. Bîrzan et al. [8] studied a sample dominated by galaxy clusters and derived the relation between cavity power and 327 MHz radio power, as well as between cavity power and the integrated 10 MHz–10 GHz radio luminosity, extending to lower frequencies their previous work [137]. While many of the observed X-ray cavities are filled with 1.4 GHz radio-emitting plasma, some are undetected at this high frequency and have been referred to as "ghost cavities." These may result from the aging of the relativistic particle population or may have been inflated by events which produced only particles of low

energy. Examinations of radio images at multiple (and low) frequencies is particularly important as the progressive loss of particle energy causes higher frequency emission to fade fastest and the spectral index to steepen, so that evidence of a former AGN activity may be reflected only at low frequencies. The lack of 1.4 GHz radio emission is observed more frequently in groups than in clusters, therefore low-frequency radio observations are crucial for galaxy groups.

Giacintucci et al. [9] selected a compilation of 18 galaxy groups, based on the presence of signs of interaction between the hot gas and the central AGN, and observed both by the Giant Metrewave Radio Telescope (GMRT) at frequencies ≤610 MHz and by *Chandra* and/or *XMM-Newton*. These authors found that nine of these groups have cavities clearly correlated with radio structures. By adding such systems to the Bîrzan et al. [8] sample, O'Sullivan et al. [165] examined the relations between jet mechanical power and radio power in a combined sample which includes the groups having the most reliable radio measurements currently available. In particular, the integrated 10 MHz–10 GHz radio luminosity estimated from the source spectral index is considered by these authors as a superior cavity power indicator compared to estimates at a single frequency, since it accounts for variations in spectral index between sources. Figure 2(b) shows the relationship between cavity power, $P_{\rm cav}$, and the integrated radio luminosity, $L_{\rm radio}$, for the combined sample. The best fitting power-law relationship is [165]

$$\log P_{\rm cav} = 0.71(\pm 0.11)\log L_{\rm radio} + 2.54(\pm 0.21), \qquad (24)$$

where $P_{\rm cav}$ and $L_{\rm radio}$ are in units of $10^{42}\,{\rm erg\,s^{-1}}$. See O'Sullivan et al. [165] for a detailed discussion of this relation.

5.4. Radio Lobe Composition. When the radio source is filling the cavities, it is possible to compare the radio pressure of the relativistic plasma internal to the lobes with the X-ray pressure of the surrounding thermal gas. The pressure of the hot gas is measured from the density and temperature derived from the X-ray data as $p \backsimeq 2n_e kT$. The total pressure in a radio lobe is the sum of the magnetic pressure, p_B, and the total particle pressure, $p_{\rm part}$, and can be written as:

$$p_{\rm radio} = p_B + p_{\rm part} = \frac{B^2}{8\pi} + \frac{1}{3}\frac{E_{\rm part}}{fV} = \frac{B^2}{8\pi} + \frac{1}{3}\frac{(1+k)E_e}{fV}, \qquad (25)$$

where B is the magnetic field, k is the ratio of the energy in protons to that in electrons (E_e), V is the volume of the radio lobe, and f is the volume filling factor of the relativistic plasma. Using the expression for E_e given in Pacholczyk [176], (25) determines the lobe pressure in terms of the magnetic field strength and the factor k/f, once the volume V of the radio lobe is known. This calculation is usually performed under the widely adopted minimum energy conditions, in which the relativistic plasma is in equipartition with the magnetic field ($B_{\rm eq}$). Further assumptions usually made in literature are $f = 1$ and $k = 0$ or $k = 1$. A volume filling factor of 1 indicates that the lobes are empty of thermal gas, which is a reasonable hypothesis when they are observed to be spatially coincident with X-ray cavities. The assumption $k = 1$ implies that half of the energy in particles is in the form of nonradiating particles, as in an electron-proton jet, whereas $k = 0$ would indicate an electron-positron jet. We stress that the uncertainties in the calculation of $B_{\rm eq}$ and $p_{\rm radio}$ come from the values of k and f, which are still largely unknown. Conversely, it is possible to constrain the ratio k/f by assuming pressure balance (see below).

For historical reasons the frequency band adopted to calculate the standard equipartition field is $\nu_1 = 10\,{\rm MHz}$–$\nu_2 = 100\,{\rm GHz}$, that is, roughly the frequency range observable with radio telescopes. From a physical point of view, the adoption of this frequency band in the calculation of the minimum energy is equivalent to the assumption that only electrons emitting between 10 MHz–100 GHz, that is, with energy between $\gamma_{\rm min} \propto (\nu_1/B_{\rm eq})^{1/2}$ and $\gamma_{\rm max} \propto (\nu_2/B_{\rm eq})^{1/2}$ are present in the radio source. This approach neglects the contribution of the electrons emitting below 10 MHz and, as a more serious bias, in radio sources with different $B_{\rm eq}$ selects different energy bands of the electron population because the energy of the electrons which emit synchrotron radiation at a given frequency depends on the magnetic field intensity [177]. A more consistent approach is to calculate the minimum energy conditions, in which $B_{\rm eq}$ does not depend on the emitted frequency band but directly on the low-energy cutoff of the electron spectrum (typically assumed to be $\gamma_{\rm min} = 100$). These so-called "revised" equipartition conditions select also the contribution to the energetics due to the low-energy electrons [178].

It is typically found that in cavity systems the X-ray pressure is more than one order of magnitude higher than the radio pressure (e.g., [108, 153, 179, 180]). It is also found that with revised equipartition the cavities are closer to pressure balance than they are with standard equipartition (e.g., [153, 165]). Vice versa, by assuming that the lobes are in pressure equilibrium with the ambient gas it is possible to constrain the particle content within the radio lobes [8, 179, 181]. In particular, one can determine the ratio $k_{\rm bal}/f$ that is required to achieve pressure balance under revised equipartition conditions. Several studies of the energetics and particle content of the radio lobes in cooling cores have found high values of $k_{\rm bal}/f$, up to several thousands (with standard equipartition) for active bubbles (e.g., [8, 138]), suggesting that a large fraction of energy must be in nonradiating particles if f is close to unity. On the other hand, the pressure imbalance found in the lobes of FR-I radio galaxies in a sample of galaxy groups appears to be linked to the radio-source morphology, that is, "plumed" sources typically have larger pressure deficits than "bridged" sources where the jets are embedded in the lobes [180]. The authors interpret this result as evidence that plumed sources have a higher entrainment rate due to the larger fraction of the jet surface which is in direct contact with the external medium, leading to an increase in k/f. Although the classification into bridged and plumed morphologies may not directly apply to radio sources at the center of cool core systems, typically having amorphous structures, this picture is consistent with the results of Dunn et al. [182] who argue that the large pressure imbalance observed in radio bubbles as those of the Perseus

cluster is more likely to be due to entrainment rather than a relativistic proton population. Recent studies show lobes having no associated X-ray cavities (e.g., [153, 165]). Assuming their detection is not limited by the sensitivity of the current *Chandra* images, this suggests the possibility of mixing between ambient gas and radio plasma in the lobes. Therefore the $k_{\rm bal}/f > 0$ values measured in such lobes is likely the results of entrainment of thermal gas through the hot gas atmosphere rather than an evidence of heavy jets ejected from the AGN.

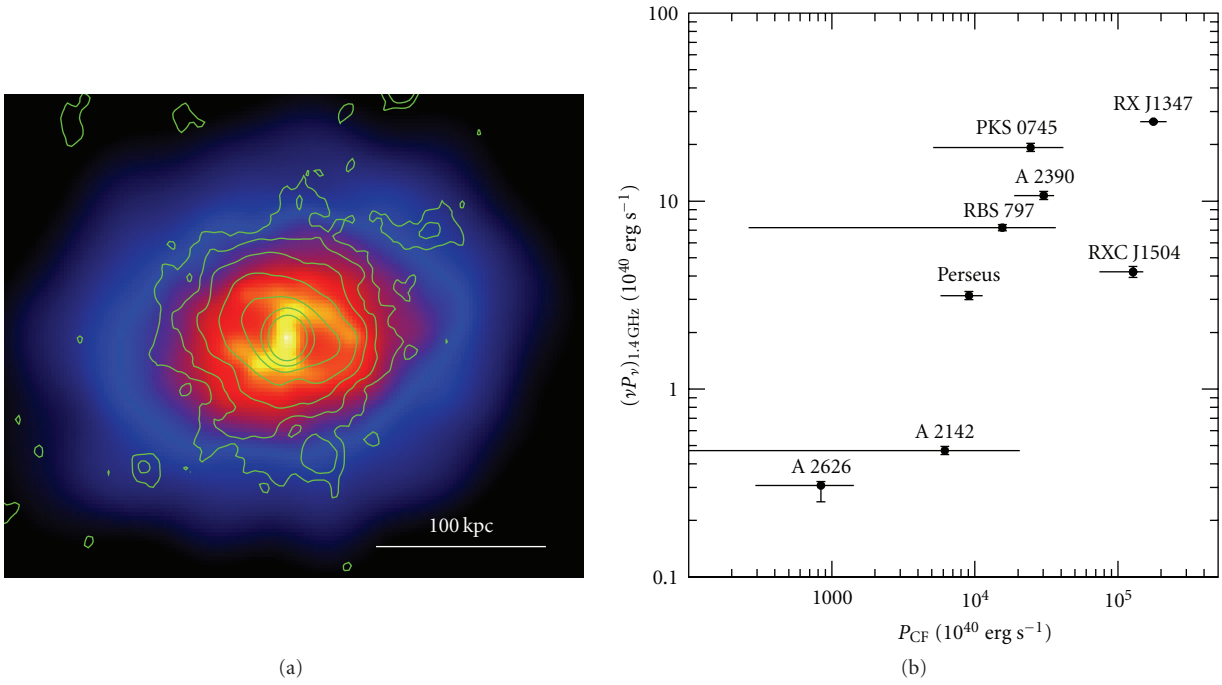
5.5. Radio Mini-Halos. In some cases, the powerful radio galaxies at the center of cool-core clusters are surrounded by diffuse radio emission on scales ∼200–500 kpc having steep radio spectra ($\alpha > 1; S_\nu \propto \nu^{-\alpha}$). These radio sources, generally referred to as "radio mini-halos," are synchrotron emission from GeV electrons diffusing through μG magnetic fields. Although the central radio galaxy is the obvious candidate for the injection of the population of relativistic electrons, mini-halos do appear quite different from the extended lobes maintained by AGN, therefore their radio emission proves that magnetic fields permeate the ICM and at the same time may be indicative of the presence of diffuse relativistic electrons. In particular, due to the fact that the radiative lifetime of radio-emitting electrons ($\sim 10^8$ yr) is much shorter than any reasonable transport time over the cluster scale, the relativistic electrons responsible for the extended radio emission from mini-halos need be continuously reenergized by various mechanisms associated with turbulence in the ICM (reaccelerated *primary* electrons), or freshly injected on a cluster-wide scale (e.g., as a result of the decay of charged pions produced in hadronic collisions, *secondary* electrons). Gitti et al. [187] developed a theoretical model which accounts for the origin of radio mini-halos as related to electron reacceleration by magnetohydrodynamic (MHD) turbulence, which is amplified by compression in the cool cores. In this model, the necessary energetics to power radio mini-halos is supplied by the cooling flow process itself, through the compressional work done on the ICM and the frozen-in magnetic field. The successful application of this model to two cool core clusters (Perseus: [187] and A 2626: [183]) has given support to a primary origin of the relativistic electrons radiating in radio mini-halos.

Radio mini-halos are rare, with only about a dozen objects known so far. Gitti et al. [183] selected an initial sample of five mini-halo clusters based on the presence of both a cool core and a diffuse, amorphous radio emission with no direct association with the central radio source: Perseus [188], A 2626 [183, 189], A 2142 [190], PKS 0745-191 [191], and A 2390 [192]. In these clusters the size of the diffuse radio emission is comparable to that of the cooling region. These criteria are now typically adopted to identify mini-halos. However, the classification of a radio source as a mini-halo is not trivial: their detection is complicated by the fact that the diffuse, low surface brightness emission needs to be separated from the strong radio emission of the central radio galaxy. Furthermore, the criteria adopted to define mini-halos are somehow arbitrary (e.g., total size, morphology, presence of cool core) and some identifications are still controversial. This said, new detections of radio mini-halos have recently been claimed in the galaxy clusters RX J1347.5-1145 [184], Z 7160 [193], A 1835 [194], A 2029 [194], Ophiuchus [194, 195], RXC J1504.1-0248 [186], and RBS 797 (Figure 3(a), see also [111]).

Radio mini-halos are still poorly understood sources. Although secondary electron models have been proposed to explain the presence of their persistent, diffuse radio emission on large-scale in the ICM (e.g., [196, 197]), a primary origin of radio mini-halos is now favored by recent statistical studies [198] and by the observed trend between the radio power of mini-halos and the maximum power of cooling flows (see Figure 3(b)). This indicates a direct connection between cooling flows and radio mini-halos, that is, the most powerful radio mini-halos are associated with the most massive cooling flows, as expected in the framework of the Gitti et al.'s [187] model. However, the origin of the turbulence necessary to trigger the electron reacceleration is still debated. The signatures of minor dynamical activity have recently been detected in some mini-halo clusters, thus suggesting that additional or alternative turbulent energy for the reacceleration may be provided by minor mergers [184, 198] and related gas sloshing mechanism in cool core clusters [199, 200]. Given the prevalence of mini-halos in clusters with X-ray cavities, another attractive possibility is that the turbulent energy is provided by a small fraction of the energy released by the bubbles rising from the central AGN (as suggested by [198]). Needless to say, a larger mini-halo sample as well as further theoretical investigations are necessary to reach a better understanding of this class of sources.

5.6. Weak Shocks and Giant Cavities. In addition to the cavity enthalpy, shocks driven by the AGN outburst may contain a large fraction of the energy released, thus working to heat the ICM. Such shocks have been long predicted by numerical simulations [201–203] but are difficult to detect since they are relatively weak (with Mach numbers ∼1-2) and are seen in projection against the cooler, brighter gas in cluster cores. We also note that to establish these surface brightness discontinuities as shocks one must measure an increase in temperature in the so-called "postshock region," as the ICM is heated by the passage of the shock. Usually the existing images are too shallow to rule out, for example, the possibility that these features are cold front edges, due to gas sloshing (e.g., [19]). Besides a very few examples of strong shocks (e.g., Centaurus A, with Mach number ∼8, [204, 205]), only recently elliptical surface brightness edges, consistent with arising from weak shock fronts driven by the cavities as they initially expanded, have become to emerge in deep *Chandra* exposures of bright clusters and groups. Beautiful examples of cocoon shocks are visible in the Hydra A cluster [116, 135] and in the NGC 5813 group [157], see left panels of Figure 1.

The recent discovery of giant cavities and associated large-scale shocks in three clusters (MS 0735 + 7421 [135], Hercules A [206], Hydra A [116]) has shown that AGN outbursts not only can affect the central regions, but also have an impact on cluster-wide scales possibly affecting the global properties of the ICM and the cluster scaling relations.



(a) (b)

Figure 3: (a) 1.4 GHz VLA radio contours (obtained by combining observations in A-, B-, and C-array configurations) overlaid onto the smoothed 0.5–2.0 keV *Chandra* X-ray image of the galaxy cluster RBS 797. The combined radio map has a resolution of 3″, and is able to reveal the morphology of the central radio source, showing its elongation in the cavity direction (see Figure 1(b)), without losing sensitivity at the larger scales. In particular, the extended radio emission is detected out to ∼90 kpc. By subtracting the contribution of the central nuclear source, the residual flux density of the diffuse radio emission is $\backsimeq 11.5 \pm 0.6$ mJy, indicating the likely presence of a radio mini-halo. (b) Integrated radio power at 1.4 GHz, $[\nu P_\nu]_{1.4\,\mathrm{GHz}}$, versus cooling flow power, $P_{\mathrm{CF}} = \dot{M}kT/\mu m_p$, for the mini-halo clusters which have relevant X-ray and radio data available (data from [111, 183–186]).

In particular, the giant cavities discovered in the galaxy cluster MS 0735 + 7421 have a diameter of about 200 kpc each. The large volume of the cavities implies a huge cavity power: this large-scale outburst is the most powerful known so far, releasing upward of 10^{61} erg into the ICM and heating the gas beyond the cooling region [135, 136]. The new, deep *Chandra* image has confirmed the presence of a weak (Mach number ∼1.3) cocoon shock surrounding the cavity system (Figure 4).

This new development may have significant consequences for several fundamental problems in astrophysics. As seen in Section 2, the observed relation between X-ray luminosity and gas temperature in clusters is steeper than expected if cluster growth were governed by gravity alone. This steepening is best explained by the addition of heat to the ICM and is therefore considered the main manifestation of nongravitational heating. The discovery of giant cavities has indicated that powerful AGN outbursts occurring at late times may contribute a significant fraction of the extra nongravitational energy. As mentioned above, this additional heating supplied by AGN could also induce the suppression of the gas cooling in massive galaxies required to explain the exponential turnover at the bright end of the luminosity function of galaxies (e.g., [11]). This would indicate a common solution for the two major heating problems associated with the ICM: those of cooling flow and galaxy formation. In the case of MS 0735 + 7421, the driving energy of the shock as determined using a spherical hydrodynamic model is $E_s \approx 5.7 \times 10^{61}$ erg [135]. As estimated by Gitti et al. [136], the AGN outburst in this cluster is heating the gas mass within 1 Mpc ($\sim 7.7 \times 10^{13} M_\odot$) at the level of about 1/4 keV per particle, and the heating level increases to ∼0.6 keV per particle when considering the gas mass within r_{2500}. This is a substantial fraction of the 1–3 keV per particle of excess energy required to heat the cluster [207]. Only a few outbursts of this magnitude erupting over the life of a cluster would be required to heat it. By contrast, MS 0735 + 7421 is found to be a factor ∼2 more luminous than expected from its average temperature on the basis of the observed *L*-*T* relation for galaxy clusters [136]. Based on the data presented in Gitti et al. [120], we found a similar result for the giant cavity cluster Hydra A (Figure 5(a)). Although caution should be taken in drawing general conclusions from the study of only a few objects, this indicates that flux limited samples of distant X-ray clusters may be biased in favor of detecting clusters with energetic AGN outbursts. We also note that powerful AGN outbursts may have a dramatic effect on the gas mass fraction measurements, due to an overestimate of the gas density [136]. The observed departure of MS 0735 + 7421 and Hydra A from the mean *L*-*T* relation is in apparent contradiction with the argument above that heat should steep the *L*-*T*

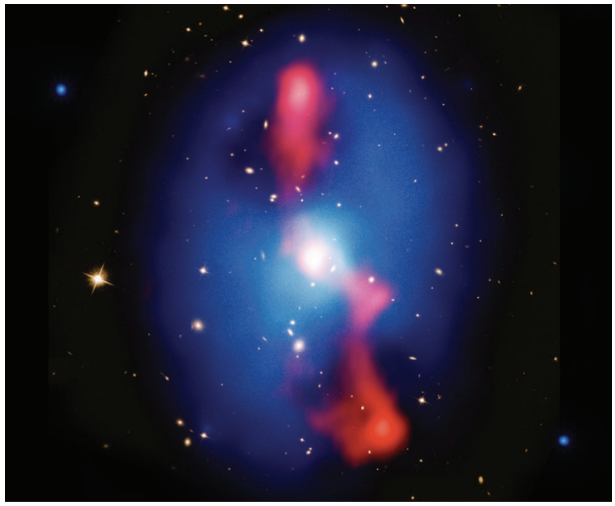


Figure 4: Deep ~500 ks *Chandra* X-ray image (blue) and Very Large Array 330 MHz radio image (red) superposed with the *Hubble Space Telescope* visual image of the galaxy cluster MS 0735 + 7421. The giant X-ray cavities, filled with radio emission, are surrounded by a cocoon shock clearly visible in the *Chandra* image as an elliptical edge. The box is roughly 800 kpc by 800 kpc.

relation, as also indicated by recent numerical simulations (e.g., [208]). However, we stress that the observed L-T relation is highly dependent on the definition of the characteristic temperature, that is, for a fixed luminosity the position of each point in the plot may vary significantly depending on the choice of the method adopted to measure the average emission-weighted temperature for each cluster. Furthermore, the possibility of building a consistent L-T scaling relation from a sample of clusters relies on the capability to correct both the temperature and the luminosity measurements for the effects of the central cooling flow in a consistent manner for the whole sample. This may not be trivial as the physical conditions can vary significantly from case to case. For example, the commonly adopted method of excluding the central 70 kpc is found to have some drawback for giant cavity systems as the cooling region and the effect of AGN feedback extend beyond this radius (see discussion in [136]).

On the other hand, based on a study of *XMM-Newton* data, Gitti et al. [136] have shown that the energetic outburst in MS 0735 + 7421 does not cause dramatic instantaneous departures from the average properties of the ICM because it has not had a measurable impact on the large-scale temperature profile, which is in fact consistent within the scatter of the profiles observed in relaxed cluster [79]. We recently found a similar result (Figure 5(b)) in the Hydra A cluster, although a sort of "plateau" standing below the typical profile in the range of radius ~0.05–0.1$r_{\rm vir}$ indicates the presence of cooler gas (see Section 5.8). In general, these results suggest that there cannot have been many previous outbursts of high magnitude in these clusters, otherwise the total energy added to the ICM outside the cooling region should have had a marked effect. Studies of cavity samples found that the prevalence of outbursts as energetic as 10^{61} erg is 3 (namely, MS 0735 + 7421, Hercules A and Hydra A) over 30 [139]. If, as it appears from our in-depth studies of MS 0735+7421 and Hydra A, such powerful outbursts are rare in individual clusters, their occurrence in ~10% of known cases hence requires that they occur a similar fraction of time in most cooling flow clusters. This picture is consistent with the model proposed by Nipoti and Binney [170], in which the AGN activity is strongly variable with time and all systems occasionally experience powerful outbursts.

5.7. SMBH Growth. AGN are powered by the release of gravitational binding energy from accretion onto massive black holes [211, 212]. The matter that reaches the black hole converts it binding energy efficiently into AGN power as $P_{\rm jet} = \epsilon \dot{M} c^2$, where ϵ ~0.1–0.4 depending on the spin of the black hole. Rapidly spinning black holes with spin parameters approaching unity are most efficient due to the smaller radius of the innermost stable circular orbit. The form of energy that is released depends on several factors including the accretion rate, the mass of the black hole, the structure of the accretion flow, and the spin of the black hole (see [213] for a review). When black hole accretion approaches the Eddington rate, the binding energy is emitted thermally from an optically thick, geometrically thin accretion disk that is morphologically classified as a quasar or Seyfert galaxy. When the accretion rate drops below a few percent of the Eddington rate a radiatively inefficient AGN is formed (i.e., an ADAF), and the energy is released primarily in the form of mechanical energy associated with a radio jet. Accretion rates in bright AGN can be estimated using the radiation emitted from the nucleus that directly (e.g., UV or X-ray emission) or indirectly (e.g., nebular or far-IR emission) trace the energetic output from the accretion disk (e.g., [214]). However, despite having mechanically powerful radio AGN [137], brightest cluster galaxies (BCGs) rarely show strong X-ray and UV emission emerging from their nuclei [215], implying that their accretion rates generally lie below a few percent of their Eddington rate.

Rafferty et al. [139] estimated the accretion rates in a sample of BCGs in clusters with prominent X-ray cavities and found this to be the case. They estimated the accretion rates using the measured output power based on the pV work done by the cavities divided by their buoyancy ages. Using this approach, Rafferty et al. [139] found that supermassive black holes centred in cool cores are growing at a rate of $\sim 0.1 M_\odot\ {\rm yr}^{-1}$. In rare instances such as the powerful AGN in MS 0735 + 7421, [135, 136, 216], the accretion rate exceeds $1 M_\odot\ {\rm yr}^{-1}$. Assuming black hole masses that are consistent with the values expected from their stellar luminosities and velocity dispersions, the accretion rates are consistent with being currently at most a few percent of Eddington. If AGN feedback in BCGs quenches cooling flows over the lifetimes of clusters, their black holes may be more massive than predicted by the $M_{\rm BH}$-σ relation.

5.8. Further Evidence for Mechanical Feedback. As we have seen in Section 5.2, it is now widely accepted that the AGN in the cool cores can reheat the ICM. Although this is certainly its main impact, AGN feedback is likely to have other

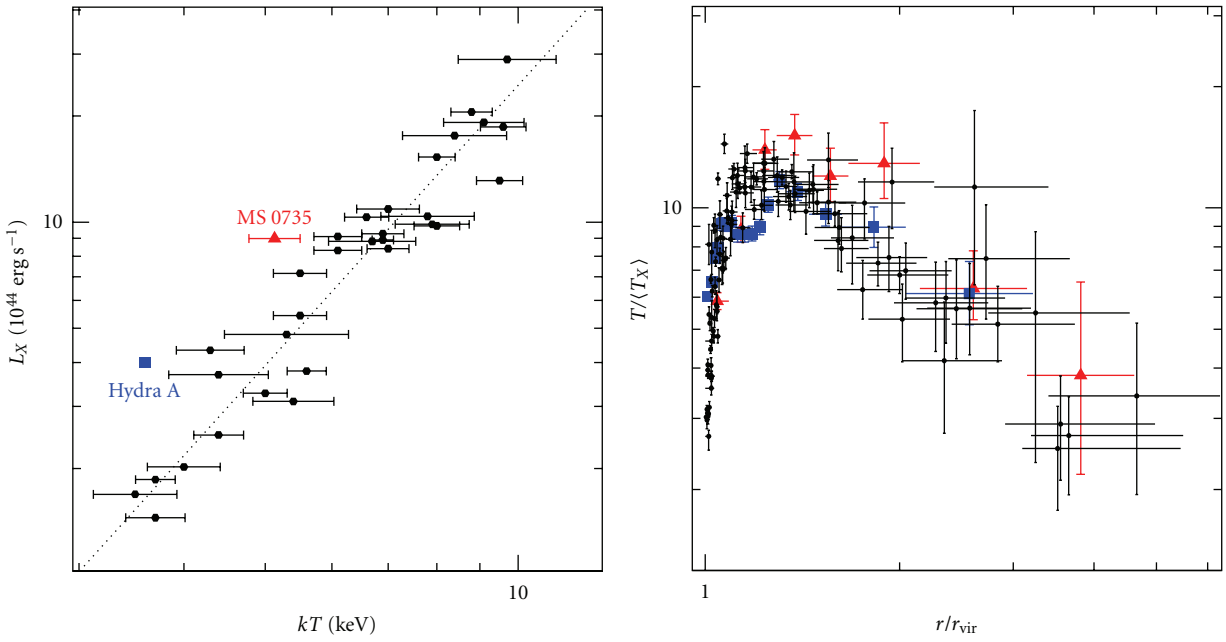


FIGURE 5: (a) Bolometric X-ray luminosities corrected for the effect of cooling flow in the central ~70 kpc versus emission-weighted temperatures derived excluding cooling flow components, from Markevitch [25]. The dashed line is the best-fit power law to the sample: $L_X = 6.35 \cdot (kT/6\ \mathrm{keV})^{2.64} \times 10^{44}\ \mathrm{erg\,s^{-1}}$. The red triangle and the blue square represent the measurements from the *XMM-Newton* and *Chandra* data of the giant cavity clusters MS 0735 + 7421 [209] and Hydra A, respectively, corrected consistently with the method adopted by Markevitch [25]. (b) Temperature profiles measured for a sample of relaxed clusters presented by Vikhlinin et al. [79]. The temperatures are scaled to the cluster emission-weighted temperature excluding the central 70 kpc regions. The profiles for all clusters are projected and scaled in radial units of the virial radius $r_{\rm vir}$, estimated from the relation $r_{\rm vir} = 2.74\ \mathrm{Mpc}\sqrt{\langle T_X \rangle / 10\,\mathrm{keV}}$ [210]. Overlaid are the cooling flow corrected, scaled temperature profiles of the giant cavity clusters MS 0735 + 7421 (red triangles, Gitti et al. [136]) and Hydra A (blue squares, Gitti et al. [120]).

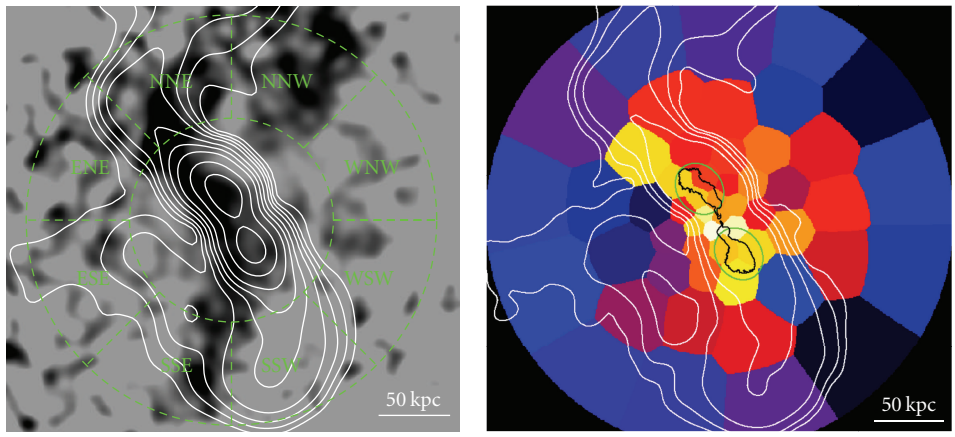


FIGURE 6: (a) High-contrast hardness ratio map of the galaxy cluster Hydra A obtained by dividing a 1.5–7.5 keV image by a 0.3–1.5 keV image. Regions in black are indicative of low temperature gas, indicating the presence of low-entropy filaments. Overlaid in green are the sectors used to study the spectral properties of the cool gas (located between radii ~70–150 kpc). Gitti et al. [120] extracted the spectra in these sectors and compared two different spectral models: a single-temperature plasma "1T" model, and a "2T" model which includes a second thermal emission component. The F statistics for the spectral fitting improvement over the single-phase model indicate that the addition of a second thermal component is most significant in sectors SSE, NNW, NNE, and ENE, thus providing evidence for the presence of multiphase gas in agreement with the hardness ratio map. Adapted from Gitti et al. [120]. (b) Metallicity map showing the central $5' \times 5'$ of Hydra A. Brighter regions represent a higher metallicity. The color scale of the metallicity Z (in solar units) is as follows: white: $Z \geq 0.75$, yellow: $Z = 0.65$–0.75, orange: $Z = 0.55$–0.65, red: $Z = 0.45$–0.55, blue: $Z = 0.3$–0.45, and black: $Z \leq 0.3$. The 1.4 GHz radio emission is shown by the black contours. The green elliptical regions indicate the position of the inner cavities. Adapted from Kirkpatrick et al. [119]. In both panels are overlaid the white contours outlining the 330 MHz radio emission from Lane et al. [164].

important effects on the ICM. Gitti et al. [120] have recently investigated this point by performing an in-depth study of the galaxy cluster Hydra A, which harbors a well-known, large-scale system of X-ray cavities embedded in a "cocoon" shock surrounding the central, powerful radio source ([107, 116], see Figure 1(a)). By means of a detailed spectral analysis of the deep (~240 ks) *Chandra* observations, Gitti et al. [120] found indication of the presence of multiphase gas along soft filaments seen in the hardness ratio map (Figure 6(a)). Interestingly, such filaments are almost spatially coincident with the radio lobes of the powerful central radio source. The cooler gas has a significant impact on the radial temperature profile of the cluster, creating a sort of "plateau" which departs from the typical profile (blue squares in Figure 5(b)). In fact, the scaled temperature profile of Hydra A measured after masking the filaments is found to agree well with the general shape of the temperature profiles observed for relaxed clusters, thus providing a confirmation that these filaments contain cool gas [120]. By performing a spectral deprojection analysis of an absorbed 2-temperature component model, Gitti et al. [120] found evidence that $\sim 10^{11} M_{\odot}$ of low-entropy material has moved upward from the central 30 kpc to the observed current position of 75–150 kpc, likely due to some form of entrainment or dredge up by the rising lobes. Assuming that the mass of cool gas, which is ~60% of the total mass of gas remaining within 30 kpc [4], was lifted out of the central cluster region by a continuous outflow or a series of bursts from the central AGN over the past 200–500 Myr (as it appears from the study of the cavity system, [117]), it would amount to outflows of a few hundred $M_{\odot}\,\mathrm{yr}^{-1}$. There would thus be a development of gas circulation that can significantly reduce the net inflow of cooling gas, as initially discussed by David et al. [4] and Nulsen et al. [110]. Therefore Gitti et al.'s [120] results show that the AGN feedback in Hydra A is acting not only by directly re-heating the gas, but also by removing a substantial amount of potential fuel for the supermassive black hole (SMBH). This provides indications of mechanical AGN feedback acting through collimated, massive outflows generated by jets or cavity dragging (e.g., [217]).

The energy required to lift the cool gas gives a lower limit on the amount of AGN outburst energy deposited in the ICM. This value can be estimated by calculating the variation in gravitational potential energy during the lifting process. By assuming that the undisturbed ICM is approximately isothermal with sound speed $c_s \approx 1000\,\mathrm{km\,s}^{-1}$ and is in a 3 hydrostatic configuration with density profile $\rho(r)$, it is possible to calculate this quantity as [218]:

$$\Delta E = \frac{M_{\mathrm{cool}} c_s^2}{\gamma} \ln\left(\frac{\rho_i}{\rho_f}\right), \tag{26}$$

where M_{cool} is the lifted mass, ρ_i and ρ_f are the initial and final densities of the surrounding ICM, and $\gamma = 5/3$ is the ratio of specific heat capacities. From the density profile presented by David et al., Gitti et al. [120] estimated that the energy required to lift the cool gas is $\gtrsim 2.2 \times 10^{60}$ erg. This value is comparable to the work required to inflate all of the cavities against the surrounding pressure [117] and is ~25% of the total energy of the large-scale shock [116]. Gitti et al. [120] also note that a good fraction of the energy required to lift the low entropy gas will be thermalized when the gas falls back inward. Given the large energy required, uplift provides a significant channel for the dissipation of outburst energy in Hydra A. There is a remarkable correlation between these low-entropy filaments and the metal-rich filaments in the iron-abundance maps measured by Simionescu et al. [219] and Kirkpatrick et al. [119], shown in Figure 6(b). The emerging picture is that Hydra A's powerful radio source is able to lift cool, metal-rich gas from the central region and distribute it throughout the X-ray atmosphere of the cluster. A similar effect of extended metal outflows in the direction of the radio lobes and X-ray cavities is observed in other systems (e.g., M87 [124, 220], A 262: [221], RBS 797 [111]). This is consistent with the results of the most current theoretical modeling of AGN feedback in massive cosmological systems, which predict the massive subrelativistic bipolar outflows and buoyant bubbles to produce a metal uplift along the jet axis (see Section 5.9).

5.9. Numerical Simulations. In the last decade the phenomenon of AGN feedback and its impact on the ICM has been the subject of many theoretical investigations (see [92, 93, 163, 174, 202, 203, 222–230], to name a few). These works focused either on several aspects of feedback physics and microphysics or with the global, long-term evolution of the ICM. However, uncertainties are still large enough that the observations must guide researchers to select the relevant mechanisms at work in the feedback process.

The body of different observational investigations set a number of key constraints on process of AGN feedback. The results presented in this paper strongly suggest that AGN feedback manifests itself as massive subrelativistic bipolar outflows which heat the ICM through weak shocks and form X-ray cavities, lift large masses of hot gas from the central region ≥100 kpc and generate abundance asymmetries along the outflow direction. Processes such as AGN Compton heating or thermal conduction are unable to explain the collection of observations described above. Thus, although possibly relevant in some respect (e.g., [231]), they likely play a secondary role in local clusters and groups.

Recent 3D hydrosimulations of outflow feedback ([163, 203]; see, among others, also [36, 202, 217, 228, 232–235]) have quantitatively verified that collimated outflows reduce the gas cooling rate below the limits set by recent *Chandra* and *XMM-Newton* observations ([5, 6] and references therein) for a timescale comparable with the cluster age. At the same time, this feedback mechanism generates ICM density and temperature profiles which reasonably agree with those of typical cool core clusters. This is not a trivial result [93, 174]. The observable effects of the feedback from bipolar AGN outflows range from creation of X-ray cavities and large-scale shocks to inducing entropy and metal abundance anisotropies, due to the lifting of central gas, relatively cool and metal rich, along the direction of the jet (see also [236, 237]). These models, far from being complete and exhaustive

(the origin of the outflows is essentially ad hoc), represent an interesting starting point for a more thorough understanding of the AGN feedback process.

6. Concluding Remarks

In order to fully understand the growth and evolution of galaxies and their central black holes, the history of star formation, and the formation of large-scale structures, it is crucial to understand first the processes of cooling, heating and the dynamical evolution of the intracluster gas. In particular, the feedback from the central black holes has turned out to be an essential ingredient that must be taken into account in any model of galaxy evolution. The main manifestation of the action of AGN feedback is in galaxy clusters and groups. Their study, which is currently a very active line of research in extragalactic astrophysics, has allowed us to make significant progresses in this field. However, many details of the AGN feedback mechanism are still unclear. It is not well understood, for instance, how the heating distributes in space and time in order to drastically reduce gas cooling, preserving at the same time the central cool core. An even more puzzling issue is the process of black hole accretion and feedback energy generation.

The last decade has represented a quantum leap in the quality of X-ray observations, thanks to the *Chandra* and *XMM-Newton* satellite telescopes. Recent results, discussed in this paper, have shaked up our understanding of the gas astrophysics in systems ranging from massive elliptical galaxies to rich galaxy clusters. They suggest that bipolar outflows emerging from the BCG core inflate large bubbles, heat the ICM and induce a circulation of gas and metals on scales of several 100 s kpc.

The current generation of X-ray observatories is still working well and can be expected to continue doing so for few more years. Unfortunately, the prospects for the future of X-ray astronomy are not clear at the moment, and X-ray astronomers must rely on the good health on the existing, but aging, X-ray telescopes. As a result, the next few years represent a narrow time window to exploit the unique opportunity to observe deeply many additional objects, thus collecting crucial information on the cluster and group evolution. At the same time, current radio instrumentation is steadily improving both at the MHz and at the mm-wave ends of the spectrum and is about to make a significant step forward with the next generation of observatories, such as Low Frequency Array (LOFAR) and Atacama Large Millimeter/submillimeter Array (ALMA). A common effort, from both the observational and theoretical sides, will allow us to widen our knowledge on this fundamental problem which is central to the entire field of extragalactic astrophysics.

Acknowledgments

The authors especially thank Ewan O'Sullivan for sharing unpublished results and for providing useful comments. They thank Clif Kirkpatrick and Scott Randall for permission to print figures from their work. M. Gitti thanks Paul Nulsen, Larry David, Jan Vrtilek, and Ewan O'Sullivan for many stimulating discussions on this exciting topic, and Stefano Ettori for carefully reading the original paper. M. Gitti acknowledges support by Grants ASI-INAF I/023/05/0 and I/088/06/0 and by *Chandra* grant GO0-11136X. B. R. McNamara acknowledges support from the Natural Sciences and Engineering Research Council of Canada.

References

[1] S. D. M. White and M. J. Rees, "Core condensation in heavy halos—a two-stage theory for galaxy formation and clustering," *Monthly Notices of the Royal Astronomical Society*, vol. 183, pp. 341–358, 1978.

[2] A. C. Fabian, "Cooling flows in clusters of galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 32, no. 1, pp. 277–318, 1994.

[3] J. R. Peterson, F. B. S. Paerels, J. S. Kaastra et al., "X-ray imaging-spectroscopy of Abell 1835," *Astronomy and Astrophysics*, vol. 365, no. 1, pp. L104–L109, 2001.

[4] L. P. David, P. E. J. Nulsen, B. R. Mcnamara et al., "A high-resolution study of the Hydra A cluster with Chandra: comparison of the core mass distribution with theoretical predictions and evidence for feedback in the cooling flow," *Astrophysical Journal*, vol. 557, no. 2, pp. 546–559, 2001.

[5] J. R. Peterson and A. C. Fabian, "X-ray spectroscopy of cooling clusters," *Physics Reports*, vol. 427, no. 1, pp. 1–39, 2006.

[6] B. R. McNamara and P. E. J. Nulsen, "Heating hot atmospheres with active galactic nuclei," *Annual Review of Astronomy and Astrophysics*, vol. 45, pp. 117–175, 2007.

[7] J. O. Burns, "The radio properties of cD galaxies in abell clusters. I. An X-ray-selected sample," *Astronomical Journal*, vol. 99, no. 1, pp. 14–30, 1990.

[8] L. Bîrzan, B. R. McNamara, P. E. J. Nulsen, C. L. Carilli, and M. W. Wise, "Radiative efficiency and content of extragalactic radio sources: toward a universal scaling relation between jet power and radio power," *Astrophysical Journal*, vol. 686, no. 2, pp. 859–880, 2008.

[9] S. Giacintucci, E. O'Sullivan, J. Vrtilek et al., "A combined low-radio frequency/X-ray study of galaxy groups. I. Giant metrewave radio telescope observations at 235 MHz and 610 MHz," *Astrophysical Journal Letters*, vol. 732, article 95, no. 2, 2011.

[10] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[11] A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole, "What shapes the luminosity function of galaxies?" *Astrophysical Journal*, vol. 599, no. 1, pp. 38–49, 2003.

[12] D. J. Croton, V. Springel, S. D. M. White et al., "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 1, pp. 11–28, 2006.

[13] M. L. Balogh, A. Babul, G. M. Voit et al., "An analytic investigation of the scatter in the integrated X-ray properties of galaxy groups and clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 366, no. 2, pp. 624–634, 2006.

[14] A. Cattaneo, S. M. Faber, J. Binney et al., "The role of black holes in galaxy formation and evolution," *Nature*, vol. 460, no. 7252, pp. 213–219, 2009.

[15] C. L. Sarazin, *X-Ray Emission from Clusters of Galaxies*, 1988.
[16] R. F. Mushotzky, "Clusters of galaxies: an X-ray perspective," http://arxiv.org/abs/astro-ph/0311105.
[17] M. Arnaud, "X-ray observations of clusters of galaxies," in *Background Microwave Radiation and Intracluster Cosmology*, F. Melchiorri and Y. Rephaeli, Eds., p. 77, 2005.
[18] G. M. Voit, "Tracing cosmic evolution with clusters of galaxies," *Reviews of Modern Physics*, vol. 77, no. 1, pp. 207–258, 2005.
[19] M. Markevitch and A. Vikhlinin, "Shocks and cold fronts in galaxy clusters," *Physics Reports*, vol. 443, no. 1, pp. 1–53, 2007.
[20] P. Rosati, S. Borgani, and C. Norman, "The evolution of X-ray clusters of galaxies," *Annual Review of Astronomy and Astrophysics*, vol. 40, pp. 539–577, 2002.
[21] W. H. Press and P. Schechter, "Formation of galaxies and clusters of galaxies by self-similar gravitational condensation," *Astrophysical Journal*, vol. 187, pp. 425–438, 1974.
[22] S. Borgani, "Cosmology with clusters of galaxies," http://arxiv.org/abs/astro-ph/0605575.
[23] N. Kaiser, "Evolution and clustering of rich clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 222, pp. 323–345, 1986.
[24] G. L. Bryan and M. L. Norman, "Statistical properties of x-ray clusters: analytic and numerical comparisons," *Astrophysical Journal*, vol. 495, no. 1, pp. 80–99, 1998.
[25] M. Markevitch, "The LX-T relation and temperature function for nearby clusters revisited," *Astrophysical Journal*, vol. 504, no. 1, pp. 27–34, 1998.
[26] M. Arnaud and A. E. Evrard, "The LX-T relation and intracluster gas fractions of X-ray clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 305, no. 3, pp. 631–640, 1999.
[27] S. Ettori, P. Tozzi, S. Borgani, and P. Rosati, "Scaling laws in X-ray galaxy clusters at redshift between 0.4 and 1.3," *Astronomy and Astrophysics*, vol. 417, no. 1, pp. 13–27, 2004.
[28] B. J. Maughan, L. R. Jones, H. Ebeling, and C. Scharf, "The evolution of the cluster X-ray scaling relations in the Wide Angle ROSAT Pointed Survey sample at $0.6 < z < 1.0$," *Monthly Notices of the Royal Astronomical Society*, vol. 365, no. 2, pp. 509–529, 2006.
[29] N. Ota, T. Kitayama, K. Masai, and K. Mitsuda, "LX-T relation and related properties of galaxy clusters," *Astrophysical Journal*, vol. 640, no. 2, pp. 673–690, 2006.
[30] A. Mantz, S. W. Allen, D. Rapetti, and H. Ebeling, "The observed growth of massive galaxy clusters—I. Statistical methods and cosmological constraints," *Monthly Notices of the Royal Astronomical Society*, vol. 406, no. 3, pp. 1759–1772, 2010.
[31] N. Kaiser, "Evolution of clusters of galaxies," *Astrophysical Journal*, vol. 383, no. 1, pp. 104–111, 1991.
[32] A. E. Evrard and J. P. Henry, "Expectations for X-ray cluster observations by the ROSAT satellite," *Astrophysical Journal*, vol. 383, no. 1, pp. 95–103, 1991.
[33] S. Cole, "Modeling galaxy formation in evolving dark matter halos," *Astrophysical Journal*, vol. 367, no. 1, pp. 45–53, 1991.
[34] S. D. M. White and C. S. Frenk, "Galaxy formation through hierarchical clustering," *Astrophysical Journal*, vol. 379, no. 1, pp. 52–79, 1991.
[35] M. L. Balogh, F. R. Pearce, R. G. Bower, and S. T. Kay, "Revisiting the cosmic cooling crisis," *Monthly Notices of the Royal Astronomical Society*, vol. 326, no. 4, pp. 1228–1234, 2001.
[36] D. Sijacki and V. Springel, "Hydrodynamical simulations of cluster formation with central AGN heating," *Monthly Notices of the Royal Astronomical Society*, vol. 366, no. 2, pp. 397–416, 2006.
[37] C. A. Metzler and A. E. Evrard, "A simulation of the intracluster medium with feedback from cluster galaxies," *Astrophysical Journal*, vol. 437, no. 2, pp. 564–583, 1994.
[38] S. Borgani, F. Governato, J. Wadsley et al., "The effect of non-gravitational gas heating in groups and clusters of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 336, no. 2, pp. 409–424, 2002.
[39] A. Finoguenov, M. Arnaud, and L. P. David, "Temperature and heavy-element abundance profiles of cool clusters of galaxies from ASCA," *Astrophysical Journal*, vol. 555, no. 1, pp. 191–204, 2001.
[40] A. Pipino, F. Matteucci, S. Borgani, and A. Biviano, "SNe heating and the chemical evolution of the intra-cluster medium," *New Astronomy*, vol. 7, no. 5, pp. 227–247, 2002.
[41] L. Tornatore, S. Borgani, V. Springel, F. Matteucci, N. Menci, and G. Murante, "Cooling and heating the intracluster medium in hydrodynamical simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 342, no. 4, pp. 1025–1040, 2003.
[42] G. M. Voit, G. L. Bryan, M. L. Balogh, and R. G. Bower, "Modified entropy models for the intracluster medium," *Astrophysical Journal*, vol. 576, no. 2, pp. 601–624, 2002.
[43] K. K. S. Wu, A. C. Fabian, and P. E. J. Nulsen, "The soft X-ray background: evidence for widespread disruption of the gas haloes of galaxy groups," *Monthly Notices of the Royal Astronomical Society*, vol. 324, no. 1, pp. 95–107, 2001.
[44] E. Scannapieco and S. P. Oh, "Quasar feedback: the missing link in structure formation," *Astrophysical Journal*, vol. 608, no. 1, pp. 62–79, 2004.
[45] M. Donahue, D. J. Horner, K. W. Cavagnolo, and G. M. Voit, "Entropy profiles in the cores of cooling flow clusters of galaxies," *Astrophysical Journal*, vol. 643, no. 2, pp. 730–750, 2006.
[46] G. M. Voit and M. Donahue, "An observationally motivated framework for agn heating of cluster cores," *Astrophysical Journal*, vol. 634, no. 2, pp. 955–963, 2005.
[47] L. Ferrarese and D. Merritt, "A fundamental relation between supermassive black holes and their host galaxies," *Astrophysical Journal*, vol. 539, no. 1, pp. L9–L12, 2000.
[48] K. Gebhardt, R. Bender, G. Bower et al., "A relationship between nuclear black hole mass and galaxy velocity dispersion," *Astrophysical Journal*, vol. 539, no. 1, pp. L13–L16, 2000.
[49] S. Tremaine, K. Gebhardt, R. Bender et al., "The slope of the black hole mass versus velocity dispersion correlation," *Astrophysical Journal*, vol. 574, no. 2, pp. 740–753, 2002.
[50] J. E. Felten, R. J. Gould, W. A. Stein, and N. J. Woolf, "X-rays from the coma cluster of galaxies," *Astrophysical Journal Letters*, vol. 146, pp. 955–958, 1966.
[51] L. J. Spitzer, *Physics of Fully Ionized Gases*, 1956.
[52] L. J. Spitzer, *Physical Processes in the Interstellar Medium*, 1978.
[53] R. K. Smith and J. P. Hughes, "Ionization equilibrium timescales in collisional plasmas," *Astrophysics Journal*, vol. 718, pp. 583–585, 2010.
[54] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics*, Wiley-VCH, 1979.
[55] W. J. Karzas and R. Latter, "Electron radiative transitions in a coulomb field," *Astrophysical Journal Letters*, vol. 6, p. 167, 1961.

[56] E. Kellogg, J. R. Baldwin, and D. Koch, "Studies of cluster X-ray sources, energy spectra for the Perseus, Virgo, and Coma clusters," *Astrophysical Journal Letters*, vol. 199, pp. 299–306, 1975.

[57] J. C. Raymond and B. W. Smith, "Soft X-ray spectrum of a hot plasma," *Astrophysical Journal Supplement Series*, vol. 35, pp. 419–439, 1977.

[58] R. Mewe and E. H. B. M. Gronenschild, "Calculated X-radiation from optically thin plasmas. IV—atomic data and rate coefficients for spectra in the range 1–270 A," *Astronomy and Astrophysics Supplement Series*, vol. 45, pp. 11–52, 1981.

[59] C. L. Sarazin and J. N. Bahcall, "X-ray line emission for clusters of galaxies. II—numerical models," *Astrophysical Journal Supplement Series*, vol. 34, pp. 451–467, 1977.

[60] J. N. Bahcall and C. L. Sarazin, "Parameters and predictions for the X-ray emitting gas of Coma, Perseus, and Virgo," *Astrophysical Journal*, vol. 213, pp. L99–L103, 1977.

[61] J. N. Bahcall and C. L. Sarazin, "X-ray line spectroscopy for clusters of galaxies. I," *Astrophysical Journal*, vol. 219, pp. 781–794, 1978.

[62] C. F. McKee and L. L. Cowie, "The evaporation of spherical clouds in a hot gas. II—effects of radiation," *Astrophysical Journal*, vol. 215, pp. 213–225, 1977.

[63] R. S. Sutherland and M. A. Dopita, "Cooling functions for low-density astrophysical plasmas," *Astrophysical Journal, Supplement Series*, vol. 88, no. 1, pp. 253–327, 1993.

[64] L. M. Voigt and A. C. Fabian, "Galaxy cluster mass profiles," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 2, pp. 518–533, 2006.

[65] M. Gitti, R. Piffaretti, and S. Schindler, "Mass distribution in the most X-ray-luminous galaxy cluster RX J1347.5-1145 studied with *XMM-Newton*," *Astronomy and Astrophysics*, vol. 472, no. 2, pp. 383–394, 2007.

[66] F. Gastaldello, D. A. Buote, P. J. Humphrey et al., "Probing the dark matter and gas fraction in relaxed galaxy groups with X-ray observations from Chandra and *XMM-Newton*," *Astrophysical Journal*, vol. 669, no. 1, pp. 158–183, 2007.

[67] S. Ettori, F. Gastaldello, A. Leccardi et al., "Mass profiles and c-MDM relation in X-ray luminous galaxy clusters," *Astronomy and Astrophysics*, vol. 524, no. 3, article 68, 2010.

[68] E. Rasia, S. Ettori, L. Moscardini et al., "Systematics in the X-ray cluster mass estimators," *Monthly Notices of the Royal Astronomical Society*, vol. 369, no. 4, pp. 2013–2024, 2006.

[69] D. Nagai, A. V. Kravtsov, and A. Vikhlinin, "Effects of galaxy formation on thermodynamics of the intracluster medium," *Astrophysical Journal*, vol. 668, no. 1, pp. 1–14, 2007.

[70] R. Piffaretti and R. Valdarnini, "Total mass biases in X-ray galaxy clusters," *Astronomy and Astrophysics*, vol. 491, no. 1, pp. 71–87, 2008.

[71] A. Cavaliere and R. Fusco-Femiano, "X-rays from hot plasma in clusters of galaxies," *Astronomy and Astrophysics*, vol. 49, pp. 137–144, 1976.

[72] M. Arnaud, "The beta of the intracluster medium," *Astronomy and Astrophysics*, vol. 500, no. 1, pp. 103–104, Commentary on: A. Cavaliere and R. Fusco-Femiano, Astronomy and Astrophysics, vol. 49, pp. 137, 1976, 2009.

[73] I. King, "The structure of star clusters. I. An empirical density law," *Astronomical Journal*, vol. 67, p. 471, 1962.

[74] S. Ettori and F. Brighenti, "On the evolution of cooling cores in X-ray galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 387, no. 2, pp. 631–638, 2008.

[75] Y. Chen, T. H. Reiprich, H. Böhringer, Y. Ikebe, and Y. Y. Zhang, "Statistics of X-ray observables for the cooling-core and non-cooling core galaxy clusters," *Astronomy and Astrophysics*, vol. 466, no. 3, pp. 805–812, 2007.

[76] S. De Grandi and S. Molendi, "Metallicity gradients in X-ray clusters of galaxies," *Astrophysical Journal*, vol. 551, no. 1, pp. 153–159, 2001.

[77] S. de Grandi and S. Molendi, "Metal abundances in the cool-cores of galaxy clusters," *Astronomy and Astrophysics*, vol. 508, pp. 565–574, 2009.

[78] S. W. Allen, R. W. Schmidt, and A. C. Fabian, "The X-ray virial relations for relaxed lensing clusters observed with Chandra," *Monthly Notices of the Royal Astronomical Society*, vol. 328, no. 3, pp. L37–L41, 2001.

[79] A. Vikhlinin, M. Markevitch, S. S. Murray, C. Jones, W. Forman, and L. Van Speybroeck, "Chandra temperature profiles for a sample of nearby relaxed galaxy clusters," *Astrophysical Journal*, vol. 628, no. 2, pp. 655–672, 2005.

[80] D. S. Hudson, R. Mittal, T. H. Reiprich, P. E. J. Nulsen, H. Andernach, and C. L. Sarazin, "What is a cool-core cluster? a detailed analysis of the cores of the X-ray flux-limited HIFLUGCS cluster sample," *Astronomy and Astrophysics*, vol. 513, article A37, p. 37, 2010.

[81] R. M. Johnstone, S. W. Allen, A. C. Fabian, and J. S. Sanders, "Chandra observations of Abell 2199," *Monthly Notices of the Royal Astronomical Society*, vol. 336, no. 1, pp. 299–308, 2002.

[82] H. Böhringer and N. Werner, "X-ray spectroscopy of galaxy clusters: studying astrophysical processes in the largest celestial laboratories," *Astronomy and Astrophysics Review*, vol. 18, no. 1-2, pp. 127–196, 2010.

[83] A. C. Fabian, R. F. Mushotzky, P. E. J. Nulsen, and J. R. Peterson, "On the soft X-ray spectrum of cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 321, no. 1, pp. L20–L24, 2001.

[84] R. G. Morris and A. C. Fabian, "Some effects of small-scale metallicity variations in cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 338, no. 4, pp. 824–836, 2003.

[85] A. C. Fabian, S. W. Allen, C. S. Crawford et al., "The missing soft X-ray luminosity in cluster cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 332, no. 3, pp. L50–L54, 2002.

[86] W. G. Mathews and F. Brighenti, "Rapid cooling of dusty gas in elliptical galaxies," *Astrophysical Journal*, vol. 590, no. 1, pp. L5–L8, 2003.

[87] R. Rosner and W. H. Tucker, "On magnetic fields, heating and thermal conduction in halos, and the suppression of cooling flows," *Astrophysical Journal*, vol. 338, pp. 761–769, 1989.

[88] G. Tabor and J. Binney, "Elliptical galaxy cooling flows without mass drop-out," *Monthly Notices of the Royal Astronomical Society*, vol. 263, p. 323, 1993.

[89] E. Churazov, M. Brüggen, C. R. Kaiser, H. Böhringer, and W. Forman, "Evolution of buoyant bubbles in M87," *Astrophysical Journal*, vol. 554, no. 1, pp. 261–273, 2001.

[90] M. Brüggen and C. R. Kaiser, "Buoyant radio plasma in clusters of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 325, no. 2, pp. 676–684, 2001.

[91] C. R. Kaiser and J. Binney, "Simple models of cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 338, no. 4, pp. 837–845, 2003.

[92] M. Ruszkowski and M. C. Begelman, "Heating, conduction, and minimum temperatures in cooling flows," *Astrophysical Journal*, vol. 581, no. 1, pp. 223–228, 2002.

[93] F. Brighenti and W. G. Mathews, "Feedback heating in cluster and galactic cooling flows," *Astrophysical Journal*, vol. 587, no. 2, pp. 580–588, 2003.

[94] W. H. Tucker and R. Rosner, "Thermal conduction and heating by nonthermal electrons in the X-ray halo of M87," *Astrophysical Journal*, vol. 267, pp. 547–550, 1983.

[95] L. M. Voigt, R. W. Schmidt, A. C. Fabian, S. W. Allen, and R. M. Johnstone, "Conduction and cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 1, pp. L7–L11, 2002.

[96] A. C. Fabian, L. M. Voigt, and R. G. Morris, "On conduction, cooling flows and galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 3, pp. L71–L74, 2002.

[97] N. L. Zakamska and R. Narayan, "Models of galaxy clusters with thermal conduction," *Astrophysical Journal*, vol. 582, no. 1, pp. 162–169, 2003.

[98] M. Markevitch, A. Vikhlinin, and P. Mazzotta, "Nonhydrostatic gas in the core of the relaxed galaxy cluster A1795," *Astrophysical Journal*, vol. 562, no. 2, pp. L153–L156, 2001.

[99] A. C. Fabian, "A gravitational contribution to the cooling flow problem," *Monthly Notices of the Royal Astronomical Society*, vol. 344, no. 2, pp. L27–L30, 2003.

[100] W. Domainko, M. Gitti, S. Schindler, and W. Kapferer, "Feedback from intra-cluster supernovae on the ICM in cooling flow galaxy clusters," *Astronomy and Astrophysics*, vol. 425, no. 2, pp. L21–L24, 2004.

[101] P. N. Best, A. von der Linden, G. Kauffmann, T. M. Heckman, and C. R. Kaiser, "On the prevalence of radio-loud active galactic nuclei in brightest cluster galaxies: implications for AGN heating of cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 379, no. 3, pp. 894–908, 2007.

[102] R. Mittal, D. S. Hudson, T. H. Reiprich, and T. Clarke, "AGN heating and ICM cooling in the HIFLUGCS sample of galaxy clusters," *Astronomy and Astrophysics*, vol. 501, no. 3, pp. 835–850, 2009.

[103] E. Churazov, S. Sazonov, R. Sunyaev, W. Forman, C. Jones, and H. Böhringer, "Supermassive black holes in elliptical galaxies: switching from very bright to very dim," *Monthly Notices of the Royal Astronomical Society*, vol. 363, no. 1, pp. L91–L95, 2005.

[104] H. Boehringer, W. Voges, A. C. Fabian, A. C. Edge, and D. M. Neumann, "A ROSAT HRI study of the interaction of the X-ray-emitting gas and radio lobes of NGC 1275," *Monthly Notices of the Royal Astronomical Society*, vol. 264, pp. L25–L28, 1993.

[105] C. L. Carilli, R. A. Perley, and D. E. Harris, "Observations of interaction between cluster gas and the radio lobes of cygnus-A," *Monthly Notices of the Royal Astronomical Society*, vol. 270, pp. 173–177, 1994.

[106] A. C. Fabian, J. S. Sanders, S. Ettori et al., "Chandra imaging of the complex X-ray core of the Perseus cluster," *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 4, pp. L65–L68, 2000.

[107] B. R. McNamara, M. Wise, P. E. J. Nulsen et al., "Chandra X-ray observations of the hydra a cluster: an interaction between the radio source and the X-ray-emitting gas," *Astrophysical Journal*, vol. 534, no. 2, pp. L135–L138, 2000.

[108] E. L. Blanton, C. L. Sarazin, B. R. Mcnamara, and M. W. Wise, "Chandra observation of the radio source/X-ray gas interaction in the cooling flow cluster abell 2052," *Astrophysical Journal*, vol. 558, no. 1, pp. L15–L18, 2001.

[109] E. L. Blanton, C. L. Sarazin, and B. R. McNamara, "Chandra observation of the cooling flow cluster abell 2052," *Astrophysical Journal*, vol. 585, no. 1, pp. 227–243, 2003.

[110] P. E. J. Nulsen, L. P. David, B. R. McNamara, C. Jones, W. R. Forman, and M. Wise, "Interaction of radio lobes with the hot intracluster medium: driving convective outflow in Hydra A," *Astrophysical Journal*, vol. 568, no. 1, pp. 163–173, 2002.

[111] A. Doria, M. Gitti, S. Ettori, F. Brighenti, P. E. J. Nulsen, and B. R. McNamara, "A chandra—VLA investigation of the X-ray cavity system and radio mini-halo in the galaxy cluster RBS 797," submitted to *Astrophysical Journal Letters*.

[112] R. J. H. Dunn and A. C. Fabian, "Investigating AGN heating in a sample of nearby clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 3, pp. 959–971, 2006.

[113] L. Bîrzan, D. A. Rafferty, B. R. McNamara, P. E. J. Nulsen, and M. W. Wise, "The detectability of AGN cavities in cooling-flow clusters," in *Proceedings of the American Institute of Physics Conference Proceedings*, S. Heinz and E. Wilcots, Eds., vol. 1201, pp. 301–304, 2009.

[114] E. Churazov, W. Forman, C. Jones, and H. Böhringer, "Asymmetric, arc minute scale structures around NGC 1275," *Astronomy and Astrophysics*, vol. 356, no. 3, pp. 788–794, 2000.

[115] A. C. Fabian, J. S. Sanders, G. B. Taylor et al., "A very deep Chandra observation of the Perseus cluster: shocks, ripples and conduction," *Monthly Notices of the Royal Astronomical Society*, vol. 366, no. 2, pp. 417–428, 2006.

[116] P. E. J. Nulsen, B. R. Mcnamara, M. W. Wise, and L. P. David, "The cluster-scale AGN outburst in Hydra A," *Astrophysical Journal*, vol. 628, no. 2, pp. 629–636, 2005.

[117] M. W. Wise, B. R. McNamara, P. E. J. Nulsen, J. C. Houck, and L. P. David, "X-ray supercavities in the hydra a cluster and the outburst history of the central galaxy's active nucleus," *Astrophysical Journal*, vol. 659, no. 2, pp. 1153–1158, 2007.

[118] A. Simionescu, E. Roediger, P. E. J. Nulsen et al., "The large-scale shock in the cluster of galaxies Hydra A," *Astronomy and Astrophysics*, vol. 495, no. 3, pp. 721–732, 2009.

[119] C. C. Kirkpatrick, M. Gitti, K. W. Cavagnolo et al., "Direct evidence for outflow of metal-enriched gas along the radio jets of hydra A," *Astrophysical Journal Letters*, vol. 707, no. 1, part 2, pp. L69–L72, 2009.

[120] M. Gitti, P. E. J. Nulsen, L. P. David, B. R. McNamara, and M. W. Wise, "A Chandra study of the large-scale shock and cool filaments in Hydra A: evidence for substantial gas dredge-up by the central outburst," *Astrophysical Journal Letters*, vol. 732, article 13, no. 1, 2011.

[121] H. Bohringer, P. E. J. Nulsen, R. Braun, and A. C. Fabian, "The interaction of the radio halo of M87 with the cooling intracluster medium of the Virgo cluster," *Monthly Notices of the Royal Astronomical Society*, vol. 274, pp. L67–L71, 1995.

[122] S. Molendi, "On the temperature structure of M87," *Astrophysical Journal*, vol. 580, no. 2, pp. 815–823, 2002.

[123] W. Forman, C. Jones, E. Churazov et al., "Filaments, bubbles, and weak shocks in the gaseous atmosphere of M87," *Astrophysical Journal*, vol. 665, no. 2, pp. 1057–1066, 2007.

[124] A. Simionescu, N. Werner, A. Finoguenov, H. Böhringer, and M. Brüggen, "Metal-rich multi-phase gas in M 87 AGN-driven metal transport, magnetic-field supported multi-temperature gas, and constraints on non-thermal emission observed with *XMM-Newton*," *Astronomy and Astrophysics*, vol. 482, no. 1, pp. 97–112, 2008.

[125] E. T. Million, N. Werner, A. Simionesc et al., "Feedback under the microscope—I. Thermodynamic structure and AGN-driven shocks in M87," *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 4, pp. 2046–2062, 2010.

[126] E. L. Blanton, C. L. Sarazin, B. R. Mcnamara, and T. E. Clarke, "Chandra observation of the central region of the cooling

flow cluster A262: a radio source that is a shadow of its former self?" *Astrophysical Journal*, vol. 612, no. 2, pp. 817–824, 2004.

[127] E. L. Blanton, S. W. Randall, E. M. Douglass, C. L. Sarazin, T. E. Clarke, and B. R. McNamara, "Shocks and bubbles in a deep chandra observation of the cooling flow cluster Abell 2052," *Astrophysical Journal*, vol. 697, no. 2, pp. L95–L98, 2009.

[128] E. L. Blanton, S. W. Randall, T. E. Clarke et al., "A very deep *Chandra* observation of abell 2052: bubbles, shocks, and sloshing," *The Astrophysical Journal*, vol. 737, no. 2, article 99, 2011.

[129] S. Schindler, A. Castillo-Morales, E. De Filippis, A. Schwope, and J. Wambsganss, "Discovery of depressions in the X-ray emission of the distant galaxy cluster RBS797 in a CHANDRA observation," *Astronomy and Astrophysics*, vol. 376, no. 3, pp. L27–L30, 2001.

[130] M. Gitti, L. Feretti, and S. Schindler, "Multifrequency VLA radio observations of the X-ray cavity cluster of galaxies RBS797: evidence of differently oriented jets," *Astronomy and Astrophysics*, vol. 448, no. 3, pp. 853–860, 2006.

[131] K. W. Cavagnolo, B. R. McNamara, M. W. Wise et al., "A powerful AGN outburst in RBS 797," *Astrophysical Journal Letters*, vol. 732, article 71, no. 2, 2011.

[132] Y. Fujita, C. L. Sarazin, J. C. Kempner et al., "Chandra observations of the disruption of the cool core in A133," *Astrophysical Journal*, vol. 575, no. 2, pp. 764–778, 2002.

[133] S. W. Randall, T. E. Clarke, P. E. J. Nulsen et al., "Radio and deep Chandra observations of the disturbed cool core cluster Abell 133," *Astrophysical Journal*, vol. 722, no. 1, pp. 825–846, 2010.

[134] T. E. Clarke, E. L. Blanton, C. L. Sarazin et al., "Tracing multiple generations of active galactic nucleus feedback in the core of abell 262," *Astrophysical Journal*, vol. 697, no. 2, pp. 1481–1492, 2009.

[135] B. R. McNamara, P. E. J. Nulsen, M. W. Wise et al., "The heating of gas in a galaxy cluster by X-ray cavities and large-scale shock fronts," *Nature*, vol. 433, no. 7021, pp. 45–47, 2005.

[136] M. Gitti, B. R. McNamara, P. E. J. Nulsen, and M. W. Wise, "Cosmological effects of powerful AGN outbursts in galaxy clusters: insights from an *XMM-Newton* observation of MS 0735+7421," *Astrophysical Journal*, vol. 660, no. 2, pp. 1118–1136, 2007.

[137] L. Bîrzan, D. A. Rafferty, B. R. McNamara, M. W. Wise, and P. E. J. Nulsen, "A systematic study of radio-induced X-ray cavities in clusters, groups, and galaxies," *Astrophysical Journal*, vol. 607, no. 2, pp. 800–809, 2004.

[138] R. J. H. Dunn, A. C. Fabian, and G. B. Taylor, "Radio bubbles in clusters of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 364, no. 4, pp. 1343–1353, 2005.

[139] D. A. Rafferty, B. R. Mcnamara, P. E. J. Nulsen, and M. W. Wise, "The feedback-regulated growth of black holes and bulges through gas accretion and starbursts in cluster central dominant galaxies," *Astrophysical Journal*, vol. 652, no. 1, pp. 216–231, 2006.

[140] S. Diehl, H. Li, C. L. Fryer, and D. Rafferty, "Constraining the nature of X-ray cavities in clusters and galaxies," *Astrophysical Journal*, vol. 687, no. 1, pp. 173–192, 2008.

[141] D. E. Harris, A. Finoguenov, A. H. Bridle, M. J. Hardcastle, and R. A. Laing, "X-ray detection of the inner jet in the radio galaxy M84," *Astrophysical Journal*, vol. 580, no. 1, pp. 110–113, 2002.

[142] A. Finoguenov, M. Ruszkowski, C. Jones, M. Brüggen, A. Vikhlinin, and E. Mandel, "In-depth Chandra study of the AGN feedback in virgo elliptical galaxy M84," *Astrophysical Journal*, vol. 686, no. 2, pp. 911–917, 2008.

[143] C. Jones, W. Forman, A. Vikhlinin et al., "Chandra observations of NGC 4636-an elliptical galaxy in turmoil," *Astrophysical Journal*, vol. 567, no. 2, pp. L115–L118, 2002.

[144] E. O'Sullivan, J. M. Vrtilek, and J. C. Kempner, "Active galactic nucleus feedback and gas mixing in the core of NGC 4636," *Astrophysical Journal*, vol. 624, no. 2, pp. L77–L80, 2005.

[145] A. Baldi, W. Forman, C. Jones et al., "The unusual X-ray morphology of NGC 4636 revealed by deep chandra observations: cavities and shocks created by past active galactic nucleus outbursts," *Astrophysical Journal*, vol. 707, no. 2, pp. 1034–1043, 2009.

[146] D. A. Buote, A. D. Lewis, F. Brighenti, and W. G. Mathews, "*XMM-Newton* and Chandra observations of the galaxy group NGC 5044. I. Evidence for limited multiphase hot gas," *Astrophysical Journal*, vol. 594, no. 2, pp. 741–757, 2003.

[147] F. Gastaldello, D. A. Buote, P. Temi, F. Brighenti, W. G. Mathews, and S. Ettori, "X-ray cavities, filaments, and cold fronts in the core of the galaxy group NGC 5044," *Astrophysical Journal Letters*, vol. 693, pp. 43–55, 2009.

[148] L. P. David, C. Jones, W. Forman et al., "Isotropic active galactic nucleus heating with small radio-quiet bubbles in the ngc 5044 group," *Astrophysical Journal*, vol. 705, no. 1, pp. 624–638, 2009.

[149] L. P. David, E. O'Sullivan, C. Jones et al., "Active-galactic-nucleus-driven weather and multiphase gas in the core of the NGC 5044 galaxy group," *Astrophysical Journal Letters*, vol. 728, article 162, no. 2, 2011.

[150] J. M. Vrtilek, L. Grego, L. P. David et al., "A Sharper Picture of X-Ray Bright Galaxy Groups: Chandra Imaging and Spectroscopy of HCG 62 and NGC 741," in *Proceedings of the American Physical Society Meeting*, Aprril 2002.

[151] U. Morita, Y. Ishisaki, N. Y. Yamasaki et al., "Chandra and XMM-Newton observations of a group of galaxies, HCG 62," *Publications of the Astronomical Society of Japan*, vol. 58, no. 4, pp. 719–742, 2006.

[152] J. Gu, H. Xu, L. Gu et al., "A high-abundance arc in the compact group of galaxies HCG 62: an AGN- or merger-induced metal outflow?" *Astrophysical Journal*, vol. 659, no. 1, pp. 275–282, 2007.

[153] M. Gitti, E. O'Sullivan, S. Giacintucci et al., "Cavities and shocks in the galaxy group HCG 62 as revealed by Chandra, *XMM-Newton*, and giant metrewave radio telescope data," *Astrophysical Journal Letters*, vol. 714, no. 1, pp. 758–771, 2010.

[154] M. E. Machacek, D. Jerius, R. P. Kraft et al., "Deep *Chandra* observations of edges and bubbles in the NGC 5846 galaxy group," *The Astrophysical Journal*, vol. 743, no. 1, article 15, 2011.

[155] V. R. Eke, C. S. Frenk, C. M. Baugh et al., "Galaxy groups in the two-degree field galaxy redshift survey: the luminous content of the groups," *Monthly Notices of the Royal Astronomical Society*, vol. 355, no. 3, pp. 769–784, 2004.

[156] S. Giodini, V. Smolić, A. Finoguenov et al., "Radio galaxy feedback in x-ray-selected groups from cosmos: the effect on the intracluster medium," *Astrophysical Journal*, vol. 714, no. 1, pp. 218–228, 2010.

[157] S. W. Randall, W. R. Forman, S. Giacintucci et al., "Shocks and cavities from multiple outbursts in the galaxy group NGC 5813: a window to active galactic nucleus feedback," *Astrophysical Journal Letters*, vol. 726, article 86, no. 2, 2011.

[158] E. O'Sullivan, S. Giacintucci, L. P. David, J. M. Vrtilek, and S. Raychaudhury, "A deep Chandra observation of the poor cluster AWM 4—I. Properties of the central radio galaxy and its effects on the intracluster medium," *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 1, pp. 321–338, 2010.

[159] E. O'Sullivan, S. Giacintucci, L. P. David, J. M. Vrtilek, and S. Raychaudhury, "A deep Chandra observation of the poor cluster AWM 4—II. The role of the radio jets in enriching the intracluster medium," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 3, pp. 1833–1842, 2011.

[160] R. Johnson, T. J. Ponman, and A. Finoguenov, "A statistical analysis of the Two-Dimensional *XMM-Newton* Group Survey: the impact of feedback on group properties," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1287–1308, 2009.

[161] M. Sun, "Every BCG with a strong radio AGN has an x-ray cool core: is the cool core-noncool core dichotomy too simple?" *Astrophysical Journal*, vol. 704, no. 2, pp. 1586–1604, 2009.

[162] R. Dong, J. Rasmussen, and J. S. Mulchaey, "A systematic search for X-ray cavities in the hot gas of galaxy groups," *Astrophysical Journal*, vol. 712, no. 2, pp. 883–900, 2010.

[163] M. Gaspari, F. Brighenti, A. D'Ercole, and C. Melioli, "AGN feedback in galaxy groups: the delicate touch of self-regulated outflows," *Monthly Notices of the Royal Astronomical Society*, vol. 415, no. 2, pp. 1549–1568, 2011.

[164] W. M. Lane, T. E. Clarke, G. B. Taylor, R. A. Perley, and N. E. Kassim, "Hydra a at low radio frequencies," *Astronomical Journal*, vol. 127, no. 1, pp. 48–52, 2004.

[165] E. O'Sullivan, S. Giacintucci, L. P. David et al., "Heating the hot atmospheres of galaxy groups and clusters with cavities: the relationship between jet power and low-frequency radio emission," *Astrophysical Journal Letters*, vol. 735, article 11, no. 1, 2011.

[166] W. G. Mathews and F. Brighenti, "Creation of the X-ray cavity jet and its radio lobe in M87/Virgo with cosmic rays: relevance to relic radio sources," *Astrophysical Journal*, vol. 676, no. 2, pp. 880–888, 2008.

[167] W. G. Mathews and F. Brighenti, "Energetics of X-ray cavities and radio lobes in galaxy clusters," *Astrophysical Journal*, vol. 685, no. 1, pp. 128–137, 2008.

[168] P. J. Mendygral, S. M. O'Neill, and T. W. Jones, "Synthetic observations of simulated active galactic nucleus jets: X-ray cavities," *Astrophysical Journal Letters*, vol. 730, article 100, no. 2, 2011.

[169] K. W. Cavagnolo, B. R. McNamara, P. E. J. Nulsen, C. L. Carilli, C. Jones, and L. Bîrzan, "A relationship between agn jet power and radio power," *Astrophysical Journal*, vol. 720, no. 2, pp. 1066–1072, 2010.

[170] C. Nipoti and J. Binney, "Time variability of active galactic nuclei and heating of cooling flows," *Monthly Notices of the Royal Astronomical Society*, vol. 361, no. 2, pp. 428–436, 2005.

[171] R. J. H. Dunn and A. C. Fabian, "Investigating heating and cooling in the BCS and B55 cluster samples," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 2, pp. 757–768, 2008.

[172] E. C. D. Pope, G. Pavlovski, C. R. Kaiser, and H. Fangohr, "Heating rate profiles in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 3, pp. 1121–1131, 2006.

[173] R. J. H. Dunn, W. Allen, G. B. Taylor et al., "The radio properties of a complete, X-ray selected sample of nearby, massive elliptical galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 404, no. 1, pp. 180–197, 2010.

[174] F. Brighenti and W. G. Mathews, "Heated cooling flows," *Astrophysical Journal*, vol. 573, no. 2, pp. 542–561, 2002.

[175] C. J. Willott, S. Rawlings, K. M. Blundell, and M. Lacy, "The emission line-radio correlation for radio sources using the 7C Redshift Survey," *Monthly Notices of the Royal Astronomical Society*, vol. 309, no. 4, pp. 1017–1033, 1999.

[176] A. G. Pacholczyk, *Radio Astrophysics. Nonthermal Processes in Galactic and Extragalactic Sources*, 1970.

[177] G. Brunetti, "Non-thermal emission from extragalactic radio sources: a high resolution broad band (radio to X-rays) approach," http://arxiv.org/abs/astro-ph/0207671.

[178] G. Brunetti, G. Setti, and A. Comastri, "Inverse Compton X-rays from strong FRII radio-galaxies," *Astronomy and Astrophysics*, vol. 325, no. 3, pp. 898–910, 1997.

[179] D. S. De Young, "The particle content of extragalactic jets," *Astrophysical Journal*, vol. 648, no. 1, pp. 200–208, 2006.

[180] J. H. Croston, M. J. Hardcastle, M. Birkinshaw, D. M. Worrall, and R. A. Laing, "An *XMM-Newton* study of the environments, particle content and impact of low-power radio galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 386, no. 3, pp. 1709–1728, 2008.

[181] R. J. H. Dunn and A. C. Fabian, "Particle energies and filling fractions of radio bubbles in cluster cores," *Monthly Notices of the Royal Astronomical Society*, vol. 355, no. 3, pp. 862–873, 2004.

[182] R. J. H. Dunn, A. C. Fabian, and A. Celotti, "Using radio bubbles to constrain the matter content of AGN jets," *Monthly Notices of the Royal Astronomical Society*, vol. 372, no. 4, pp. 1741–1748, 2006.

[183] M. Gitti, G. Brunetti, L. Feretti, and G. Setti, "Particle acceleration in cooling flow clusters of galaxies: the case of Abell 2626," *Astronomy and Astrophysics*, vol. 417, no. 1, pp. 1–11, 2004.

[184] M. Gitti, C. Ferrari, W. Domainko, L. Feretti, and S. Schindler, "Discovery of diffuse radio emission at the center of the most X-ray-luminous cluster RX J1347.5-1145," *Astronomy and Astrophysics*, vol. 470, no. 3, pp. L25–L28, 2007.

[185] H. Böhringer, V. Burwitz, Y. Y. Zhang, P. Schuecker, and N. Nowak, "Chandra reveals galaxy cluster with the most massive nearby cooling core: RXC J1504.1-0248," *Astrophysical Journal*, vol. 633, no. 1, pp. 148–153, 2005.

[186] S. Giacintucci, M. Markevitch, G. Brunetti, R. Cassano, and T. Venturi, "A radio minihalo in the extreme cool-core galaxy cluster RXC J1504.1-0248," *Astronomy and Astrophysics*, vol. 525, no. 7, article L10, 2011.

[187] M. Gitti, G. Brunetti, and G. Setti, "Modeling the interaction between ICM and relativistic plasma in cooling flows: the case of the Perseus cluster," *Astronomy and Astrophysics*, vol. 386, no. 2, pp. 456–463, 2002.

[188] J. O. Burns, M. E. Sulkanen, G. R. Gisler, and R. A. Perley, "Where have all the cluster halos gone?" *Astrophysical Journal*, vol. 388, no. 2, pp. L49–L52, 1992.

[189] E. Rizza, C. Loken, M. Bliton, K. Roettiger, J. O. Burns, and F. N. Owen, "X-ray and radio interactions in the cores of cooling flow clusters," *Astronomical Journal*, vol. 119, no. 1, pp. 21–31, 2000.

[190] G. Giovannini and L. Feretti, "Halo and relic sources in clusters of galaxies," *New Astronomy*, vol. 5, no. 6, pp. 335–347, 2000.

[191] S. A. Baum and C. P. O'Dea, "Multifrequency VLA observations of PKS 0745—191—the archetypal 'cooling flow' radio source?" *Monthly Notices of the Royal Astronomical Society*, vol. 250, pp. 737–749, 1991.

[192] M. Bacchi, L. Feretti, G. Giovannini, and F. Govoni, "Deep images of cluster radio halos," *Astronomy and Astrophysics*, vol. 400, no. 2, pp. 465–476, 2003.

[193] T. Venturi, S. Giacintucci, D. Dallacasa et al., "GMRT radio halo survey in galaxy clusters at z = 0.2-0.4 II. the eBCS clusters and analysis of the complete sample," *Astronomy and Astrophysics*, vol. 484, no. 2, pp. 327–340, 2008.

[194] F. Govoni, M. Murgia, M. Markevitch et al., "A search for diffuse radio emission in the relaxed, cool-core galaxy clusters A1068, A1413, A1650, A1835, A2029, and Ophiuchus," *Astronomy and Astrophysics*, vol. 499, no. 2, pp. 371–383, 2009.

[195] M. Murgia, D. Eckert, F. Govoni et al., "GMRT observations of the Ophiuchus galaxy cluster," *Astronomy and Astrophysics*, vol. 514, no. 12, article 76, 2010.

[196] C. Pfrommer and T. A. Enßlin, "Constraining the population of cosmic ray protons in cooling flow clusters with γ-ray and radio observations: are radio mini-halos of hadronic origin?" *Astronomy and Astrophysics*, vol. 413, no. 1, pp. 17–36, 2004.

[197] U. Keshet and A. Loeb, "Using radio halos and minihalos to measure the distributions of magnetic fields and cosmic rays in galaxy clusters," *Astrophysical Journal*, vol. 722, no. 1, pp. 737–749, 2010.

[198] R. Cassano, M. Gitti, and G. Brunetti, "A morphological comparison between giant radio halos and radio mini-halos in galaxy clusters," *Astronomy and Astrophysics*, vol. 486, no. 3, pp. L31–L34, 2008.

[199] P. Mazzotta and S. Giacintucci, "Do radio core-halos and cold fronts in non-major-merging clusters originate from the same gas sloshing?" *Astrophysical Journal*, vol. 675, no. 1, pp. L9–L12, 2008.

[200] J. ZuHone, M. Markevitch, and G. Brunetti, "Testing the connection between radio mini-halos and core gas sloshing with MHD simulations," http://arxiv.org/abs/1101.4627.

[201] S. Heinz, "The interaction of relativistic jets with their environment," *New Astronomy Reviews*, vol. 47, no. 6-7, pp. 565–567, 2003.

[202] F. Brighenti and W. G. Mathews, "Stopping cooling flows with jets," *Astrophysical Journal*, vol. 643, no. 1, pp. 120–127, 2006.

[203] M. Gaspari, C. Melioli, F. Brighenti, and A. D'Ercole, "The dance of heating and cooling in galaxy clusters: three-dimensional simulations of self-regulated active galactic nuclei outflows," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 1, pp. 349–372, 2011.

[204] R. P. Kraft, P. E. J. Nulsen, M. Birkinshaw et al., "A Chandra study of the lobe/interstellar medium interactions around the inner radio lobes of centaurus A: constraints on the temperature structure and transport processes," *Astrophysical Journal*, vol. 665, no. 2, pp. 1129–1137, 2007.

[205] J. H. Croston, R. P. Kraft, M. J. Hardcastle et al., "High-energy particle acceleration at the radio-lobe shock of Centaurus A," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 4, pp. 1999–2012, 2009.

[206] P. E. J. Nulsen, D. C. Hambrick, B. R. McNamara et al., "The powerful outburst in Hercules A," *Astrophysical Journal*, vol. 625, no. 1, pp. L9–L12, 2005.

[207] K. K. S. Wu, A. C. Fabian, and P. E. J. Nulsen, "Non-gravitational heating in the hierarchical formation of X-ray clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 3, pp. 889–912, 2000.

[208] E. Puchwein, D. Sijacki, and V. Springel, "Simulations of AGN feedback in galaxy clusters and groups: impact on gas fractions and the L-T scaling relation," *Astrophysical Journal*, vol. 687, pp. L53–L56, 2008.

[209] M. Gitti, B. R. McNamara, P. E. J. Nulsen, and M. W. Wise, "Cosmological effects of powerful AGN outbursts in galaxy clusters: insights from an XMM-Newton observation of MS 0735+7421," *Astrophysical Journal*, vol. 660, no. 2, pp. 1118–1136, 2007.

[210] A. E. Evrard, C. A. Metzler, and J. F. Navarro, "Mass estimates of X-ray clusters," *Astrophysical Journal*, vol. 469, no. 2, pp. 494–507, 1996.

[211] D. Lynden-Bell, "Galactic nuclei as collapsed old quasars," *Nature*, vol. 223, no. 5207, pp. 690–694, 1969.

[212] M. C. Begelman, R. D. Blandford, and M. J. Rees, "Theory of extragalactic radio sources," *Reviews of Modern Physics*, vol. 56, no. 2, pp. 255–351, 1984.

[213] R. Narayan and J. E. McClintock, "Advection-dominated accretion and the black hole event horizon," *New Astronomy Reviews*, vol. 51, no. 10–12, pp. 733–751, 2008.

[214] A. J. Barger, L. L. Cowie, M. W. Bautz et al., "Supermassive black hole accretion history inferred from a large sample of Chandra hard X-ray sources," *Astronomical Journal*, vol. 122, no. 5, pp. 2177–2189, 2001.

[215] J. Hlavacek-Larrondo and A. C. Fabian, "Investigating a sample of strong cool core, highly luminous clusters with radiatively inefficient nuclei," *Monthly Notices of the Royal Astronomical Society*, vol. 413, pp. 313–321, 2011.

[216] B. R. McNamara, F. Kazemzadeh, D. A. Rafferty et al., "An energetic AGN outburst powered by a rapidly spinning supermassive black hole or an accreting ultramassive black hole," *Astrophysical Journal*, vol. 698, no. 1, pp. 594–605, 2009.

[217] E. C. D. Pope, A. Babul, G. Pavlovski, R. G. Bower, and A. Dotter, "Mass transport by buoyant bubbles in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 406, no. 3, pp. 2023–2037, 2010.

[218] C. S. Reynolds, E. A. Casper, and S. Heinz, "A deep Chandra observation of Abell 4059: a new face to "radio-mode" AGN feedback?" *Astrophysical Journal*, vol. 679, no. 2, pp. 1181–1191, 2008.

[219] A. Simionescu, N. Werner, H. Böhringer et al., "Chemical enrichment in the cluster of galaxies Hydra A," *Astronomy and Astrophysics*, vol. 493, no. 2, pp. 409–424, 2009.

[220] N. Werner, A. Simionescu, E. T. Million et al., "Feedback under the microscope-II. Heating, gas uplift and mixing in the nearest cluster core," *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 4, pp. 2063–2074, 2010.

[221] C. C. Kirkpatrick, B. R. McNamara, and K. W. Cavagnolo, "Anisotropic metal-enriched outflows driven by active galactic nuclei in clusters of galaxies," *Astrophysical Journal Letters*, vol. 731, article L23, no. 2, 2011.

[222] C. S. Reynolds, S. Heinz, and M. C. Begelman, "The hydrodynamics of dead radio galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 332, no. 2, pp. 271–282, 2002.

[223] M. Ruszkowski, M. Brüggen, and M. C. Begelman, "Cluster heating by viscous dissipation of sound waves," *Astrophysical Journal*, vol. 611, no. 1, pp. 158–163, 2004.

[224] T. W. Jones and D. S. De Young, "Magnetohydrodynamic simulations of relic radio bubbles in clusters," *Astrophysical Journal*, vol. 624, no. 2, pp. 586–605, 2005.

[225] N. Soker and F. Pizzolato, "Feedback heating with slow jets in cooling flow clusters," *Astrophysical Journal*, vol. 622, no. 2, pp. 847–852, 2005.

[226] S. Heinz, M. Briiggen, A. Young, and E. Levesque, "The answer is blowing in the wind: simulating the interaction of jets with dynamic cluster atmospheres," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 1, pp. L65–L69, 2006.

[227] M. Brüggen, S. Heinz, E. Roediger, M. Ruszkowski, and A. Simionescu, "Shock heating by Fanaroff-Riley type I radio sources in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 1, pp. L67–L70, 2007.

[228] A. Sternberg, F. Pizzolato, and N. Soker, "Inflating fat bubbles in clusters of galaxies by wide jets," *Astrophysical Journal*, vol. 656, no. 1, pp. L5–L8, 2007.

[229] M. Brüggen and E. Scannapieco, "Self-regulation of active galactic nuclei in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 2, pp. 548–560, 2009.

[230] B. J. Morsony, S. Heinz, M. Brüggen, and M. Ruszkowski, "Swimming against the current: simulations of central AGN evolution in dynamic galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 2, pp. 1277–1289, 2010.

[231] M. Ruszkowski and S. P. Oh, "Galaxy motions, turbulence and conduction in clusters of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 2, pp. 1493–1507, 2011.

[232] H. Omma, J. Binney, G. Bryan, and A. Slyz, "Heating cooling flows with jets," *Monthly Notices of the Royal Astronomical Society*, vol. 348, no. 4, pp. 1105–1119, 2004.

[233] C. Zanni, G. Murante, G. Bodo, S. Massaglia, P. Rossi, and A. Ferrari, "Heating groups and clusters of galaxies: the role of AGN jets," *Astronomy and Astrophysics*, vol. 429, no. 2, pp. 399–415, 2005.

[234] J. C. Vernaleo and C. S. Reynolds, "AGN feedback and cooling flows: problems with simple hydrodynamic models," *Astrophysical Journal*, vol. 645, no. 1, pp. 83–94, 2006.

[235] A. Cattaneo and R. Teyssier, "AGN self-regulation in cooling flow clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 376, no. 4, pp. 1547–1556, 2007.

[236] M. Brüggen, "The effect of mixing on metallicity gradients in the intracluster medium," *Astrophysical Journal*, vol. 571, no. 1, pp. L13–L16, 2002.

[237] E. Roediger, M. Brüggen, P. Rebusco, H. Böhringer, and E. Churazov, "Metal mixing by buoyant bubbles in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 375, no. 1, pp. 15–28, 2007.

Hindawi Publishing Corporation
Advances in Astronomy
Volume 2012, Article ID 853701, 19 pages
doi:10.1155/2012/853701

Review Article

Clustering of X-Ray-Selected AGN

N. Cappelluti,[1,2] V. Allevato,[3] and A. Finoguenov[2,4]

[1] *Osservatorio Astronomico di Bologna, INAF, Via Ranzani 1, 40127 Bologna, Italy*
[2] *University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA*
[3] *Max-Planck-Institut für Plasmaphysik and Excellence Cluster Universe, Boltzmannstrasse 2, 85748 Garching, Germany*
[4] *Max-Planck-Institute für Extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany*

Correspondence should be addressed to N. Cappelluti, nico.cappelluti@oabo.inaf.it

Received 30 August 2011; Revised 1 December 2011; Accepted 26 January 2012

Academic Editor: Angela Bongiorno

The study of the angular and spatial structure of the X-ray sky has been under investigation since the times of the *Einstein* X-ray Observatory. This topic has fascinated more than two generations of scientists and slowly unveiled an unexpected scenario regarding the consequences of the angular and spatial distribution of X-ray sources. It was first established from the clustering of sources making the CXB that the source spatial distribution resembles that of optical QSO. It then became evident that the distribution of X-ray AGN in the Universe was strongly reflecting that of Dark Matter. In particular, one of the key results is that X-ray AGNs are hosted by dark matter halos of mass similar to that of galaxy groups. This result, together with model predictions, has lead to the hypothesis that galaxy mergers may constitute the main AGN-triggering mechanism. However, detailed analysis of observational data, acquired with modern telescopes, and the use of the new halo occupation formalism has revealed that the triggering of an AGN could also be attributed to phenomena-like tidal disruption or disk instability and to galaxy evolution. This paper reviews results from 1988 to 2011 in the field of X-ray-selected AGN clustering.

1. Introduction

After about 50 years from the opening of the X-ray window on the Universe with the discovery of Sco-X1 and the Cosmic X-ray background (CXB, [1]), our knowledge of high-energy processes in the Universe has dramatically improved. One of the leading mechanisms for the production of X-ray in the Universe is accretion onto compact objects. For this reason, the study of astrophysical X-ray sources is a powerful tool for studying matter under the effects of extreme gravity. As the efficiency of converting matter into energy in accretion processes is proportional to the "compactness" of the object (i.e., $\propto$ M/R), it is clear that the strongest sources powered by accretion are super-massive black holes (SMBH). It also became a cornerstone of astrophysics that every galaxy with a bulge-like component hosts a SMBH at its centre and that the BH mass and the bulge velocity dispersion are strictly related [2]. It is also believed that black holes reach those high masses via one or more phases of intense accretion activity and therefore shining as active galactic nuclei (AGN). It is believed that an AGN basically shines mostly from the power emitted by a thin, viscous, accretion disk orbiting the central SMBH Shakura and Sunyaev [3]. Such a disk produces a high amount of X-rays both from its hot inner regions (as far as the soft X-ray emission is concerned) and from a nonthermal source which is supposed to be the primary source of X-rays (both soft and hard).

Since its discovery, the nature of the CXB has been strongly debated, but soon the community converged into interpreting most of the CXB as the integrated emission of AGN across the cosmic time. While the discrete nature of the CXB has been proposed [4] and rapidly unveiled by experiments like *Einstein* [1] and ROSAT (see, e.g., [5]), little cosmological information has been obtained from samples of AGN because of the scarce number of detected sources in the X-ray band. Structure formation models and numerical simulations have shown that structures in the Universe have undergone a hierarchical growth starting from the denser peaks in the primordial Gaussian matter distribution. The large-scale structures (LSS) of the Universe

are gravitationally dominated by dark matter (DM), and we can consider it as the responsible and one of the main drivers of the Cosmological structures evolution. Dark matter is believed to clump in large-scale halos (DMH Navarro, [6]) which are populated by galaxies. Thus, galaxies can be considered as tracers of the DM distribution in the Universe, and the study of their spatial clustering led us to a most comprehensive view of the LSS. On the other hand AGN/Quasar, as phase of the galactic evolution, is a quite rare phenomenon in the Universe as their space density of these objects is about 1/100–1/1000 lower than that of galaxies. This means that AGN/Quasar survey requires large field of view and/or deep exposure to provide statistically significant samples.

The study of their clustering and its evolution is a powerful tool to understand, from a statistical point of view, what kind of environment is more likely to host AGN. This is not just an academic question, but this is strictly related to the mechanism of AGN activation. We know that one of the candidate mechanisms for triggering an AGN is galaxy merger (see, e.g., [7–10]). The probability of such an event is definitely dependent on the environment inhabited by the host galaxy. Even if the mean distance between galaxies is relatively small, in high-density (mass) environments, they have a high velocity dispersion, and, therefore, the likelihood of a major merger is very low. On the contrary, in the field, the likelihood of galaxy mergers is low because of the large average distance between galaxies. The most favorable place to detect a merger is therefore a moderately low-density (mass) environment like a group (see, e.g., [11]).

In fact, merger-driven models (see, e.g., [7]) accurately predict the observed large-scale clustering of quasars as a function of redshift up to $z\sim4$. The clustering is precisely that predicted for small group halos in which major mergers of gas-rich galaxies should proceed most efficiently. Thus, it is well established empirically and with theoretical predictions that quasar clustering traces a characteristic host halo mass $\sim 4 \times 10^{12} h^{-1} M_{\odot}$, supporting the scenario in which major mergers dominate the bright quasar populations.

In addition, other phenomena like secular processes may become dominant at lower luminosities as suggested by Milosavljević et al. [12]; Hopkins et al. [13]; Hopkins and Henquist [9]. Low-luminosity AGN could be triggered in more common nonmerger events, like stochastic encounters of the black holes and molecular clouds, tidal disruption, or disk instability. This leads to the expectation of a characteristic transition to merger-induced fueling around the traditional quasar-Seyfert luminosity division (growth of BH masses above/below $\sim 10^7 M_{\odot}$). However, the triggering mechanism of the SMBH growth must be compliant with M_{BH}-σ relation that links the growth of the SMBH with growth of the bulge of the host galaxy [2].

As shown in Hopkins et al. [8], the predicted large-scale bias of quasars triggered by secular processes is, at all redshifts, lower than the bias estimated for quasars fueled by major mergers. This implies that low-luminosity Seyfert galaxies live in DMHs that never reach the characteristic mass associated with small group scales.

On the other hand, the majority of the results on the clustering of X-ray-selected AGN suggest a picture where moderate-luminosity AGN live in massive DMHs ($12.5 < \log M_{\mathrm{DMH}}\,[h^{-1}\,M_{\odot}] < 13.5$) up to $z \sim 2$, that is, X-ray-selected AGN samples appear to cluster more strongly than bright quasars. The reason for this is not completely clear, but several studies argued that these large bias and DMH masses could suggest a different AGN-triggering mechanism respect to bright quasars characterized by galaxy merger-induced fueling.

This paper reviews results of clustering of X-ray-selected AGN from the first *Einstein* to the most recent *Chandra* and XMM-*Newton* surveys. We give a detailed description of the methods used in this kind of analysis from simple power-law to halo models. In addition, we discuss the results of X-ray AGN clustering in the framework of AGN evolution and triggering. We adopt a ΛCDM cosmology with $\Omega_{\Lambda} = 0.7$, $\Omega_m = 0.3$, $\mathrm{H}_0 = 100\,\mathrm{h}^{-1}$ km/s/Mpc with h = 0.7 and $\sigma_8 = 0.8$ ([14], WMAP-7).

2. Previous Measures of X-Ray Clustering Amplitude

As far as the X-ray source clustering results are concerned, the development of the field has always been driven by the performance of the telescopes. In particular, while first results studied the angular distribution of the unresolved CXB under the assumption that Quasars were its main contributors, recent *Chandra* and XMM-*Newton* surveys sample clustering of AGN with a precision comparable to that achievable with redshift galaxy surveys.

In the following section, we will use the following convention for reporting results of clustering analysis in the case of power-law representation of the auto(cross)-correlation function: if the clustering is measured in the angular space, we will use

$$w(\theta) = \left(\frac{\theta}{\theta_0}\right)^{1-\gamma}, \tag{1}$$

where θ_0 is the angular correlation length. If the measurements have been performed in the real (redshift) space this becomes

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}, \qquad \left(\xi(s) = \left(\frac{s}{s_0}\right)^{-\gamma}, \quad \text{in } z-\text{space}\right), \tag{2}$$

where γ is the 3D correlation slope and r_0 or s_0 are the correlation lengths. Barcons and Fabian [15] measured with *Einstein* a clustering signal of the CXB on scales $\leq 5'$ corresponding to an angular correlation length $\theta_0 \sim 4'$. They have shown the importance of studying the angular structure of the CXB by pointing out that a large fraction of the CXB could have been attributed to sources with a redshift distribution similar to optical QSOs. In addition, the first prediction was not consistent with the hypothesis that the CXB was also partly produced by a diffuse hot intergalactic medium (IGM) component. It was also proposed that these

sources were actually clustered on comoving scales of the order of $\sim 10\,h^{-1}$ Mpc.

Carrera and Barcons [16], Georgantopoulos et al. [17], and Soltan and Hasinger [18] observed that the CXB was highly isotropic on scales of the order of 2°–25°. The first attempt of measuring the clustering of X-ray-selected AGN was performed by Boyle and Mo [19] that measured a barely significant signal by using a sample of 183 EMSS sources, mostly local AGN ($z < 0.2$). These evidences have brought the attention to the study of the clustering of the CXB down to the arcminute scale. The first significant upward turn for the measurement of AGN clustering in the X-ray band has been brought to light by ROSAT. By using a set of ROSAT-PSPC pointing on an area of $\sim 40\,\mathrm{deg}^2$, Vikhlinin and Forman [20] measured, for the first time, an angular correlation signal of faint (ROSAT) X-ray sources on scales $<10'$. By using the Limber equation (see Appendix B and [21]) they have deprojected their angular correlation function into a real-space correlation function and found that, under the assumption that the redshift distribution of the sources was the same as that of optical QSOs, the spatial correlation length was in the range 6–10 h^{-1} Mpc. With such a result, they confirmed the hypothesis that the CXB was mostly produced by sources with a redshift distribution comparable to that of optically selected QSO, though with almost double source density. By using the results of Vikhlinin and Forman [20] and Akylas et al. [22] (who obtained similar results), Barcons et al. [23] have shown for the first time that X-ray-selected AGNs are highly biased tracers of the underlying LSS at $z < 1$ by showing a redshift evolving bias factor as large as $b \sim 2$.

However, it is worth to consider that the deprojection of the angular correlation function into a 3D correlation relies on several assumptions, like the model-dependent expected redshift distribution, which may lead to a biased estimate of the real-space clustering. It is, however, worth noticing that angular correlation can be very useful to provide a first overview in the early phase of surveys, when optical identifications are not available, especially sampling new part of the parameter space of sources, like that is, new unexplored luminosity/flux limits and therefore source classes. Detailed physical models are, however, much better investigated by more sophisticated techniques as shown in the following parts.

The first firm detection of 3D spatial clustering of X-ray-selected AGN has been claimed by Mullis et al. [24] by using data of the ROSAT-NEP survey. They detected on an area of $\sim 81\,\mathrm{deg}^2$ a 3σ significant signal in the redshift space autocorrelation function of soft X-ray-selected sources at $\langle z \rangle \sim 0.22$. They have shown that, at that redshift AGN cluster with a typical correlation length, $r_0 = 7.4 \pm 1.9\,h^{-1}$ Mpc. Their results suggest that the population of AGN in such a sample is consistent with an unbiased population with respect to the underlying matter. Their result suggested that, at that redshift, AGNs were hosted in DMHs of mass of the order of $10^{13}\,h^{-1}\,M_\odot$.

With the development of *Chandra* and XMM-*Newton* surveys and thanks to the high source surface densities (i.e., >400–1000 deg^{-2}), our capabilities in tracing the LSS have dramatically increased. One of the first evidences that AGNs are highly correlated with the underlying LSS has been pointed out by Cappi et al. [25] and Cappelluti et al. [26] and references therein, who showed that, around massive high-z galaxy clusters, the source surface density of *Chandra* point sources is significantly, up to two times, higher than that of the background. More recently, Koulouridis and Plionis [27] showed that, although the X-ray source surface density of AGN around galaxy clusters is larger than in the background, the amplitude of their overdensities is about 4 times lower than that of galaxies in the same fields. This has been interpreted as a clear indication of an environmental influence on the AGN activity. Silverman et al. [10] in the COSMOS field and Koss et al. [28] in the *Swift*-BAT all-sky survey have shown that the AGN fraction in galaxy pairs is higher relative to isolated galaxies of similar stellar mass providing an additional evidence of the influence of the environment on AGN activity.

Chandra and XMM-*Newton* performed several blanck sky extragalactic surveys, and most of them dedicated part of their efforts in the study of the LSS traced by AGN to unveil their coevolution. Basilakos et al. [29, 30] by using data of the XMM-*Newton* 2dF-survey have measured an unexpected high correlation length both in the angular ($\theta_0 \sim 10''$) and, by projection, in the real space ($r_0 \sim 16\,h^{-1}$ Mpc). Such a high correlation length has been detected in this field only, thus one can explain such a measurement as a statistical fluctuation. With the same technique, Gandhi et al. [31] obtained a marginal 2-3σ detection of angular clustering in the XMM-LSS survey and obtained $\theta_0 = 6.3(42) \pm 3(^{+7}_{-13})$ in the 0.5–2 (2–10) keV bands and a slope $\gamma \sim 2.2$. Puccetti et al. [32] measured the clustering of X-ray sources in the XMM-*Newton* ELAIS-S1 survey in the soft and hard energy bands with a sample of 448 sources. They obtained $\theta_0 = 5.2 \pm 3.8\ 4''$ and $\theta_0 = 12.8 \pm 7.8\ 4''$ in the two bands, respectively. These measurements have been deprojected with the Limber's inversion in the real space and obtained $r_0 = 9.8$–$12.8\,h^{-1}$ Mpc and $r_0 = 13.4$–$17.9\,h^{-1}$ Mpc in the two bands, respectively.

In the *Chandra* era, Gilli et al. [33] measured the real space autocorrelation function of point sources in the CDFS-CDFN. They have measured, in the CDFS, $r_0 = 8.6 \pm 1.2\,h^{-1}$ Mpc at $z = 0.73$, while, in the CDFN, they obtained $r_0 = 4.2 \pm 0.4\,h^{-1}$ Mpc. The discrepancy of these measurements has been explained with variance introduced by the relatively small field of view and the consequent random sampling of LSSs in the field. In the CLASXS survey, Yang et al. [34] obtained a measurement of the clustering at $z = 0.94$ with $r_0 = 8.1^{+1.2}_{-2.2}\,h^{-1}$ Mpc which proposes that AGNs are hosted by DMH of mass of $10^{12.1}\,h^{-1}\,M_\odot$ (see Section 3). In addition, they proposed that AGN clustering evolves with luminosity and they found that the bias factor evolves with the redshift. Such a behavior is similar to that found in optically selected quasars. The XMM-*Newton* [35–37] and *Chandra* [38, 39] survey of the COSMOS field have provided a leap forward to the field of X-ray AGN clustering by surveying a $2\,\mathrm{deg}^2$ field of view. The key of the success of this project is a redshift survey *zCOSMOS* [40] performed simultaneously with the X-ray

survey, together with observations in more than 30 energy bands from radio to X-ray that allowed to measure either the spectroscopic or the photometric redshift of every source. In the X-ray band, the survey covers 2 $\deg^2$ with XMM-*Newton* with a depth of ~60 ks with the addition of a central 0.9 $\deg^2$ observed by *Chandra* with ~150 ks exposure. The first sample of ~1500 X-ray sources [36] has been used by Miyaji et al. [41] to determine their angular correlation function, without knowing their distance, and just assuming a theoretical redshift distribution for the purpose of Limber's deprojection. Significant positive signals have been detected in the 0.5–2 band, in the angular range of 0.5′–24′, while the positive signals were at the ~2σ and 3σ levels in the 2–4.5 and 4.5–10 keV bands, respectively. With power-law fits to the ACFs without the integral constraint term, they have found correlation lengths of $\theta_0 = 1.9 \pm 0.3''$, $0.8^{+0.5}_{-0.4}$, and $6 \pm 2''$ for the three bands, respectively, for a fixed slope $\gamma = 1.8$. The inferred comoving correlation lengths were $r_0 = 9.8 \pm 0.7$, $5.8^{+1.4}_{-1.7}$, and $12 \pm 2\,h^{-1}$ Mpc at the effective redshifts of $z = 1.1$, 0.9, and 0.6, respectively. Comparing the inferred rms fluctuations of the spatial distribution of AGNs $\sigma_{8,\mathrm{AGN}}$ (see Appendix D) with those of the underlying dark matter, the bias parameters of the X-ray source clustering at these effective redshifts were found in the range $b = 1.5$–4. Such a result leads to the conclusion that the typical mass of the DMH hosting an AGN is of the order $M_{\mathrm{DMH}} \sim 10^{13}\,M_\odot\,h^{-1}$. Similar results have been found by Ebrero et al. [42] using the angular correlation function of 30000 X-ray sources in the AXIS survey. In the XMM-LSS survey, Elyiv et al. [43] measured the clustering of ~5000 AGN and computed via Limber's deprojection the obtained $r_0 = 7.2 \pm 0.8$ Mpc/h and $r_0 = 10.1 \pm 0.8$ Mpc/h and $\gamma \sim 2$ in the 0.5–2 keV and 2–10 keV energy bands, respectively. In the XMM-COSMOS field, Gilli et al. [44] measured the clustering of 562 X-ray selected and spectroscopically confirmed AGN. They have obtained that the correlation length of these source, $r_0 = 8.6 \pm 0.5\,h^{-1}$ Mpc, and slope of $\gamma = 1.88 \pm 0.07$. They also found that, if source in redshift spikes removed, the correlation length decreases to about 5-6 h^{-1} Mpc. Even if not conclusively, they also showed that narrow-line AGN and broad-line AGN cluster in the same way, indicating that both classes of sources share the same environment, an argument in favor of the unified AGN model which predicts that obscuration, and therefore the Type-I/Type II dichotomy is simply a geometrical problem. However, it is worth noticing that such a procedure may artificially reduce the clustering signal and the effects of such a cut in the sample may lead to an unreliable estimate of the clustering signal.

Even if the results of Gilli et al. [44] provide a quite complete overview of the environments of the AGN in the COSMOS field, Allevato et al. [45] analyzed the same field by using the halo model formalism (see Section 3). Their results show that AGNs selected in the X-ray band are more biased than the more luminous optically selected QSO. This observation significantly deviates from the prediction of models of merger-driven AGN activity [13, 46], indicating that other mechanisms like disk/bar instability of tidal disruptions may trigger an AGN. They also found that Type 1 AGN are more biased than Type 2 AGNs up to redshift of ~1.5.

In the Böotes field, Hickox et al. [47] explored the connection between different classes of AGN and the evolution of their host galaxies, by deriving host galaxy properties, clustering, and Eddington ratios of AGN selected in the radio, X-ray, and infrared (IR) wavebands from the wide-field (9 $\deg^2$) Böotes survey. They noticed that radio and X-ray AGNs reside in relatively large DMHs ($M_{\mathrm{DMH}} \sim 3 \times 10^{13}$ and $10^{13}\,M_\odot\,h^{-1}$, resp.) and are found in galaxies with red and green colors. In contrast, IR AGNs are in less luminous galaxies, have higher Eddington ratios, and reside in halos with $M_{\mathrm{DMH}} < 10^{12}\,M_\odot\,h^{-1}$.

On the same line, Coil et al. [48] measured the clustering of nonquasar X-ray active galactic nuclei at $z = 0.7$–1.4 in the AEGIS field. Using the cross-correlation of Chandra-selected AGN with 5000 DEEP2 galaxies, they measured a correlation length of $r_0 = 5.95 \pm 0.90\,h^{-1}$ Mpc and slope $\gamma = 1.66 \pm 0.22$. They also concluded that X-ray AGNs have a similar clustering amplitude as red, quiescent, and "green" transition galaxies at $z \sim 1$ and are significantly more clustered than blue, star-forming galaxies. In addition, they proposed a "sequence" of X-ray AGN clustering, where its strength is primarily determined by the host galaxy color; AGNs in red host galaxies are significantly more clustered than AGNs in blue host galaxies, with a relative bias that is similar to that of red to blue DEEP2 galaxies. They did not observe any dependence of clustering on optical brightness, X-ray luminosity, or hardness ratio. In addition, they obtained evidence that galaxies hosting X-ray AGN are more likely to reside in groups and more massive DMHs than galaxies of the same color and luminosity without an X-ray AGN. Allevato et al. [45], Coil et al. [48] and Mountrichas and Georgakakis [49] concluded that DEEP2 X-ray AGN at $z \sim 1$ are more clustered than optically selected quasars (with a 2.6σ significance) and therefore may reside in more massive DMHs. In an evolutionary picture, their results are consistent with galaxies undergoing a quasar phase while in the blue cloud before settling on the red sequence with a lower-luminosity X-ray AGN, if they are similar objects at different evolutionary stages [47]. At lower redshift, Krumpe et al. [50] confirmed the results of Coil et al. [48]. Various recent works have presented indications and/or evidences, of varying significance, regarding a correlation between the X-ray Luminosity and the AGN clustering amplitude, based either on the spatial [34, 44, 48, 50–52] or the angular [53] correlation function.

Note that luminosity-dependent clustering is one of the key features of merger-triggered AGN activity and is one of the prime motivations for AGN clustering analyses. Low L_X AGNs have been found to cluster in a similar way as blue star forming galaxies while high L_X AGN cluster like red passive galaxies. Such a result has been confirmed by Cappelluti et al. [51] using the Swift-BAT all-sky survey at $z \sim 0$. They detected both a L_X dependence of AGN clustering amplitude and a larger clustering of Type I AGN than that of Type II AGN. Krumpe et al. [50, 52] confirm the weak dependence of the clustering strength on AGN X-ray luminosity at a 2σ level for $z < 0.5$.

Table 1: Cappelluti, Allevato, and Finoguenov.

Survey	Band keV	N_{obj}	z	θ_0 arcsec	r_0 h^{-1} Mpc	γ	$b(z)^a$	$\mathrm{Log}(M_{DMH})^b$ $M/(M_\odot h)$
EMSS	0.5–2	183	<0.2	X	<10	X	X	X
RASS	0.1–2.4	2158	1–1.5	~10	<10	1.7 ± 0.3	X	X
RASS	0.1–2.4	2096	0.1	~3.7	6.0 ± 1.6	1.9 ± 0.31	X	X
ROSAT-NEP	0.1–2.4	220	0.22	X	$7.5\ ^{+2.7}_{-4.2}$	$1.85\ ^{+1.90}_{-0.80}$	$1.83\ ^{+1.88}_{-0.61}$	$13.51\ ^{+0.91}_{-0.79}$
AXIS[1]	0.5–2	31288	0.96	22.9 ± 2.0	6.54 ± 0.12	1.12 ± 0.04	2.48 ± 0.07	$13.20\ ^{+0.11}_{-0.12}$
AXIS[1]	2–10	9188	0.94	$29.2^{+5.1}_{-5.7}$	9.9 ± 2.4	$2.33\ ^{+0.10}_{-0.11}$	2.38 ± 0.51	$13.14\ ^{+0.28}_{-0.41}$
AXIS[1]	5–10	1259	0.77	$40.9^{+19.6}_{-29.3}$	5.1 ± 4.1	$1.47\ ^{+0.43}_{-0.57}$	2.14 ± 1.88	$13.17\ ^{+0.84}_{-2.44}$
ELAIS-S1	0.5–2	392	0.4	5.2 ± 3.8	$9.8\ ^{+2.7}_{-4.3}$	1.8	X	X
ELAIS-S1	2–10	205	0.4	12.8 ± 7.8	$13.4\ ^{+2.7}_{-4.3}$	1.8	X	X
CDFS	0.5–2	97	0.84	X	8.6 ± 1.2	1.33 ± 0.11	$2.64\ ^{+0.29}_{-0.30}$	$13.41\ ^{+0.55}_{-0.18}$
CDFN[2]	0.5–2	164	0.96	X	4.2 ± 0.4	1.42 ± 0.07	$1.87\ ^{+0.14}_{-0.16}$	$12.73\ ^{+0.12}_{-0.17}$
XMM-2dF[3]	0.5–2	432	1.2	10.8 ± 1.9	~16	1.8	1.9–2.7	12.5–13.1
XMM-LSS	0.5–2	1130	0.7	6.3 ± 3	6 ± 3	2.2 ± 0.2	X	X
XMM-LSS	2–10	413	0.7	4.2^{+7}_{-13}	6 ± 3	$3.1\ ^{+1.1}_{-0.5}$	X	X
CLASXS	0.5–8	233	1.2	X	$8.1\ ^{+1.2}_{-2.2}$	2.1 ± 0.5	$3.58\ ^{+2.49}_{-1.38}$	$12.86\ ^{+0.61}_{-0.16}$
CDFN[4]	0.5–8	252	0.8	X	$5.8\ ^{+1.0}_{-1.5}$	$1.38\ ^{+0.12}_{-0.14}$	$1.77\ ^{+0.80}_{-0.15}$	$13.53\ ^{+0.63}_{-0.71}$
XMM-COSMOS[5]	0.5–2	1037	1.1	2.9 ± 0.6	11.8 ± 1.1,	1.8	3.7 ± 0.3	13.6 ± 0.1
XMM-COSMOS[5]	2–4.5	545	0.9	$1.2^{+1.1}_{-0.9}$	$6.9\ ^{+2.2}_{-3.1}$,	1.8	$2.5\ ^{+0.7}_{-1.0}$	$13.3\ ^{+0.3}_{-0.7}$
XMM-COSMOS[5]	4.5–10	151	0.6	$6.5^{+3.0}_{-2.7}$	$12.7\ ^{+2.3}_{-2.7}$	1.8	$3.8\ ^{+0.6}_{-0.8}$	13.9 ± 0.2
XMM-COSMOS[6]	0.5–2	538	0.98	X	$8.65\ ^{+0.41}_{-0.48}$	$1.88\ ^{+0.06}_{-0.07}$	3.08 ± 0.14	$13.51\ ^{+0.05}_{-0.07}$
XMM-COSMOS[7]	0.5–2	593	1.21	X	$7.12\ ^{+0.28}_{-0.18}$	$1.81\ ^{+0.04}_{-0.03}$	2.71 ± 0.14	$13.10\ ^{+0.06}_{-0.07}$
SWIFT-BAT	15–55	199	0.045	X	$5.56\ ^{+0.49}_{-0.43}$	$1.64\ ^{+0.07}_{-0.08}$	$1.21\ ^{+0.06}_{-0.07}$	$13.15\ ^{+0.09}_{-0.13}$
AEGIS	0.5–2	113	0.9	X	5.95 ± 0.90	1.66 ± 0.22	$1.97\ ^{+0.26}_{-0.25}$	$13.0\ ^{+0.1}_{-0.4}$
AGES	0.5–2	362	0.51	X	4.5 ± 0.6	1.6 ± 0.1	$1.35\ ^{+0.06}_{-0.07}$	$12.60\ ^{+0.1}_{-0.1}$
ROSAT + SDSS	0.1–2.4	1552	0.27	X	$4.28^{+0.44}_{-0.54}$	$1.67\ ^{+0.13}_{-0.12}$	$1.11\ ^{+0.10}_{-0.12}$	$12.58\ ^{+0.20}_{-0.33}$
XMM-LSS	0.5–2	4360	1.1	3.2 ± 0.5	7.2 ± 0.8	1.93 ± 0.03	2.7 ± 0.3	13.2 ± 0.3
XMM-LSS	2–10	1712	1.0	9.9 ± 0.4	10.1 ± 0.9	1.98 ± 0.04	3.3 ± 0.3	13.7 ± 0.3

X: Unconstrained or undetermined, [a]: Bias factors converted to a common cosmology ($\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\sigma_8 = 0.8$), [b]: DMH masses estimated using van den Bosch [54] and Sheth et al. [55], [1]: Ebrero et al. [42], fit ID = 2, assuming no redshift evolution of the correlation length, [2]: Gilli et al. [33], [3]: Basilakos et al. [30], using the LDDE model, [4]: Yang et al. [34], [5]: Miyaji et al. [41], fit ID = 6 with integral constrain, assuming redshift evolution of the correlation length, [6]: Gilli et al. [44], [7]: Allevato et al. [45].

Table 1 summarizes all the discussed results on the clustering of AGN in X-ray surveys with bias factors converted to a common cosmology ($\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\sigma_8 = 0.8$) in the EMSS, Boyle and Mo [19]; RASS, Vikhlinin and Forman [20], Akylas et al. [22]; ROSAT-NEP, Mullis et al. [24]; AXIS, Ebrero et al. [42]; ELAIS-S1, Puccetti et al. [32]; CDFS, Gilli et al. [33]; CDFN, Gilli et al. [33], Yang et al. [34]; XMM-2dF, Basilakos et al. [30]; XMM-LSS, Gandhi et al. [31]; CLASXS, Yang et al. [34]; COSMOS, Gilli et al. [44] Allevato et al. [45]; Swift-BAT, Cappelluti et al. [51]; AEGIS, Coil et al. [48]; AGES, Hickox et al. [47]; ROSAT-SDSS, Krumpe et al. [50], while Figure 3 shows the redshift evolution of the correlation length r_0 as estimated in previous works, according to the legend.

2.1. Techniques of Investigation. The continuously increasing volume and quality of data allowed a parallel improvement of the techniques of investigation. The first surveys of *Einstein* (see, e.g., [15]) used the autocorrelation function of the unresolved CXB and linked it to the clustering properties of the clustering of X-ray source that produced it.

Modern surveys have mostly estimated correlation function with estimators that use random samples and real data pairs and then estimating physical clustering properties by fitting the correlation function functions with simple power-law models in the form of (2). A detailed description of the method to estimate correlation functions is given in the appendix A. Considering its power, here we give a detailed description of halo modeling which is by far the most reliable formalism to describe clustering of AGN/Galaxies and to determine the environment of a specific DMH tracer.

3. Halo Model

In the hierarchical model of cosmological structure formation, galaxies, group of galaxies, clusters, and so on are built

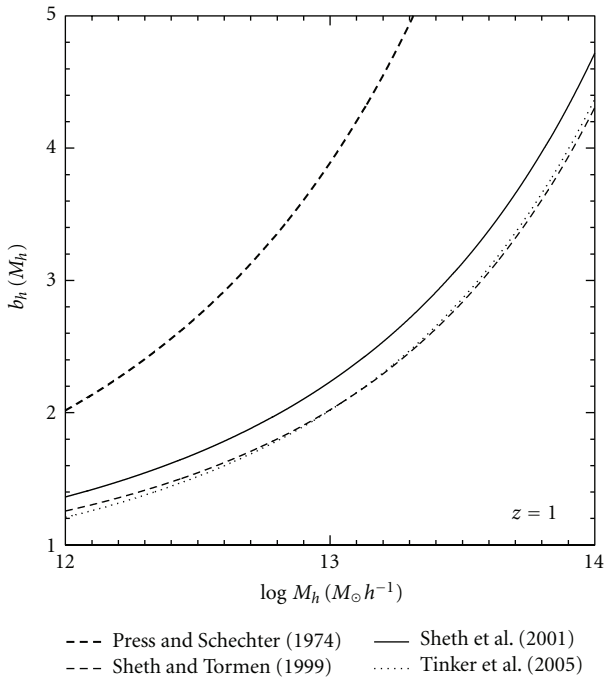


Figure 1: Halo bias as function of halo masses for a fixed redshift $z = 1$ and the corresponding predictions of Press and Shechter [56] (long-dashed line), Sheth and Tormen [57] (dashed line), Sheth et al. [55] (solid line), and Tinker et al. [58] (dotted line).

from the same initial perturbation in the underlying dark matter density field. Regions of dark matter denser-than-average collapse to form halos in which structures form. Galaxies and AGN, as well as, groups and clusters are believed to populate the collapsed DMHs.

The theoretical understanding of galaxy clustering has been greatly enhanced through the framework of the *halo model* [58, 64–67]. One can fill DMHs with objects based on a statistical *halo occupation distribution* (HOD), allowing one to model the clustering of galaxies within halos (and thus at nonlinear scales) while providing a self-consistent determination of the bias at linear scales. Similarly the problem of discussing the abundance and spatial distribution of AGN can be reduced to studying how they populate their host halos.

The HOD analysis recasts AGN-clustering measurements into a form that is more physically informative and conducive for testing galaxy/AGN formation theories.

Thus, one can use measurements of AGN two-point correlation functions to constrain the HOD of different sets of AGN and gain information on the nature of DMH in which they live. In fact, the power of the HOD modeling is the capability to transform data on AGN pair counts at small scales into a physical relation between AGN and DMH at the level of individual halos.

The key ingredient needed to describe the clustering properties of AGN is their *halo occupation distribution function* $P_N(M_h)$, which gives the probability of finding N AGN within a single halo as a function of the halo mass, M_h. In the most general case, $P_N(M_h)$ is entirely specified by all its moments which, in principle, could be observationally determined by studying AGN clustering at any order. Regrettably, AGNs are so rare that their two-point function is already poorly determined, so that it is not possible to accurately measure higher-order statistics. One overcomes this problem by assuming a predefined functional form for the lowest-order moments of $P_N(M_h)$, defining the *halo occupation number* $N(M_h)$ which is the mean value of the halo occupation distribution $N(M_h) = \langle N \rangle (M_h) = \sum_N NP_N(M_h)$. It is convenient to describe $N(M_h)$ in terms of a few parameters whose values will then be constrained by the data.

An accurate description of matter clustering on the basis of the halo approach requires three major ingredients: these halo mass function $n(M_h)$ (the number of DMHs per unit mass and volume), the mass-dependent biasing factor $b(M_h)$, and the density profile of halos. These terms, along with a parametrization of $N(M_h)$, allow us to calculate some useful quantities; the number density of AGN:

$$n_{\mathrm{AGN}} = \int n(M_h) N(M_h) dM_h, \tag{3}$$

the large-scale bias:

$$b = \frac{\int b_h(M_h) N(M_h) n(M_h) dM_h}{\int N(M_h) n(M_h) dM_h}, \tag{4}$$

and the average mass of the host dark halo:

$$M = \frac{\int M_h N(M_h) n(M_h) dM_h}{\int N(M_h) n(M_h) dM_h}. \tag{5}$$

The number density and clustering properties of the DMHs can be easily computed, at any redshift, by means of a set of analytical tools which have been tested and calibrated against numerical simulations [55, 57, 58, 68–72]. Popular choices for both $n(M_h)$ and $b(M_h)$ are the analytical spherical collapse [57] or an ellipsoidal collapse model ([55], see Section 4 for more details). A detailed description of HOD mathematical formalism is given in Appendix B.

3.1. Occupation Number. In the past ten years, a very successful framework for modeling the nonlinear clustering properties of galaxies has been developed and a number of halo models have been presented in the literature. These have been successfully used to describe the abundance and clustering properties of galaxies at both low [58, 65, 73–82] and high [83–87] redshifts, as well as whether these galaxies occupy the centers of the DMH or are satellite galaxies [67, 88].

Partially due to the low number density of AGN, there have been few results in the literature interpreting AGN correlation function using HOD modeling, where the small-scale clustering measurements are essential. Porciani et al. [89] studied the clustering of 2QZ QSO with the halo model to infer the mean number of optically selected quasars which are harboured by a virialized halo of given mass and the characteristic quasar lifetime. Padmanabhan et al. [90] discussed qualitative HOD constraints on their LRG-optical

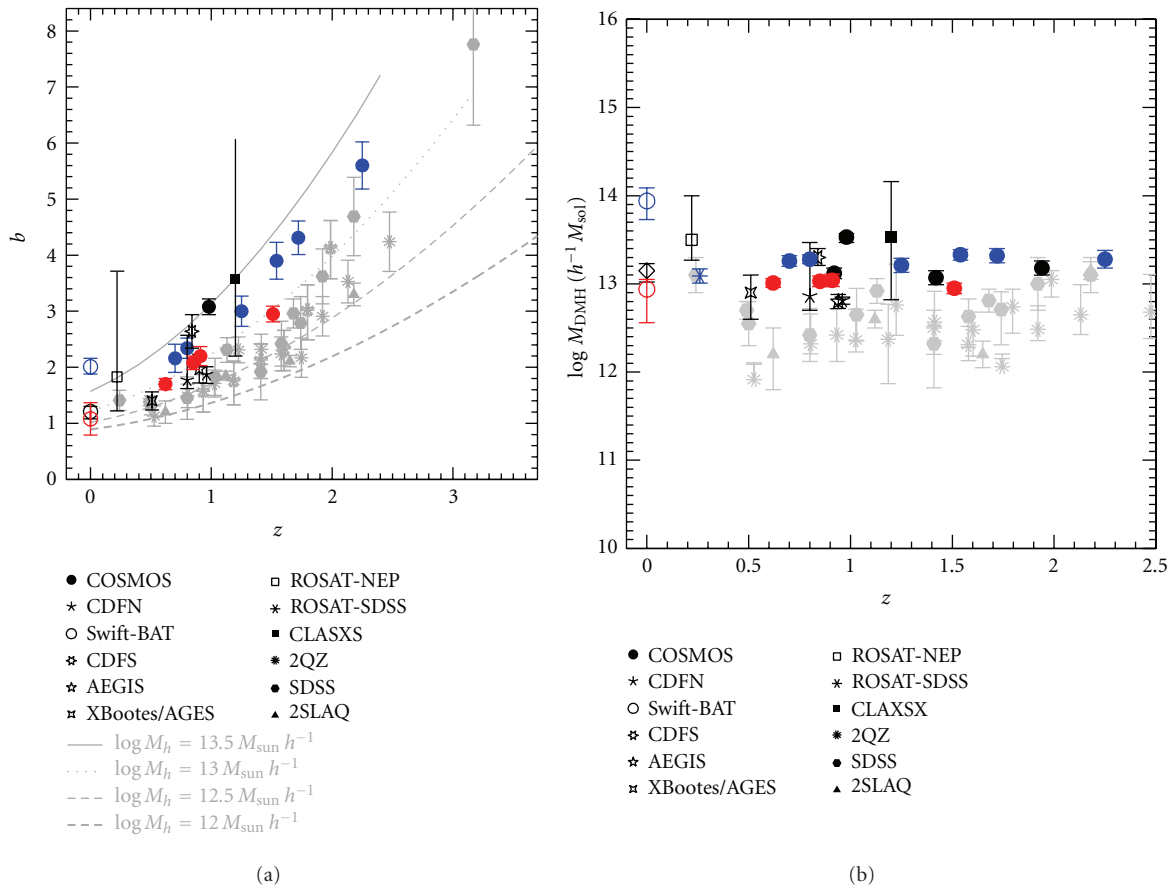


FIGURE 2: Bias factor (a) and mass of AGN hosting halos (b) as a function of redshift for X-ray-selected AGN (black data points), X-ray-selected Type 1 AGN (blue data points), and X-ray selected Type 2 AGN (red data points) as estimated in different surveys (COSMOS, Gilli et al. [44], Allevato et al. [45]; CDFN, Gilli et al. [33], Yang et al. [34]; Swift-BAT, Cappelluti et al. [51]; CDFS, Gilli et al. [33]; AEGIS, Coil et al. [48]; AGES, Hickox et al. [47]; ROSAT-NEP, Mullis et al. [24]; ROSAT-SDSS, Krumpe et al. [50]; CLASXS, Yang et al. [34]). The dashed lines show the expected $b(z)$ of DMHs with different masses according to the legend, based on Sheth et al.[55]. The grey points show results from quasar-quasar correlation measurements using spectroscopic samples from SDSS [59, 60], 2QZ [61, 62], and 2SLAQ [63]. All the previous studies infer the picture that X-ray-selected AGN which are moderate luminosity AGN compared to bright quasars inhabit more massive DMHs than optically selected quasars in the range $z = 0.5$–2.25.

QSO cross-correlation function (CCF), and Shen et al. [91] modelled with the HOD the observed two-point correlation function of 15 binary quasars at $z > 2.9$.

The standard halo approach used for quasars and galaxies is based on the idea that the elements of HOD can be effectively decomposed into two components, separately describing the properties of central and satellite galaxies within the DMH. A simple parametric form used to describe the galaxy HOD is to model the mean occupation number for central galaxies as a step function, that is, $\langle N_{cen} \rangle = 1$ for halos with mass $M \geq M_{min}$ and $\langle N_{cen} \rangle = 0$ for $M < M_{min}$, while the distribution of satellite objects can be well approximated by a Poisson distribution with the mean following a power law, $\langle N_{sat} \rangle = (M/M_1)^{\alpha}$. Previously derived HOD of galaxies show α values ~ 1–1.2 which imply a number of satellite galaxies approximately proportional to M_h.

The clustering properties of X-ray-selected AGN have been modelled with the HOD in two previous works for sources in the *Bootes* field Starikova et al. [92] and in the *ROSAT All-Sky Survey* Miyaji et al. [93]. Starikova et al. [92] used the the projections of the two-point correlation function both on the sky plane and in the line of sight to show that *Chandra/Bootes* AGNs are located at the center of DM halos with $M > M_{min} = 4 \times 10^{12}\,h^{-1}\,M_{\odot}$, assuming a halo occupation described by a step function (zero AGN per halo/subhalo below M_{min} and one above it). They also showed that Chandra/Boötes AGNs are located at the centers of DMHs, limiting the fraction of AGN in noncentral galaxies to be <0.09 at the 95% CL. The central locations of the AGN host galaxies are expected in the merger trigger model because mergers of equally sized galaxies preferentially occur at the centers of DMH [8].

Miyaji et al. [93] modelled the AGN HOD testing the effects of having or not AGN in central galaxies by using the RASS AGN-LRG cross-correlation. In the first scenario, they assumed that all the AGNs are satellites and they visualized the HOD of the LRG as a step function with a step at $\log M_h[h^{-1}\,M_{\odot} = 13.5]$. While formally they assumed that

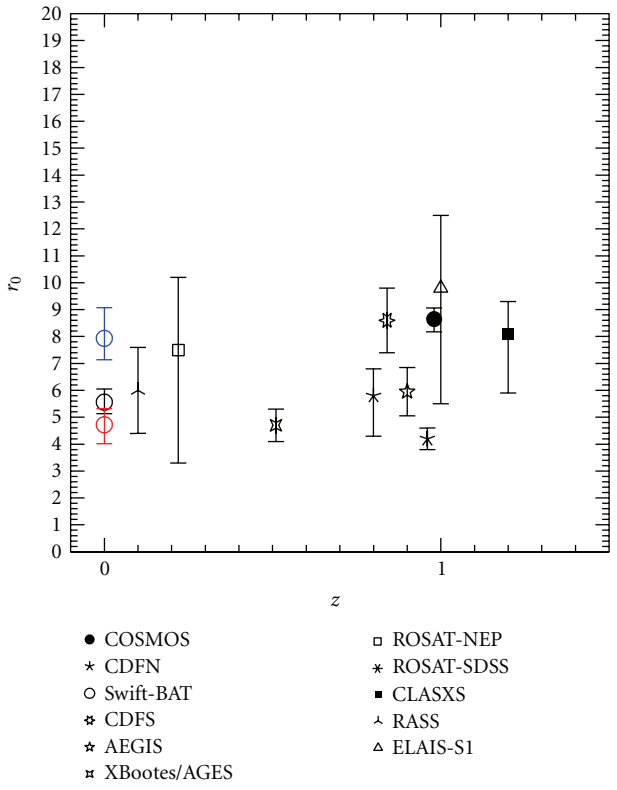


Figure 3: Redshift evolution of the correlation length r_0 as estimated in different X-ray surveys (COSMOS, Gilli et al. [44], Allevato et al.[45]; CDFN, Gilli et al. [33], Yang et al. [34]; Swift-BAT, Cappelluti et al. [51]; CDFS, Gilli et al. [33]; AEGIS, Coil et al. [48]; AGES, Hickox et al. [47]; ROSAT-NEP, Mullis et al. [24]; ROSAT-SDSS, Krumpe et al. [50]; CLASXS, Yang et al. [34]; RASS, Akylas et al. [22]; ELAIS-S1, Puccetti et al. [32]).

all AGNs are not in central galaxies, the HOD constraints obtained from this assumption can be applied to satellite and central AGN if the AGN activity in central galaxies of high-mass halos ($\log M_h[h^{-1}\,M_\odot > 13.5]$) is suppressed. In particular, they used a truncated power-law satellite HOD, with two parameters: the critical DMH mass below which the AGN HOD is zero and the slope α of the HOD for $M_h > M_{\rm cr}$. They also investigated a model where the central HOD is constant and the satellite HOD has a power-law form, both at masses above $M_{\rm min}$. In all the cases, they rejected $\alpha \sim 1$, finding a marginal preference for an AGN fraction among satellite galaxies which decreases with increasing M_h. They argued that this result might be explained by a decrease of the cross-section for galaxy merging in the environment of richer groups or clusters. In fact, previous observations infer that the AGN fraction is smaller in clusters than in groups [27, 94–96].

It is important to stress that the small number statistics has so far limited the accuracy of correlation function of X-ray AGN at small-scales, especially through the autocorrelation function of the AGN themselves. The situation can be improved by measuring the cross-correlation function of AGN with a galaxy sample that has a much higher space density, with common sky and redshift coverage as the AGN redshift surveys. The AGN clustering through cross-correlation function with galaxies is emerging in the last years [47, 48, 90, 97–99] and can be used to improve our understanding of how AGNs populate DMH [52, 93].

4. Bias and DMH Mass

In the literature, the bias parameter is often calculated with the power-law fits [24, 34, 41, 48, 50, 100] over scales of 0.1–$0.3 < r_p < 10$–$20\,h^{-1}$ Mpc. The power-law models of the ACF are usually converted to the rms fluctuation over $8\,h^{-1}$ Mpc spheres or are averaged up to the distance of $20\,h^{-1}$ Mpc. While some authors use only large scales ($r_p > 1$-$2\,h^{-1}$ Mpc) to ensure that the linear regime is used, others include smaller scales to have better statistics. As an example, Hickox et al. [47] fitted their data with a biased DMH-projected correlation function.

In the HOD analysis, the bias factor only comes from the 2-halo term ($r_p > 1$-$2\,h^{-1}$ Mpc). Miyaji et al. [93] compared the bias of RASS-AGN from the full HOD model (D.4) with the one estimated using the power-law best fits parameters, finding that the bias estimates are consistent within 1σ. Moreover, using (D.1), one introduces large statistical errors. Allevato et al. [45] found a similar results in comparing the bias of X-ray AGN in COSMOS field from the 2-halo term with (D.3) and the one estimated from the power-law best fits parameters. In Appendix C, we describe the mathematical procedures for the bias parameter calculation commonly used in the literature.

Most of the authors [47, 50, 51] used an analytical expression (as the one described in [55, 57, 58, 69]) to assign a characteristic DMH mass to the hosting halos. The large-scale bias is directly related to the mass function of halos, so that the mass of a halo dictates the halo clustering and the number of such halos. The halo mass can be quantified in terms of the peak height $\nu = \delta_c/\sigma(M_h, z)$, which characterizes the amplitude of density fluctuations from which a halo of mass M_h forms at a given redshift. In generals one assumes $\delta_c = 1.686$ and $\sigma(M_h, z)$ is the linear overdensity variance in spheres enclosing a mean mass M_h. The traditional choice of the mass function and then of the bias has been that of Press and Schechter [56]:

$$b^{PS} = 1 + \frac{\nu^2 - 1}{\delta_c}. \tag{6}$$

A commonly used prescription was derived by Sheth and Tormen [57]:

$$b^{ST} = 1 + \frac{a\nu^2 - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (a\nu^2)^p}, \tag{7}$$

where $a = 0.707$ and $p = 0.3$ or the ellipsoidal collapse formula of Sheth et al. [55]:

$$b^{\rm SMT} = 1 + \frac{1}{\sqrt{a}\delta_c}\left[\sqrt{a}(a\nu^2) + \sqrt{a}b(a\nu^2)^{1-c} - \frac{(a\nu^2)^c}{(a\nu^2)^c} + b(1-c)\left(\frac{1-c}{2}\right)\right], \tag{8}$$

where $a = 0.707$, $b = 0.5$, $c = 0.6$ or the recalibrated parameters $a = 0.707$, $b = 0.35$, $c = 0.8$ of Tinker et al. [58]. The ν parameter can be estimated following the appendix of Van den Bosch [54]. Figure 1 shows the bias as function of the halo mass M_h, at $z = 1$, following the predictions of Press and Schechter [56], Sheth and Tormen [57], Sheth et al. [55], and Tinker et al. [58].

Allevato et al. [45] argued that this approach reveals an incongruity due to the fact that the AGN bias used in the formulas above is the average bias of a given AGN sample at a given redshift. In fact, following this approach, one cannot take into account that the average bias is sensitive to the entirety of the mass distribution; different mass distributions with different average masses can give rise to the same average bias.

On the contrary, by using the halo model, the average bias and the average mass of the sample, (D.4), and (5) properly account for the shape of the mass distribution: the average bias depends on the halo number density and on the AGN HOD, integrated over the mass range of the particular AGN sample. They introduced a new method that uses the 2-halo term in estimating the AGN bias factor assuming an AGN HOD described by a δ-function. Following this approach, they properly took into account for the sample variance and the growth of the structures over time associated with the use of large redshift interval of the AGN sample.

On the other hand, Miyaji et al. [93] and Krumpe et al. [52] applied the HOD modeling technique to the RASS AGN-LRG CCF in order to move beyond determining the typical DMH mass based on the clustering signal strength and instead constrain the full distribution of AGN as a function of DMH mass. Along with a parametrization of $N(M_h)$, they estimated the large-scale bias and the typical mass of hosting DM halos using (D.4) and (5). This method improves the clustering analysis because it properly uses the nonlinear growth of matter in the *1-halo* term through the formation and growth of DMHs. These results are significant improvements with respect to the standard method of fitting the signal with a phenomenological power law or using the 2-halo term (see Appendix C).

4.1. X-Ray-Selected AGN Bias, Bias Evolution, and Mass of the Hosting Halos. The majority of the X-ray surveys agree with a picture where X-ray AGNs are typically hosted in DM halos with mass of the order of $12.5 < \log M_{\rm DMH}[{\rm h}^{-1}\,M_\odot] < 13.5$, at low ($z < 0.4$) and high ($z \sim 1$) redshift [33, 34, 44, 47, 48, 50–52, 92, 93].

At high redshift, Gilli et al. [33] measured the clustering of X-ray AGN with $z = 0$–4 in both the $\sim 0.1\,{\rm deg}^2$ CDFs, finding $b = 1.87^{+0.14}_{-0.16}$ for 240 sources in the northern field and $b = 2.64^{+0.29}_{-0.30}$ for 124 sources in the southern field. At $z \sim 1$, Yang et al. [34] measured the clustering of 233 spectroscopic sources in the $0.4\,{\rm deg}^2$ *Chandra* CLASXS area and of 252 spectroscopic sources from the CDFN, both at $z = 0.1$–3. They found $b = 3.58^{+2.49}_{-1.38}$ for the CLASXS AGN and $b = 1.77^{+0.80}_{-0.15}$ for the CDFN field. Gilli et al. [44] studied 538 XMM-COSMOS AGN with $0.1 < z < 3$, and they found a bias factor $b = 3.08^{+0.14}_{-0.14}$ at $\bar{z} \sim 1$. Using the Millennium simulations, they suggested that XMM-COSMOS AGNs reside in DMH with mass $M_{\rm DMH} > 2.5 \times 10^{12}\,h^{-1}\,M_\odot$. Coil et al. [48] measured the clustering of X-ray AGN at $z = 0.7$–1.4 in the AEGIS field, and they estimated $b = 1.85^{+0.28}_{-0.28}$. Following Zheng et al. [87], they infer from the bias factor that, at $z = 0.94$, the minimum DM halo mass of the X-ray AGN is $>10^{12}\,M_\odot\,h^{-1}$. These results combined with Mountrichas and Georgakakis [49] show that moderate luminosity X-ray-selected AGN live in DMHs with masses $M_h \sim 10^{13}\,h^{-1}\,M_\odot$ at all redshifts since $z \sim 1$. At lower redshift, Hickox et al. [47] analysed 362 AGES X-ray AGN at $\langle z \rangle = 0.51$. The bias factor equal to $b = 1.40 \pm 0.16$ indicates that X-ray AGNs inhabit DM halos of typical mass $\sim 10^{13}\,M_\odot\,h^{-1}$.

In the local Universe, Cappelluti et al. [51] estimated for ~ 200 Swift-BAT AGN a bias equal to $b = 1.21^{+0.07}_{-0.06}$ which corresponds to $\log M_{\rm DM} = 13.15\ ^{+0.09}_{-0.13}\,h^{-1}\,M_\odot$.

Allevato et al. [45] estimated an average mass of the XMM-COSMOS AGN hosting halos equal to $\log M_0[h^{-1}Mpc] = 13.10 \pm 0.06$ at $z \sim 1.2$. They also measured the bias of Type 1 and Type 2 AGN, finding that the latter resides in less massive halos than Type 1 AGN. Only two other works [50, 51] analysed the clustering properties of X-ray-selected Type 1 AGN and Type 2 AGN. Cappelluti et al. [51] estimated the typical DM halo mass hosting type 1 and type 2 Swift-BAT AGN at $z \sim 0$. They measured that these two different samples are characterized by halos with mass equal to $\log M_{\rm DM}[h^{-1}\,M_\odot] \sim 13.94^{+0.15}_{-0.21}$ and $\sim 12.92^{+0.11}_{-0.38}$, respectively. However, the lack of small separation pair of Type I AGN in the local Universe may have produced systematic deviations which were not accounted in their fits. In Krumpe et al. [50], the bias factor of BL RASS AGN at $z = 0.27$ is consistent with BL AGN residing in halos with mass $\log M_{\rm DM}[h^{-1}\,M_\odot] = 12.58^{+0.20}_{-0.33}$.

Using the HOD model, Starikova et al. [92] suggested that X-ray Chandra/Bootes AGN is located at the center of DM halos with $M > M_{\rm min} = 4 \times 10^{12}\,h^{-1}\,M_\odot$, while Miyaji et al. [93] estimated for RASS AGN at $z = 0.25$, $b = 1.32 \pm 0.08$, and a typical mass of the host halos of 13.09 ± 0.08.

The redshift evolution of the clustering of X-ray-selected AGN has been first studied by Yang et al. [34] in the CLAXS+CDFN fields. They measured an increase of the bias factor with redshift, from $b = 0.95 \pm 0.15$ at $z = 0.45$ to $b = 3.03 \pm 0.83$ at $z = 2.07$, corresponding to an average halo mass of $\sim 12.11\,h^{-1}\,M_\odot$.

Allevato et al. [45] studied the redshift evolution of the bias for a sample of XMM-COSMOS AGN at $z < 2$. They found a bias evolution with time from $b(z = 0.92) = 1.80 \pm 0.19$ to $b(z = 1.94) = 2.63 \pm 0.21$ with a DM halo mass consistent with being constant at $\log M[h^{-1}\,M_\odot] \sim 13.1$ at all redshifts $z < 2$. They also found evidence of a redshift evolution of the bias factor of XMM-COSMOS Type 1 AGN and Type 2. The bias evolves with redshift at constant average halo mass $\log M_0[h^{-1}\,M_\odot] \sim 13.3$ for Type 1 AGN and $\log M_0[h^{-1}\,M_\odot] \sim 13$ for Type 2 AGN at $z < 2.25$ and $z < 1.5$, respectively. In particular, Allevato et al. [45] argued that X-ray selected Type 1 AGNs reside in more massive DMHs compared to X-ray-selected Type 2 AGN at all redshifts at

$\sim 2.5\sigma$ level, suggesting that the AGN activity is a mass-triggered phenomenon and that different AGN classes are associated with the DM halo mass, irrespective of redshift z.

Krumpe et al. [52] measured the clustering amplitudes of both X-ray RASS and optically selected SDSS broad-line AGNs, as well as for X-ray-selected narrow-line RASS/SDSS AGNs through cross-correlation functions with SDSS galaxies and derive the bias by applying the HOD model directly to the CCFs. They estimated typical DMH masses of broad-line AGNs in the range $\log(M_h/[\mathrm{h}^{-1}\,M_\odot]) = 12.4$–$13.4$, consistent with the halo mass range of typical non-AGN galaxies at low redshifts, and they found no significant difference between the clustering of X-ray-selected narrow-line AGNs and broad-line AGNs up to $z \sim 0.5$.

Figure 2(a) shows the bias parameter and Figure 2(b) the mass of the AGN hosting halos as a function of redshift for X-ray-selected AGN (black data points), X-ray-selected Type 1 AGN (blue data points), and X-ray-selected Type 2 AGN (red data points) as estimated for different surveys (see the legend). The dashed lines show the expected b(z) of typical DM halo masses M_{DMH} based on Sheth et al. [55]. The masses are given in $\log M_{\mathrm{DMH}}$ in units of $h^{-1}\,M_\odot$.

There have been several studies of the bias evolution of optical quasar with the redshift as shown in Figure 2 (grey data points), based on large survey samples such as 2QZ, 2SLAQ, and SDSS [59–63]. These previous studies infer the picture that X-ray-selected AGNs which are moderate luminosity AGN compared to bright quasars inhabit more massive DMHs than optically selected quasars in the range $z = 0.5$–2.25.

Recently, Krumpe et al. [52] verified that the clustering properties between X-ray and optically selected AGN samples are not significantly different in three redshift bins below $z = 0.5$ (the differences are 1.5σ, 0.1σ, and 2.0σ).The reason for the fact that X-ray-selected AGN samples appear to cluster more strongly than optically selected AGNs is still unclear. Allevato et al. [45] and Mountrichas and Georgakakis [49] suggested that the difference in the bias and then in the host DMH masses is due to the different fueling mode of those sources from that of the X-ray-selected moderate luminosity AGN. On the contrary, Krumpe et al. [52] suggested that some of the X-ray clustering studies significantly underestimate their systematic uncertainties and then it may turn out that these measurements are consistent with optical AGN clustering measurements. More high-z AGN clustering measurements based on larger samples are needed to gain a clearer picture.

4.2. AGN Life Time. One of the most important tests for studying the evolution models of AGN is understanding their lifetime. It is widely accepted that AGN is phase of the galaxy life necessary to explain the coevolution of the bulge and the black hole. After a triggering event of which we do not know the nature, yet, the central black hole begins its accretion phase and it is believed that it undergoes several regimes of Eddington rates and bolometric luminosity. Martini and Weinberg [101] proposed a method to derive the AGN life time by knowing their space density and their DMH host mass.

By knowing the AGN and DMH halo space density at a given luminosity and mass (n_{AGN}, n_{DMH}), one can estimate the duty cycle of the AGN, $\tau_{\mathrm{AGN}}(z) = (n_{\mathrm{AGN}}(L,z)/n_{\mathrm{DMH}}(M,z))(\tau_H(z))$, where $\tau(H(z))$ is the Hubble time at a given redshift. Actually, this method provides only an upper limit since it assumes that the life of halo of a given mass is similar to the Hubble time. A more exhaustive formulation would be $\tau_{\mathrm{AGN}}(z) = (n_{\mathrm{AGN}}(L,z)/n_{\mathrm{DMH}}(M,z))(\tau_{\mathrm{DMH}}(z))$, where $\tau_{\mathrm{DMH}}(z)$ is the age of a DMH at given redshift. Unfortunately, this quantity cannot be estimated analytically but could be estimated in a statistical way by using hydrodynamic simulations. Several results can be mentioned for these quantities, but their dispersion is very large, therefore we report only some example. At $z = 1$, Gilli et al. [44] obtains that the typical duty cycle of AGN is <1 Gyr. At $z = 0$, Cappelluti et al. [51] have measured a duty cycle in the range 0.2 Gyr–5 Gyr with an expectation value of 0.7 Gyr. Both the measurements are fairly larger than the 40 million years determined by Martini and Weinberg [101] at $z = 2$-3. These differences, however, are not surprising if we assume that the different populations of AGN grow with a different Eddington rate as function of their typical luminosities and/or redshifts [102].

5. Discussion

In this paper, we reviewed the results in the field of X-ray AGN clustering, for energies between 0.1 keV to 55 keV over a period of more than 20 years. The literature has produced an increasingly convincing and consistent picture of the physical quantities derivable from this kind of study. Most of the advancements in the field have been achieved with the improvement of survey capabilities and instruments sensitivity. The availability of simultaneously wide and deep fields, coupled with multiwavelength information, has produced larger and larger samples of spectroscopically confirmed sources. This allowed several teams to refine the techniques needed to estimate the two-point ACF and the quantities derived form it. In particular, we are entering a phase where, at least at $z < 2$, AGN clustering studies will not probably provide any new result unless evaluated with the HOD formalism. Open questions as what is the AGN occupation number and the evolution of HOD define a new barrier which is necessary to break in order to understand the history of X-ray emission from accretion onto AGN. In this respect, samples of X-ray-selected AGN always need a spectroscopical followup to provide a solid base to compute clustering in the real space rather than in the angular space.

Summarizing, the current picture is that X-ray-selected AGNs are highly biased objects with respect to the underlined matter distribution. Such an evidence is clearer when measuring the redshift dependence of AGN bias. At every redshift from $z = 0$ to $z = 2$, AGNs cluster in way similar to DMH of mass of the order of $\log(M_\odot\,\mathrm{h}^{-1}) = 13$. The spread of such a value is of the order 0.3–0.5 dex at 1σ. This

means that the determination of what kind of environment is inhabited by AGN is relatively well constrained and identical at every redshift sampled by X-ray surveys. This allows us to formulate the hypothesis that every phase of AGN activity is mass-triggered phenomenon (i.e., each AGN evolutionary phase is characterized by a critical halo mass).

It is believed that major mergers of galaxies is one of the dominant mechanisms for fueling quasars at high redshift and bright luminosities, while minor interactions, bar instabilities, or tidal disruptions are important at low redshift ($z \lesssim 1$) and low luminosities ($L \lesssim 10^{44}\,\mathrm{erg\,s^{-1}}$) [9, 13, 103, 104]. In the local Universe, for example, the study of the environment of Swift BAT Seyfert galaxies [28] finds a larger fraction of BAT AGNs with disturbed morphologies or in close physical pairs (<30 kpc) compared to matched control galaxies or optically selected AGNs. The high rate of apparent mergers (25%) suggests that AGN activity and merging are critically linked for the moderate luminosity AGN in the BAT sample. Moreover, models of major mergers appear to naturally produce many observed properties of quasars, as the quasar luminosity density, the shape, and the evolution of the quasar luminosity function and the large-scale quasar clustering as a function of L and z (e.g., [8, 46, 105–109]). Quasar clustering at all redshift is consistent with halo masses similar to group scales, where the combination of low velocity dispersion and moderate galaxy space density yields to the highest probability of a close encounter [8, 11]. Moreover, recent detections of an L_X-dependent clustering play in favor of major mergers being the dominant AGN triggering mechanism.

On the other hand, it has became clear that many AGNs are not fueled by major mergers and only a small fraction of AGNs are associated with morphologically disturbed galaxies. Georgakakis et al. [110] and Silverman et al. [96] found that AGNs span a broad range of environments, from the field to massive groups and thus major mergers of galaxies, possibly relevant for the more luminous quasar phenomenon, may not be the primary mechanism for fueling these moderate luminosity AGN.

Georgakakis et al. [111] suggest that bar instabilities and minor interactions are more efficient in producing luminous AGN at $z \lesssim 1$ and not only Seyfert galaxies and low-luminosity AGN as the Hopkins and Henquist [9] model predicts. Cisternas et al. [112] analysed a sample of X-ray-selected AGN host galaxies and a matched control sample of inactive galaxies in the COSMOS field. They found that mergers and interactions involving AGN hosts are not dominant and occur no more frequently than for inactive galaxies. Over 55% of the studied AGN sample that is characterized by $L_{\mathrm{BOL}} \sim 10^{45}\,\mathrm{erg\,s^{-1}}$ and by mass of the host galaxies $M_* \gtrsim 10^{10}\,M_\odot$ are hosted by disk-dominated galaxies, suggesting that secular fuelling mechanisms can be highly efficient.

Moreover, several works on the AGN host galaxies [113–118] show that the morphologies of the AGN host galaxies do not present a preference for merging systems.

At high redshift ($z \sim 2$), recent findings of Schlegel et al. [119] and Rosario et al. [120], who examined a smaller sample of AGN in the ERS-II region of the GOODS-South field, inferred that late-type morphologies are prevalent among the AGN hosts. The role that major galaxy mergers play in triggering AGN activity at $1.5 < z < 2.5$ was also studied in the CDF-S. At $z = 1.5$–3, Schawinski et al. [121] showed that, for X-ray-selected AGN in the Chandra Deep Field South and with typical luminosities of $10^{42}\,\mathrm{erg\,s^{-1}} < L_X < 10^{44}\,\mathrm{erg\,s^{-1}}$, the majority (80%) of the host galaxies of these AGNs have low Srsic indices indicative of disk-dominated light profiles, suggesting that secular processes govern a significant fraction of the cosmic growth of black holes. That is, many black holes in the present-day Universe grew much of their mass in disk-dominated galaxies and not in early-type galaxies or major mergers.

Later, Kocevski et al. [122] found that X-ray-selected AGNs at $z \sim 2$ do not exhibit a significant excess of distorted morphologies while a large fraction reside in late-type galaxies. They also suggested that these late-type galaxies are fueled by the stochastic accretion of cold gas, possibly triggered by a disk instability or minor interaction.

Allevato et al. [45] argued that for moderate luminosity X-ray AGN secular processes such as tidal disruptions or disk instabilities might play a much larger role than major mergers up to $z \sim 2.2$.

It becomes important to study the clustering properties of AGN at high redshift when we assume the peak of the merger-driven accretion. Moreover, given the complexity of AGN triggering, a proper selection of AGN samples, according to the luminosity or the mass of the host galaxies, can help to test a particular model boosting the fraction of AGN host galaxies associated with morphologically disturbed galaxies.

From the evolutionary point of view, the evidence of a bias segregation of optically and X-ray-selected AGN might be a sufficient proof to claim that the two phenomena are sensitive to different environments and therefore likely driven by different triggering mechanisms. A more comprehensive picture will be available when the clustering of different phases of AGN activity will be studied and compared.

Hickox et al. [47] interpreted their clustering results in terms of a general picture for AGN and galaxy evolution which is reproduced in Figure 4. The picture consists of an evolutionary sequence that occurs at different redshifts for halos with different masses. In this scenario, luminous AGN accretion occurs preferentially (through a merger or some secular process) when a host DMH reaches a critical M_{DMH} between 10^{12} and $10^{13}\,M_\odot\,h^{-1}$ (this phase is indicated by the solid ovals). Once a large halo reaches this critical mass, it becomes visible as a ULIRG or SMG (owing to a burst of dusty star formation) or (perhaps subsequently) as a luminous, unobscured quasar. The ULIRG/quasar phase is associated with rapid growth of the SMBH and formation of a stellar spheroid and is followed by the rapid quenching of star formation in the galaxy. Subsequently, the young stellar population in the galaxy ages (producing "green" host galaxy), and the galaxy experiences declining nuclear accretion that may be associated with an X-ray AGN. Eventually, the aging of the young stars leaves a "red" and "dead" early-type galaxy, which experiences intermittent

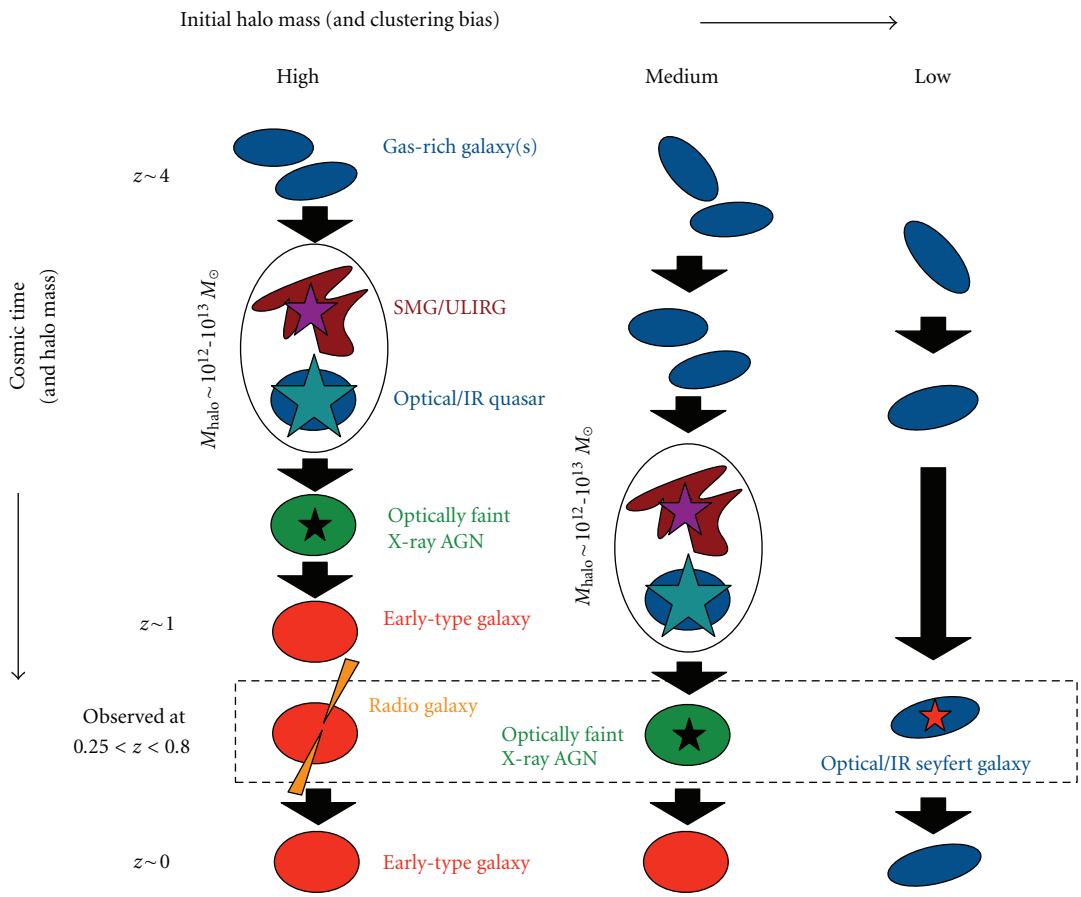


FIGURE 4: Schematic for a simple picture of AGN and host galaxy evolution, taken from Hickox et al. [47] and motivated by the AGN host galaxy and clustering results presented in that study.

"radio-mode" AGN outbursts that heat the surrounding medium. For "medium" initial DMHs, the quasar phase and formation of the spheroid occur later than for the systems with high halo mass, so that, at $z \sim 0.5$, we may observe the green X-ray AGN phase. Even smaller halos never reach the threshold mass for quasar triggering; these still contain star-forming disk galaxies at $z \lesssim 0.8$, and we observe some of them as optical or IR-selected Seyfert galaxies. The dashed box indicates the AGN types (in their characteristic DMH) that would be observable in the redshift range $0.25 < z < 0.8$.

Further steps in the field will require the study of clustering of AGN from $z = 3$ to $z = 6$-7. This will likely lead to the determination of the mass of early DM spheroids who hosted primordial black holes seeds. However, this is a very challenging task since it requires a very deep and wide survey with an almost complete optical followup.

BOSS [123] and BigBOSS [119] will detect high redshift AGNs at $z \sim 2.2$, which will improve AGN clustering measurements at higher redshifts. The only approved mission that at the moment will allow to study the $z = 3$–5 X-ray Universe is eROSITA ([124], launch Dec. 2013) for which an estimate of the completeness of the typical followup is still unavailable. Additionally, the Large Synoptic Survey Telescope ([125], LSST) is expected to identify ∼2 million AGNs in optical bands. eROSITA and LSST have the potential to significantly improve AGN clustering measurements at low and high redshifts, though only if there are dedicated large spectroscopic follow-up programs. Another strong contribution will come from either Nustar that will likely provide a better view of AGN clustering without the selection biases introduced by photoelectric absorption. Athena, the proposed ESA new generation telescope that will mount a wide field imager on a very large collecting area telescope, will provide a further view on the deep X-ray sky and likely push our knowledge of the high-z X-ray Universe.

In addition to better model the evolution of SMBH environments, a fundamental point to start is to establish the nature of BH seeds at $z = 10$. Such a determination will likely come with the new generation of telescope like JWST and ESO-ELT.

Appendices

A. Deriving the Two-Point Autocorrelation Function

The two-point autocorrelation function ($\xi(r)$, ACF) describes the excess probability over random of finding

a pair with an object in the volume dV_1 and another in the volume dV_2, separated by a distance r so that $dP = n^2[1 + \xi(r)]dV_1dV_2$, where n is the mean space density. A known effect when measuring pairs separations is that the peculiar velocities combined with the Hubble flow may cause a biased estimate of the distance when using the spectroscopic redshift. To avoid this effect, it is usually computed the projected ACF [126]: $w(r_p) = 2\int_0^{\pi_{\max}} \xi(r_p, \pi)d\pi$, where r_p is the distance component perpendicular to the line of sight and π parallel to the line of sight [127]. It can be demonstrated that, if the ACF is expressed as $\xi(r) = (r/r_0)^{-\gamma}$, then

$$w\left(r_{\mathrm{p}}\right) = A(\gamma) r_0^{\gamma} r_p^{1-\gamma}, \tag{A.1}$$

where $A(\gamma) = \Gamma(1/2)\Gamma[(\gamma - 1)/2]/\Gamma(\gamma/2)$ [21].

The ACF is mostly estimated by using the minimum variance estimator described by Landy and Szalay [128]:

$$\xi\left(r_p, \pi\right) = \frac{\mathrm{DD} - 2\mathrm{DR} + \mathrm{RR}}{\mathrm{RR}}, \tag{A.2}$$

where DD, DR, and RR are the normalized number of data-data, data-random, and random-random source pairs, respectively. Equation (A.2) indicates that an accurate estimate of the distribution function of the random samples is crucial in order to obtain a reliable estimate of $\xi(r_p, \pi)$. Note that other estimators have been proposed in the literature, but the Landy and Szalay [128] one has been shown to provide the smallest statistical variance. Such a formalism can be easily adopted when computing the angular or the redshift space correlation function, with the only difference that the evaluation is made on a single dimension. Several observational biases must be taken into account when generating a random sample of objects in a X-ray-flux limited survey. In particular, in order to reproduce the selection function of the survey, one has to carefully reproduce the space and flux distributions of the sources, since the sensitivity in X-ray surveys is not homogeneous on the detector and therefore on the sky. This points out the necessity of creating a random sample which includes as many selection effects as possible since the estimate of $\xi(r)$ (or $w(\theta)$) is strongly dependent on RR (see (A.2)). Moreover, in several cases, optical followup of the X-ray source is not 100% complete, therefore one must carefully reproduce the mask effect. What is usually done is that to create random samples in 3D, sources are placed at the same angular position of the real sources and redshift are randomly drawn from a smoothed redshift distribution of the real sources. If instead the spectral completeness is close to 100%, then the right procedures are to occupy the survey volume with random sources drawn from a L-z dependent luminosity function and accept check if they would be observable using a sensitivity map. An important choice for obtaining a reliable estimate of $w(r_p)$ is to set $\pi_{\max}$ in the calculation of the integral above. One should avoid values of $\pi_{\max}$ too large since they would add noise to the estimate of $w(r_p)$. If, instead, $\pi_{\max}$ is too small one could not recover all the signal. Uncertainties in the ACF are usually evaluated with a bootstrap resampling technique, but it is worth noting that, in the literature, several methods are adopted for errors estimates in two-point statistics, (See, [129] for a detailed description). It is known that Poisson estimators generally underestimate the variance because they do consider that points in ACF are not statistically independent. Jackknife resampling method, where one divides the survey area in many sub fields and iteratively recomputes correlation functions by excluding one subfield at a time, generally gives a good estimate of errors. But it requires that sufficient number of almost statistically independent subfields, this is not the case for most of X-ray surveys where the source statistics is moderately low. Coil et al. [48] estimated the error bars on the two-point correlation function including both Poisson and cosmic variance errors estimated, using DEEP2 mock catalogs derived from the Millenium Run simulations.

B. Limber's Deprojection

The 2D angular correlation function (ACF) is a projection of the real-space 3D ACF of the sources along the line of sight. In the following discussions and thereafter, r is in comoving coordinates. The relation between the 2D (angular) ACF and the 3D ACF is expressed by the Limber equation (e.g., [21]). Under the assumption that the scale length of the clustering is much smaller than the distance to the object, this reduces to

$$w(\theta)N^2 = \int \left(\frac{dN}{dZ}\right)^2 \int \xi\left(\sqrt{[\mathrm{d}_A(z)\theta^2) + l^2(1+z]}\right)\left(\frac{dl}{dz}\right)^{-1} dl\, dz, \tag{B.1}$$

where $d_{\mathrm{A}}(z)$ is the angular distance, N is the total number of sources, and dN/dz is the redshift distribution (per z) of the sources. The redshift evolution of the 3D correlation function is customarily expressed by

$$\xi(r, z) = \left(\frac{r}{r_0}\right)^{-\gamma}(1+z)^{-3-\epsilon+\gamma}, \tag{B.2}$$

where $\epsilon = -3$ and $\epsilon = \gamma - 3$ correspond to the case where the correlation length is constant in physical and comoving coordinates, respectively. In these notations, the zero-redshift 3D correlation length r_0 can be related to the angular correlation length θ_0 by

$$\begin{aligned} r_0^{\gamma} &= \left(\frac{N^2}{S}\right)\theta_0^{\gamma-1}, \\ S &= H_\gamma \int \left(\frac{dN}{dZ}\right)^2 \left[\frac{cd\tau(z)}{dz}\right]^{-1} \\ &\quad d_A^{1-\gamma}(1+z)^{-3-\epsilon} dz, \\ H_\gamma &= \frac{\Gamma[(\gamma-1)/2]\Gamma(1/2)}{\Gamma(1/2)}, \end{aligned} \tag{B.3}$$

where $\tau(z)$ is the look-back time. We also define the comoving correlation length

$$r_0\left(z_{\mathrm{eff}}^{-}\right) = r_0\left(1 + z_{\mathrm{eff}}^{-}\right)^{-3-\epsilon+\gamma}, \tag{B.4}$$

at the effective redshift z_{eff}^{-}, which is the median redshift of the contribution to the angular correlation (the integrand of the second term). An essential ingredient of the deprojection process is the redshift distribution of the sources, and, when individual redshifts are not available, this is derived from integration of the luminosity function.

C. 1-Halo and 2-Halo Terms in the HOD Formalism

In the halo model approach, the two-point correlation function of AGN is the sum of two contributions: the first term (*1-halo term*) is due to the correlation between objects in the same halo and the second term (*2-halo term*) arises because of the correlation between two distinct halos:

$$\xi(r) = \xi_{1h}(r) + \xi_{2h}(r). \tag{C.1}$$

Recent articles prefer to express $w = (1 + \xi_{1h}) + \xi_{2h}$ [58, 67, 130], instead of $\xi = \xi_{1h} + \xi_{2h}$, as used in older articles. This is because $1+\xi$ represents a quantity that is proportional to the number of pairs $\propto [1 + \xi_{1h}] + [1 + \xi_{2h}]$. In this new convention, the projected correlation function ξ_{1h} represents the projection of $1 + \xi_{1h}$ rather than ξ_{1h}.

Similarly, one expresses the power spectrum of the distribution of the AGN in terms of the 1- and 2-halo term contributions:

$$P(k) = P_{1h}(k) + P_{2h}(k), \tag{C.2}$$

and then the projected correlation function as

$$\begin{aligned} w_{p,1h}\left(r_p\right) &= \int \frac{k}{2\pi} P_{1h}(k) J_0\left(kr_p\right) dk, \\ w_{p,2h}\left(r_p\right) &= \int \frac{k}{2\pi} P_{2h}(k) J_0\left(kr_p\right) dk, \end{aligned} \tag{C.3}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind.

Several parametrizations exist in literature for representing the DMH profile [66, 131, 132], and the Navarro et al. [6] (NFW) profile is a popular choice. If $y(k, M_h)$ expresses the Fourier transform of the NFW profile of the DMH with mass M_h, normalized such that volume integral up to the virial radius is unity, then the one-halo term of the power spectrum can be written as

$$P_{1h}(k) = \frac{1}{n_{\mathrm{AGN}}^2} \int n(M_h) N(M_h) \left|y(k, M_h)\right|^2 dM_h. \tag{C.4}$$

Assuming the linear halo bias model [68], the two-halo term of the power spectrum reduces to

$$P_{2h}(k) = P_m(k) \left[\frac{1}{n_{\mathrm{AGN}}} \int n(M_h) b(M_h) y(k, M_h) dM_h\right]^2. \tag{C.5}$$

Since the clustering on large scales is dominated by the two-halo term, it is fairly insensitive to the assumption of AGN distribution inside the hosting halo [75]. It should be noted that since $y \sim 1$ on large scales (e.g., scales much larger than the virial radius of halos), on such scales the two-halo term can be rewritten as

$$P_{2h}(k) \approx b^2 P_m(k, z), \tag{C.6}$$

or, in terms of projected correlation function,

$$w_{p,2h}\left(r_p\right) = b^2 w_{m,2h}\left(r_p\right), \tag{C.7}$$

where b is the bias parameter of the sample and $w_{m,2h}$ is the DM-projected correlation function. For the matter power spectrum, $P_m(k)$, one can use the primordial power spectrum with a fixed n_s and a transfer function calculated using the fitting formula of Eisenstein and Hu [133] or the nonlinear form given by Smith et al. [134] and Tinker et al. [58].

D. Bias Parameter Calculation

In the majority of works on clustering of X-ray AGN [24, 33, 34, 48, 50, 51], the standard approaches used to estimate the bias are based on the power-law fit parameters of the AGN correlation function. This method assumes that the projected correlation function is well fitted by a power-law and the bias factors are derived from the best fit parameters r_0 and γ of the clustering signal at large scale. Using the power-law fit, one can estimate the AGN bias factor using the power-law best fit parameters:

$$b_{\mathrm{PL}} = \frac{\sigma_{8,\mathrm{AGN}}(z)}{\sigma_{\mathrm{DM}}(z)}, \tag{D.1}$$

where $\sigma_{8,\mathrm{AGN}}(z)$ is the rms fluctuations of the density distribution over the sphere with a comoving radius of $8\,\mathrm{Mpc}\,h^{-1}$, $\sigma_{\mathrm{DM}}(z)$ is the dark matter correlation function evaluated at $8\,\mathrm{Mpc}\,h^{-1}$, normalized to a value of σ_{DM} $(z = 0) = 0.8$. For a power-law correlation function, this value can be calculated by [21]:

$$\left(\sigma_{8,\mathrm{AGN}}\right)^2 = J_2(\gamma) \left(\frac{r_0}{8\,\mathrm{Mpc}\,h^{-1}}\right)^{\gamma}, \tag{D.2}$$

where $J_2(\gamma) = 72/[(3-\gamma)(4-\gamma)(6-\gamma)2^{\gamma}]$.

Differently in the halo model approach, the 2-halo term of the projected correlation function, which dominates at large scales, can be considered in the regime of linear density fluctuations. In the linear regime, AGNs are biased tracers of the dark matter distribution and the bias factor is described by:

$$b = \left(\frac{w_{p,1h}\left(r_p\right)}{w_{m,2h}\left(r_p\right)}\right)^{1/2}. \tag{D.3}$$

HOD modeling is currently the optimal method to establish the large-scale bias parameter, provided the parametrization of $N(M_h)$, by using

$$b = \frac{\int b_h(M_h) N(M_h) n(M_h) dM_h}{\int N(M_h) n(M_h) dM_h} \tag{D.4}$$

assuming the halo mass function $n(M_h)$ and the halo bias factor $b(M_h)$.

In fact, power-law fit bias measurements commonly use smaller scales (<1-2 h^{-1} Mpc) that are in the 1-halo term in order to increase the statistical significance. If power-law fits are restricted only to larger scales, the method suffers from the problem that the lowest scale, where the linear biasing scheme can still be applied, varies from sample to sample and remains ambiguous.

HOD modeling allows, in principle, the use of the full range of scales since the method first determines the 1- and 2-halo terms and then constrains the linear using data down to the smallest r_p values that are dominated by the 2-halo term for each individual sample.

Krumpe et al. [52] estimated the RASS-AGN bias following the power-law (D.1) and the HOD (D.4) approach, pointing out that, using the first method, the errors on the bias are much larger, but the values are statistically consistent which those derived from the HOD model fits. Allevato et al. [45] found similar results in estimating the COSMOS-AGN bias following (D.1) and (D.3).

In order to derive a reliable picture of AGN clustering, bias parameters should be inferred from HOD modeling, or at least from the comparison of the correlation function with that of the DM only in the linear regime, because systematic errors based on power-law bias parameters will be larger than the statistical uncertainties of the clustering measurement.

Acknowledgments

N. Cappelluti thanks the INAF-Fellowship program for support. N. Cappelluti thanks the Della Riccia and Blanceflor-Lodovisi-Boncompagni foundation for partial support. V. Allevato is supported by the DFG cluster of excellence Origin and Structure of the Universe (http://www.universe-cluster.de/). N. Cappelluti, V. Allevato, and A. Finoguenov thank the referees, Ryan Hickox and Manolis Plionis, for valuable suggestions for improving the paper.

References

[1] R. Giacconi, J. Bechtold, G. Branduardi et al., "A high-sensitivity X-ray survey using the Einstein Observatory and the discrete source contribution to the extragalactic X-ray background," *Astrophysical Journal*, vol. 234, pp. L1–L17, 1979.

[2] J. Magorrian, S. Tremaine, D. Richstone et al., "The demography of massive dark objects in galaxy centers," *Astronomical Journal*, vol. 115, no. 6, pp. 2285–2305, 1998.

[3] N. I. Shakura and R. A. Sunyaev, "A theory of the instability of disk accretion on to black holes and the variability of binary X-ray sources, galactic nuclei and quasars," *Monthly Notices of the Royal Astronomical Society*, vol. 175, pp. 613–32, 1976.

[4] R. Bergamini, P. Londrillo, and G. Setti, "The cosmic black-body radiation and the inverse compton effect in the radio galaxies: the X-ray background," *Il Nuovo Cimento B Series*, vol. 52, no. 2, pp. 495–506, 1967.

[5] G. Hasinger, R. Burg, R. Giacconi et al., "A deep X-ray survey in the lockman-hole and the soft X-ray N-log," *Astronomy and Astrophysics*, vol. 275, no. 1, p. 1, 1993.

[6] J. F. Navarro, C. S. Frenk, and S. D. M. White, "A universal density profile from hierarchical clustering," *Astrophysical Journal*, vol. 490, no. 2, pp. 493–508, 1997.

[7] P. F. Hopkins, A. Lidz, L. Hernquist et al., "The co-formation of spheroids and quasars traced in their clustering," *Astrophysical Journal*, vol. 662, no. 1, pp. 110–130, 2007.

[8] P. F. Hopkins, L. Hernquist, T. J. Cox, and D. Kerbs, "A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. i. galaxy mergers and quasar activity," *Astrophysical Journal, Supplement Series*, vol. 175, no. 2, pp. 356–389, 2008.

[9] P. F. Hopkins and L. Hernquist, "A characteristic division between the fueling of quasars and seyferts: five simple tests," *Astrophysical Journal*, vol. 694, no. 1, pp. 599–609, 2009.

[10] J. D. Silverman, P. Kampczyk, K. Jahnke et al., "The impact of galaxy interactions on active galactic nucleus activity in zCOSMOS," *The Astrophysical Journal*, vol. 743, no. 1, article 2, 2011.

[11] D. H. McIntosh, Y. Guo, H. J. Mo, F. van den Bosch, and X. Yan, "The SDSS view of galaxy mergers and their environments," *Bulletin of the American Astronomical Society*, vol. 41, p. 244, 2009.

[12] M. Milosavljević, D. Merritt, and L. C. Ho, "Contribution of stellar tidal disruptions to the X-ray luminosity function of active galaxies," *Astrophysical Journal*, vol. 652, no. 1, pp. 120–125, 2006.

[13] P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, and V. Springel, "A unified, merger-driven model of the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes, and galaxy spheroids," *Astrophysical Journal, Supplement Series*, vol. 163, no. 1, pp. 1–49, 2006.

[14] D. Larson, J. Dunkley, G. Hinshaw et al., "Seven-year wilkinson microwave anisotropy probe (WMAP) observations: power spectra and WMAP-derived parameters," *Astrophysical Journal, Supplement Series*, vol. 192, article 16, 2011.

[15] X. Barcons and A. C. Fabian, "Fluctuations in the X-ray background and the large-scale structure of the universe," *Monthly Notices of the Royal Astronomical Society*, vol. 230, pp. 189–206, 1988.

[16] F. J. Carrera and X. Barcons, "The spatial distribution of cosmic X-ray sources from the isotropy of the soft X-ray background," *Monthly Notices of the Royal Astronomical Society*, vol. 257, no. 3, pp. 507–512, 1992.

[17] I. Georgantopoulos, G. C. Stewart, T. Shanks, R. E. Griffiths, and B. J. Boyle, "A deep ROSAT survey. II—observations of the isotropy of the 1-2 keV X-ray background," *Monthly Notices of the Royal Astronomical Society*, vol. 262, no. 3, pp. 619–626, 1993.

[18] A. Soltan and G. Hasinger, "The angular correlation function of the soft X-ray background," *Astronomy and Astrophysics*, vol. 288, no. 1, pp. 77–88, 1994.

[19] B. J. Boyle and H. J. Mo, "The clustering of QSOs at low redshift," *Monthly Notices of the Royal Astronomical Society*, vol. 260, no. 4, pp. 925–928, 1993.

[20] A. Vikhlinin and W. Forman, "Detection of the angular correlation of faint X-ray sources," *Astrophysical Journal*, vol. 455, no. 2, pp. L109–L113, 1995.

[21] P. J. E. Peebles, *The Large Scale Structure of the Universe*, Princeton University Press, Princeton, NJ, USA, 1980.

[22] A. Akylas, I. Georgantopoulos, and M. Plionis, "The angular correlation function of the ROSAT all-sky survey bright source catalogue," *Monthly Notices of the Royal Astronomical*

Society, vol. 318, no. 4, pp. 1036–1040, 2000.

[23] X. Barcons, F. J. Carrera, M. T. Ceballos, and S. Mateos, "X-ray sources as tracers of the large-scale structure in the universe ," in *Proceedings of the X-Ray Astronomy: Stellar Endpoints,AGN, and the Diffuse X-ray Background*, vol. 599, pp. 3–12, 2001.

[24] C. R. Mullis, J. P. Henry, I. M. Gioia et al., "Spatial correlation function of X-ray-selected active galactic nuclei," *Astrophysical Journal*, vol. 617, no. 1, pp. 192–208, 2004.

[25] M. Cappi, P. Mazzotta, M. Elvis et al., "Chandra study of an overdensity of x-ray sources around two distant ($z \sim 0.5$) clusters," *Astrophysical Journal*, vol. 548, no. 2, pp. 624–638, 2001.

[26] N. Cappelluti, M. Cappi, M. Dadina et al., "X-ray source overdensities in *Chandra* distant cluster fields: a new probe to map the cosmic tapestry?" *Astronomy and Astrophysics*, vol. 430, no. 1, pp. 39–45, 2005.

[27] E. Koulouridis and M. Plionis, "Luminous X-ray active galactic nuclei in clusters of galaxies," *Astrophysical Journal*, vol. 714, no. 2, pp. L181–L184, 2010.

[28] M. Koss, R. Mushotzky, S. Veilleux, and L. Winter, "Merging and clustering of the swift bat AGN sample," *Astrophysical Journal*, vol. 716, no. 2, pp. L125–L130, 2010.

[29] S. Basilakos, M. Plionis, A. Georgakakis et al., "The XMM-Newton/2dF survey—III. Comparison between optical and X-ray cluster detection methods," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 3, pp. 989–996, 2004.

[30] S. Basilakos, M. Plionis, A. Georgakakis, and I. Georgantopoulos, "The XMM-Newton/2dF survey—VI. Clustering and bias of the soft X-ray point sources," *Monthly Notices of the Royal Astronomical Society*, vol. 356, no. 1, pp. 183–191, 2005.

[31] P. Gandhi, O. Garcet, L. Disseau et al., "The XMM large scale structure survey: properties and two-point angular correlations of point-like sources," *Astronomy and Astrophysics*, vol. 457, no. 2, pp. 393–404, 2006.

[32] S. Puccetti, F. Flore, V. D'Elia et al., "The XMM-Newton survey of the ELAIS-S1 field I. Number counts, angular correlation function and X-ray spectral properties," *Astronomy and Astrophysics*, vol. 457, no. 2, pp. 501–515, 2006.

[33] R. Gilli, E. Daddi, G. Zamorani et al., "The spatial clustering of X-ray selected AGN and galaxies in the Chandra Deep Field South and North," *Astronomy and Astrophysics*, vol. 430, no. 3, pp. 811–825, 2005.

[34] Y. Yang, R. F. Mushotzky, A. J. Barger, and L. L. Cowie, "Spatial correlation function of the Chandra-selected active galactic nuclei," *Astrophysical Journal*, vol. 645, no. 1, pp. 68–82, 2006.

[35] G. Hasinger, N. Cappelluti, H. Brunner et al., "The XMM-Newton wide-field survey in the COSMOS field. I. Survey description," *Astrophysical Journal, Supplement Series*, vol. 172, no. 1, pp. 29–37, 2007.

[36] N. Cappelluti, G. Hasinger, M. Brusa et al., "The XMM-Newton wide-field survey in the COSMOS field. II. X-ray data and the log N-log S relations," *Astrophysical Journal, Supplement Series*, vol. 172, no. 1, pp. 341–352, 2007.

[37] N. Cappelluti, M. Brusa, G. Hasinger et al., "The XMM-Newton wide-field survey in the COSMOS field," *Astronomy and Astrophysics*, vol. 497, no. 2, pp. 635–648, 2009.

[38] M. Elvis and Chandra-COSMOS Team, "The chandra-COSMOS survey: first results," *Bullettin of the American Astronomical Society*, vol. 39, no. 4, p. 899, 2007.

[39] S. Puccetti, C. Vignali, and N. Cappelluti, "The chandra survey of the cosmos field. II. source detection and photometry," *The Astrophysical Journal Supplement Series*, vol. 185, no. 2, article 586, 2009.

[40] S. J. Lilly, O. Le Fèvre, A. Renzini et al., "zCOSMOS: a large VLT/VIMOS redshift survey covering $0 < z < 3$ in the COSMOS field 1," *Astrophysical Journal, Supplement Series*, vol. 172, no. 1, pp. 70–85, 2007.

[41] T. Miyaji, G. Zamorani, N. Cappelluti et al., "The XMM-Newton wide-field survey in the COSMOS field. V angular clustering of the X-ray point sources," *Astrophysical Journal, Supplement Series*, vol. 172, no. 1, pp. 396–405, 2007.

[42] J. Ebrero, S. Mateos, G. C. Stewart, F. J. Carrera, and M. G. Watson, "High-precision multi-band measurements of the angular clustering of X-ray sources," *Astronomy and Astrophysics*, vol. 500, no. 2, pp. 749–762, 2009.

[43] A. Elyiv, N. Clerc, M. Plionis et al., "Angular correlation functions of X-ray point-like sources in the full exposure XMM-LSS field," *Astronomy and Astrophysics*, vol. 537, article A131, 2012.

[44] R. Gilli, G. Zamorani, T. Miyaji et al., "The spatial clustering of X-ray selected AGN in the XMM-COSMOS field," *Astronomy and Astrophysics*, vol. 494, no. 1, pp. 33–48, 2009.

[45] V. Allevato, A. Finoguenov, N. Cappelluti et al., "The XMM-newton wide field survey in the cosmos field: redshift evolution of agn bias and subdominant role of mergers in triggering moderate-luminosity AGNs at redshifts up to 2.2," *Astrophysical Journal*, vol. 736, no. 2, article 99, 2011.

[46] S. Bonoli, F. Marulli, V. Springel, S. D. M. White, E. Branchini, and L. Moscardini, "Modelling the cosmological co-evolution of supermassive black holes and galaxies—II. the clustering of quasars and their dark environment," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 1, pp. 423–438, 2009.

[47] R. C. Hickox, C. Jones, W. R. Forman et al., "Host galaxies, clustering, Eddington ratios, and evolution of radio, X-ray, and infrared-selected AGNs," *Astrophysical Journal Letters*, vol. 696, no. 1, pp. 891–919, 2009.

[48] A. L. Coil, A. Georgakakis, J. A. Newman et al., "Aegis: the clustering of x-ray active galactic nucleus relative to galaxies at $z \sim 1$," *Astrophysical Journal Letters*, vol. 701, no. 2, pp. 1484–1499, 2009.

[49] G. Mountrichas and A. Georgakakis, "The clustering of X-ray-selected active galactic nuclei at $z = 0.1$," *Monthly Notices of the Royal Astronomical Society*, vol. 420, no. 1, pp. 514–525, 2012.

[50] M. Krumpe, T. Miyaji, and A. L. Coil, "The spatial clustering of rosat all-sky survey AGNs. I. the cross-correlation function with sdss luminous red galaxies," *Astrophysical Journal*, vol. 713, no. 1, pp. 558–572, 2010.

[51] N. Cappelluti, M. Burlon D. Aiello et al., "Active galactic nuclei clustering in the local universe: an unbiased picture from swift-BAT," *The Astrophysical Journal Letters*, vol. 716, no. 2, pp. L209–L213, 2010.

[52] M. Krumpe, T. Miyaji, A. L. Coil, and H. Aceves, "The spatial clustering of ROSAT All-Sky Survey active galactic nuclei. III. Expanded sample and comparison with optical active galactic nuclei," *Astrophysical Journal*, vol. 746, article 1, 2012.

[53] M. Plionis, M. Rovilos, S. Basilakos, I. Georgantopoulos, and F. Bauer, "Luminosity-dependent X-ray active galactic nucleus clustering?" *The Astrophysical Journal*, vol. 674, no. 1, pp. L5–L8, 2008.

[54] F. C. van den Bosch, "The universal mass accretion history

of cold dark matter haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 331, no. 1, pp. 98–110, 2002.

[55] R. K. Sheth, H. J. Mo, and G. Tormen, "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 323, no. 1, pp. 1–12, 2001.

[56] W. H. Press and P. Schechter, "Formation of galaxies and clusters of galaxies by self-similar gravitational condensation," *Astrophysical Journal*, vol. 187, pp. 425–438, 1974.

[57] R. K. Sheth and G. Tormen, "Large-scale bias and the peak background split," *Monthly Notices of the Royal Astronomical Society*, vol. 308, no. 1, pp. 119–126, 1999.

[58] J. L. Tinker, D. H. Weinberg, Z. Zheng, and I. Zehavi, "On the mass-to-light ratio of large-scale structure," *Astrophysical Journal*, vol. 631, no. 1, pp. 41–58, 2005.

[59] N. P. Ross, Y. Shen, M. A. Strauss et al., "Clustering of low-redshift ($z \leq 2.2$) quasars from the sloan digital sky survey," *Astrophysical Journal*, vol. 697, no. 2, pp. 1634–1655, 2009.

[60] Y. Shen, M. A. Strauss, N. P. Ross et al., "Quasar clustering from SDSS DR5: dependences on physical properties," *Astrophysical Journal Letters*, vol. 697, no. 2, pp. 1656–1673, 2009.

[61] S. M. Croom, B. J. Boyle, T. Shanks et al., "The 2dF QSO redshift survey - XIV. Structure and evolution from the two-point correlation function," *Monthly Notices of the Royal Astronomical Society*, vol. 356, no. 2, pp. 415–438, 2005.

[62] C. Porciani and P. Norberg, "Luminosity- and redshift-dependent quasar clustering," *Monthly Notices of the Royal Astronomical Society*, vol. 371, no. 4, pp. 1824–1834, 2006.

[63] J. Da Ângela, T. Shanks, S. M. Croom et al., "The 2dF-SDSS LRG and QSO survey: QSO clustering and the L-z degeneracy," *Monthly Notices of the Royal Astronomical Society*, vol. 383, no. 2, pp. 565–580, 2008.

[64] G. Kauffmann, A. Nusser, and M. Steinmetz, "Galaxy formation and large-scale bias," *Monthly Notices of the Royal Astronomical Society*, vol. 286, no. 4, pp. 795–811, 1997.

[65] J. A. Peacock and R. E. Smith, "Halo occupation numbers and galaxy bias," *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 4, pp. 1144–1156, 2000.

[66] A. Cooray and R. Sheth, "Halo models of large scale structure," *Physics Report*, vol. 372, no. 1, pp. 1–129, 2002.

[67] Z. Zheng, A. A. Berlind, D. H. Weinberg et al., "Theoretical models of the halo occupation distribution: separating central and satellite galaxies," *Astrophysical Journal*, vol. 633, no. 2, pp. 791–809, 2005.

[68] H. J. Mo and S. D. M. White, "An analytic model for the spatial clustering of dark matter haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 282, no. 2, pp. 347–361, 1996.

[69] S. Basilakos, M. Plionis, and C. Ragone-Figueroa, "The halo mass-bias redshift evolution in the ACDM cosmology," *Astrophysical Journal*, vol. 678, no. 2, pp. 627–634, 2008.

[70] J. L. Tinker, B. E. Robertson, A. V. Kravtsov et al., "The large-scale bias of dark matter halos: numerical calibration and model tests," *Astrophysical Journal*, vol. 724, no. 2, pp. 878–886, 2010.

[71] A. Pillepich, C. Porciani, and O. Hahn, "Halo mass function and scale-dependent bias from N-body simulations with non-Gaussian initial conditions," *Monthly Notices of the Royal Astronomical Society*, vol. 402, no. 1, pp. 191–206, 2010.

[72] C.-P. Ma, M. Maggiore, A. Riotto, and J. Zhang, "The bias and mass function of dark matter haloes in non-Markovian extension of the excursion set theory," *Monthly Notices of the Royal Astronomical Society*, vol. 411, no. 4, pp. 2644–2652, 2011.

[73] Seljak, "Analytic model for galaxy and dark matter clustering," *Monthly Notices of the RoyalAstronomical Society*, vol. 318, p. 2035, 2000.

[74] R. Scoccimarro, R. K. Sheth, L. Hui, and B. Jain, "How many galaxies fit in a halo? Constraints on galaxy formation efficiency from spatial clustering," *Astrophysical Journal*, vol. 546, no. 1, pp. 20–34, 2001.

[75] A. A. Berlind and D. H. Weinberg, "The halo occupation distribution: toward an empirical determination of the relation between galaxies and mass," *Astrophysical Journal*, vol. 575, no. 2, pp. 587–616, 2002.

[76] C. Marinoni and M. J. Hudson, "The mass-to-light function of virialized systems and the relationship between their optical and X-ray properties," *Astrophysical Journal*, vol. 569, no. 1, pp. 101–111, 2002.

[77] M. Magliocchetti and C. Porciani, "The halo distribution of 2dF galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 346, no. 1, pp. 186–198, 2003.

[78] F. C. van den Bosch, X. Yang, and H. J. Mo, "Linking early- and late-type galaxies to their dark matter haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 340, no. 3, pp. 771–792, 2003.

[79] X. Yang, H. J. Mo, and F. C. Van den Bosch, "Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 339, no. 4, pp. 1057–1080, 2003.

[80] I. Zehavi, D. H. Weinberg, Z. Zheng et al., "On departures from a power law in the galaxy correlation function," *Astrophysical Journal Letters*, vol. 608, no. 1, pp. 16–24, 2004.

[81] S. Phleps, J. A. Peacock, K. Meisenheimer, and C. Wolf, "Galaxy clustering from COMBO-17: the halo occupation distribution at z = 0.6," *Astronomy and Astrophysics*, vol. 457, no. 1, pp. 145–155, 2006.

[82] Z. Zheng, I. Zehavi, D. J. Eisenstein, D. H. Weinberg, and Y. P. Jing, "Halo occupation distribution modeling of clustering of luminous red galaxies," *Astrophysical Journal*, vol. 707, no. 1, pp. 554–572, 2009.

[83] J. S. Bullock, R. H. Wechsler, and R. S. Somerville, "Galaxy halo occupation at high redshift," *Monthly Notices of the Royal Astronomical Society*, vol. 329, no. 1, pp. 246–256, 2002.

[84] L. A. Moustakas and R. S. Somerville, "The masses, ancestors, and descendants of extremely red objects: constraints from spatial clustering," *Astrophysical Journal*, vol. 577, no. 1, pp. 1–10, 2002.

[85] T. Hamana, M. Ouchi, K. Shimasaku, I. Kayo, and Y. Suto, "Properties of host haloes of Lyman-break galaxies and Lyman α emitters from their number densities and angular clustering," *Monthly Notices of the Royal Astronomical Society*, vol. 347, no. 3, pp. 813–823, 2004.

[86] Z. Zheng, "Interpreting the observed clustering of red galaxies at $z \sim 3$," *Astrophysical Journal Letters*, vol. 610, no. 1, pp. 61–68, 2004.

[87] Z. Zheng, A. L. Coil, and I. Zehavi, "Galaxy evolution from halo occupation distribution modeling of DEEP2 and SDSS galaxy clustering," *Astrophysical Journal*, vol. 667, no. 2, pp. 760–779, 2007.

[88] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler et al., "The dark side of the halo occupation distribution," *Astrophysical Journal*, vol. 609, no. 1, pp. 35–49, 2004.

[89] C. Porciani, M. Magliocchetti, and P. Norberg, "Cosmic evolution of quasar clustering: implications for the host haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 355, no. 3, pp. 1010–1030, 2004.

[90] N. Padmanabhan, M. White, P. Norberg, and C. Porciani, "The real-space clustering of luminous red galaxies around z < 0.6 quasars in the Sloan Digital Sky Survey," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 4, pp. 1862–1875, 2009.

[91] Y. Shen, J. F. Hennawi, F. Shankar et al., "Binary quasars at high redshift. II. Sub-Mpc clustering at z ~ 3-4," *Astrophysical Journal Letters*, vol. 719, no. 2, pp. 1693–1698, 2010.

[92] S. Starikova, R. Cool, D. Eisenstein et al., "Constraining halo occupation properties of x-ray active galactic nuclei using clustering of Chandra sources in the Botes survey region," *Astrophysical Journal*, vol. 741, no. 1, article 15, 2011.

[93] T. Miyaji, M. Krumpe, A. L. Coil, and H. Aceves, "The spatial clustering of rosat all-sky survey AGNs. II. halo occupation distribution modeling of the cross-correlation function," *Astrophysical Journal Letters*, vol. 726, no. 2, article 83, 2011.

[94] T. J. Arnold, P. Martini, J. S. Mulchaey, A. Berti, and T. E. Jeltema, "Active galactic nuclei in groups and clusters of galaxies: detection and host morphology," *Astrophysical Journal*, vol. 707, no. 2, pp. 1691–1706, 2009.

[95] P. Martini, G. R. Sivakoff, and J. S. Mulchaey, "The evolution of active galactic nuclei in clusters of galaxies to redshift 1.3," *Astrophysical Journal*, vol. 701, no. 1, pp. 66–85, 2009.

[96] J. D. Silverman, K. Kova, and C. Knobel, "The environments of active galactic nuclei within the zCOSMOS density field," *The Astrophysical Journal*, vol. 695, no. 1, p. 171, 2009.

[97] C. Li, G. Kauffmann, L. Wang, S. D. M. White, T. M. Heckman, and Y. P. Jing, "The clustering of narrow-line AGN in the local Universe," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 2, pp. 457–468, 2006.

[98] A. L. Coil, J. F. Hennawi, J. A. Newman, M. C. Cooper, and M. Davis, "The DEEP2 galaxy redshift survey: clustering of quasars and galaxies at z = 1," *Astrophysical Journal Letters*, vol. 654, no. 1, pp. 115–124, 2007.

[99] G. Mountrichas, U. Sawangwit, T. Shanks et al., "QSO-LRG two-point cross-correlation function and redshift-space distortions," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 4, pp. 2050–2064, 2009.

[100] R. C. Hickox, A. D. Myers, M. Brodwin et al., "Clustering of obscured and unobscured quasars in the Boötes field: placing rapidly growing black holes in the cosmic web," *Astrophysical Journal Letters*, vol. 731, no. 2, article 117, 2011.

[101] P. Martini and D. H. Weinberg, "Quasar clustering and the lifetime of quasars," *Astrophysical Journal*, vol. 547, no. 1, pp. 12–26, 2001.

[102] A. C. Fabian, R. V. Vasudevan, R. F. Mushotzky, L. M. Winter, and C. S. Reynolds, "Radiation pressure and absorption in AGN: results from a complete unbiased sample from *Swift*," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 1, pp. L89–L92, 2009.

[103] P. F. Hopkins, T. J. Cox, D. Kerbš, and L. Hernquist, "A cosmological framework for the co-evolution of Quasars, supermassive black holes, and elliptical galaxies. II. Formation of red ellipticals," *Astrophysical Journal, Supplement Series*, vol. 175, no. 2, pp. 390–422, 2008.

[104] G. Hasinger, "Absorption properties and evolution of active galactic nuclei," *Astronomy and Astrophysics*, vol. 490, no. 3, pp. 905–922, 2008.

[105] Y. Shen, "Supermassive black holes in the hierarchical universe: a general framework and observational tests," *Astrophysical Journal Letters*, vol. 704, no. 1, pp. 89–108, 2009.

[106] F. Shankar, D. H. Weinberg, and J. Miralda-Escudé, "Self-consistent models of the AGN and black hole populations: duty cycles, accretion rates, and the mean radiative efficiency," *Astrophysical Journal*, vol. 690, no. 1, pp. 20–41, 2009.

[107] F. Shankar, C. Martin, M.-E. Jordi, F. Pablo, H. W. David et al., "On the radiative efficiencies, eddington ratios, and duty cycles of luminous high-redshift quasars," *The Astrophysical Journal* , vol. 718, no. 1, article 231, 2010.

[108] F. Shankar, "Merger-induced quasars, their light curves, and their host halos," in *Proceedings of the International Astronomical Union (IAUS '10)*, vol. 267, pp. 248–253, 2010.

[109] E. Treister, C. M. Urry, K. Schawinski, C. N. Cardamone, and D. B. Sanders, "Heavily obscured active galactic nuclei in high-redshift luminous infrared galaxies," *Astrophysical Journal*, vol. 722, no. 2, pp. L238–L243, 2010.

[110] A. Georgakakis, K. Nandra, E. S. Laird et al., "AEGIS: the environment of X-ray sources at z ~ 1," *Astrophysical Journal Letters*, vol. 660, no. 1, pp. L15–L18, 2007.

[111] A. Georgakakis, A. L. Coil, E. S. Laird et al., "Host galaxy morphologies of X-ray selected AGN: assessing the significance of different black hole fuelling mechanisms to the accretion density of the Universe at z ~ 1," *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 2, pp. 623–633, 2009.

[112] M. Cisternas, K. Jahnke, K. J. Inskip et al., "The bulk of the black hole growth since z ~ 1 occurs in a secular universe: no major merger-AGN connection," *Astrophysical Journal Letters*, vol. 726, no. 2, article 57, 2011.

[113] J. S. Dunlop, R. J. McLure, M. J. Kukula, S. A. Baum, C. P. O'Dea, and D. H. Hughes, "Quasars, their host galaxies and their central black holes," *Monthly Notices of the Royal Astronomical Society*, vol. 340, no. 4, pp. 1095–1135, 2003.

[114] N. A. Grogin, C. J. Conselice, E. Chatzichristou et al., "AGN host galaxies at z ~ 0.4-1.3: bulge-dominated and lacking merger-AGN connection," *Astrophysical Journal*, vol. 627, no. 2, pp. L97–L100, 2005.

[115] C. M. Pierce, J. M. Lotz, E. S. Laird et al., "AEGIS: host galaxy morphologies of X-ray-selected and infrared-selected active galactic nuclei at $0.2 \leq z < 1.2$," *Astrophysical Journal Letters*, vol. 660, no. 1, pp. L19–L22, 2007.

[116] J. M. Gabor, C. D. Impey, K. Jahnke et al., "Active galactic nucleus host galaxy morphologies in cosmos," *The Astrophysical Journal*, vol. 691, no. 1, article 705, 2009.

[117] T. A. Reichard, T. M. Heckmas, and G. Rudnick, "The lopsidedness of present-day galaxies: connections to the formation of stars, the chemical evolution of galaxies, and the growth of black holes," *The Astrophysical Journal*, vol. 691, no. 2, article 1005, 2009.

[118] T. Tal, P. G. Van Dokkum, J. Nelan, and R. Bezanson, "The frequency of tidal features associated with nearby luminous elliptical galaxies from a statistically complete sample," *Astronomical Journal*, vol. 138, no. 5, pp. 1417–1427, 2009.

[119] D. J. Schlegel, F. Abdalla, and T. Abraham, "The BigBOSS experiment," *Astrophysics*, vol. 1, no. 11, p. 2865, 2011.

[120] D. J. Rosario, R. C. McGurk, and C. E. Max, "Adaptive optics imaging of QSOs with double-peaked narrow lines: are they dual AGNs?" *The Astrophysical Journal*, vol. 739, no. 1, article 44, 2011.

[121] K. Schawinski, M. Urry, E. Treister, B. Simmons, P. Natarajan, and E. Glikman, "Evidence for three accreting black holes in a galaxy at z ∼ 1.35: a snapshot of recently formed black hole seeds?" *Astrophysical Journal Letters*, vol. 743, no. 2, article L37, 2011.

[122] D. D. Kocevski, S. M. Faber, M. Mozena et al., "Candels: constraining the AGN-merger connection with host morphologies at z ∼ 2," *Astrophysical Journal*, vol. 744, no. 2, article 148, 2012.

[123] D. J. Eisenstein, D. H. Weinberg, and E. Agol, "SDSS-III: massive spectroscopic surveys of the distant universe, the milky way galaxy, and extra-solar planetary systems," *The Astronomical Journal*, vol. 142, no. 3, article 72, 2011.

[124] P. Predehl, R. Andritschke, W. Bornemann et al., "eROSITA," in *UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XV*, vol. 6686 of *Proceedings of SPIE*, August 2007.

[125] Z. Ivezic, J. A. Tyson, E. Acosta et al., "LSST: from science drivers to reference design and anticipated data products," *Bulletin of the American Astronomical Society*, vol. 41, p. 366, 2008.

[126] M. Davis and P. J. E. Peebles, "A survey of galaxy redshifts. V—the two-point position and velocity correlations," *Astrophysical Journal*, vol. 267, pp. 465–482, 1983.

[127] K. B. Fisher, M. Davis, M. A. Strauss, A. Yahil, and J. P. Huchra, "Clustering in the 1.2 Jy IRAS galaxy redshift survey II: redshift distortions and $\xi(r_p, \pi)$," *Monthly Notices of the Royal Astronomical Society*, vol. 267, pp. 927–948, 1994.

[128] S. D. Landy and A. S. Szalay, "Bias and variance of angular correlation functions," *Astrophysical Journal*, vol. 412, no. 1, pp. 64–71, 1993.

[129] P. Norberg, C. M. Baugh, E. Gaztañaga, and D. J. Croton, "Statistical analysis of galaxy surveys—I. Robust error estimation for two-point clustering statistics," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 1, pp. 19–38, 2009.

[130] C. Blake, A. Collister, and O. Lahav, "Halo-model signatures from 380 000 Sloan Digital Sky Survey luminous red galaxies with photometric redshifts," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 3, pp. 1257–1269, 2008.

[131] S. R. Knollmann, C. Power, and A. Knebe, "Dark matter halo profiles in scale-free cosmologies," *Monthly Notices of the Royal Astronomical Society*, vol. 385, no. 2, pp. 545–552, 2008.

[132] J. Stadel, D. Potter, B. Moore et al., "Quantifying the heart of darkness with GHALO—a multibillion particle simulation of a galactic halo," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 1, pp. L21–L25, 2009.

[133] D. J. Eisenstein and W. Hu, "Power spectra for cold dark matter and its variants," *Astrophysical Journal Letters*, vol. 511, no. 1, pp. 5–15, 1999.

[134] R. E. Smith, J. A. Peacock, A. Jenkins et al., "Stable clustering, the halo model and non-linear cosmological power spectra," *Monthly Notices of the Royal Astronomical Society*, vol. 341, no. 4, pp. 1311–1332, 2003.