

Research Article

A Principle Component Analysis of Galaxy Properties from a Large, Gas-Selected Sample

Yu-Yen Chang,^{1,2,3} Rikon Chao,^{2,4} Wei-Hao Wang,³ and Pisin Chen^{1,2,5}

¹Department of Physics and Graduate Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan

²Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan

³Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan

⁴Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan

⁵Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94305, USA

Correspondence should be addressed to Yu-Yen Chang, rainswl@gmail.com

Received 11 January 2012; Revised 17 March 2012; Accepted 19 March 2012

Academic Editor: Gary Wegner

Copyright © 2012 Yu-Yen Chang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disney et al. (2008) have found a striking correlation among global parameters of HI-selected galaxies and concluded that this is in conflict with the CDM model. Considering the importance of the issue, we reinvestigate the problem using the principal component analysis on a fivefold larger sample and additional near-infrared data. We use databases from the Arecibo Legacy Fast Arecibo *L*-band Feed Array Survey for the gas properties, the Sloan Digital Sky Survey for the optical properties, and the Two Micron All Sky Survey for the near-infrared properties. We confirm that the parameters are indeed correlated where a single physical parameter can explain 83% of the variations. When color (*g-i*) is included, the first component still dominates but it develops a second principal component. In addition, the near-infrared color (*i-J*) shows an obvious second principal component that might provide evidence of the complex old star formation. Based on our data, we suggest that it is premature to pronounce the failure of the CDM model and it motivates more theoretical work.

1. Introduction

One way to understand our universe is to gain insights into the structure of galaxies. For one thing, it helps to reveal the role of dark matter in their formation and dynamics. The cosmological model that consists of a cosmological constant and the cold dark matter in addition to the ordinary baryon matter and radiation (Λ CDM) has been able to successfully explain the evolution of the cosmic structures especially at large scales. By measuring CMB fluctuations COBE [4, 5], WMAP [6–9], Type Ia supernovae [10], and gravitational lensing [11], this model of cosmology has by now been established as the standard model of cosmology. Among its various implications, it suggests a hierarchical, bottom-up history of structure formation that evolves from small fluctuations to galaxies, clusters, and eventually superclusters. On the other side, the success of

Λ CDM has not been as clearcut at small scales. There still exist several inconsistencies between Λ CDM and the observations at small scales. For example, the simulations based on Λ CDM have revealed more number of galactic satellites [12–15] and less number of disk galaxies [16] than what have been observed. Besides, the degree of emptiness in the voids is also inconsistent between theory and observation [17]. While the Λ CDM can explain the galaxy rotational curves at large radii [18], the relatively higher density at the galactic core than that predicted by the Λ CDM, that is, the so-called cusp-core problem, is still unresolved [19].

The hierarchical galaxy formation scenario has been actively investigated previously (e.g., [20–22]), and the tight correlation between selected galactic parameters has been studied in the past [23–30]. Moreover, the overall correlation among all major galactic variables has also been thoroughly

investigated recently by Disney et al. ([31]; hereafter D08) and Garcia-Appadoo et al. ([32]; hereafter G09). Based on 195 galaxies, the two studies found remarkably strong correlations between six galactic variables: the 90%-light radius (R_{90}), the 50%-light radius (R_{50}), the H I mass (M_{HI}), the dynamical mass (M_d), the luminosity (L), and the color. G09 found strong correlations among the galactic properties, including a common dynamical mass-to-light ratio within the optical radii, a correlation between surface brightness and luminosity, and a common H I surface density. D08 further showed that all the parameters can be included in one correlation and suggested that this is in conflict with the notion of the hierarchical structure formation scenario and the Λ CDM model.

Considering the importance of this issue, we set out to reinvestigate the analysis made by D08 and G09. We note that their claim actually involves two separate subissues. First, there is the issue of whether the galactic dynamics is indeed controlled by a single parameter. Second, even if this is true, there still remains the issue of whether such a fact necessarily concludes the failure of the hierarchical structure formation scenario and the cold dark matter model. In our view the positive conclusion of the former does not necessarily lead to the affirmative conclusion of the latter. In this paper we demonstrate tentative evidence that there may exist more than one principal component among the global parameters of galaxies with regard to the first issue. As for the second issue, we argue about the importance of the nongravitational baryon physics in galactic structure formation, which renders the naive extrapolation of the hierarchical structure formation scenario from cosmic to galactic scales questionable.

In order to scrutinize the correction issue, we performed a similar analysis on global parameters of galaxies with a significantly larger database and two additional parameters, L_J and M_J , based on the infrared J band. We include a total of 1022 galaxies from the Arecibo Legacy Fast Arecibo L -band Feed Array Survey (ALFALFA; [33–36]) for the atomic gas properties, and the Sloan Digital Sky Survey (SDSS; [37]) for the optical properties. Recently, Toribio et al. published two papers ([38, 39]; hereafter T11) with these two surveys. In T11, they assembled three samples to analyse the data. They argued that H I emission provides the most reliable way to determine the morphologies and conclude that color and H mass of gas-rich galaxies cannot be very closed. In addition, they also found the correlation between mass, radius, luminosity and ($g-r$) by principal component analysis (PCA; [40]) for one of their subsamples. Among our 1022 galaxies, 479 of them have also been detected by the Two Micron All Sky Survey (2MASS; [41]). We use these galaxies to study their near-infrared properties, which contain more stellar information.

Near-infrared studies of H I-selected galaxies had been attempted by various groups (e.g., G09; [42, 43]). Our motivation of adding the near-infrared data in our analysis is the following. The optical emission is sensitive to young stars. The near-infrared emission, on the other hand, is less affected by the young stars and is therefore a better tracer of the total stellar mass, which dominates the baryonic matter at galactic

scales. We believe that the inclusion of the infrared data would provide us additional and independent information on the baryonic mass assembly history of the galaxies.

Through the PCA, we confirm that except the color, all other observables, from H I, optical to near-infrared bands, are highly correlated and dominated by a single parameter. This is true both in the optical and in the near-infrared bands and this confirms the results in D08. In addition, we also see a second component from color, especially in our near-infrared data. Because near-infrared color provides the information of integrated star formation history, it may be an evidence for complex formation history, whereas a valid structure formation theory needs to explain this observation.

The organization of this paper is as follows. We describe the data and sources in Section 2. Then several variables are adopted and applied to statistical analysis in Section 3. In Section 4, we summarize and discuss our results.

2. Data

Our samples are the blind 21 cm survey from ALFALFA. This selects gas rich galaxies which also contain low luminosity and low surface brightness galaxies in higher proportion than those in an optical selection. The optical data for this study come from the SDSS DR7, which covers 12,000 deg² for imaging and provides spectra of 930,000 galaxies. Here we briefly describe how we select SDSS counterparts to the ALFALFA sources, and we refer to G09 for more detailed discussion about identification.

The Arecibo Telescope has a beam size of 3.5' at 21 cm. Since the majority of the H I detections have S/N > 10 [34–36], we adopted a conservative searching radius of 10". We found 1233 SDSS galaxies that appear to be detected by ALFALFA. We then excluded Virgo galaxies, because their neutral hydrogen is known to undergo strong environmental impact (e.g., [44]). We also excluded galaxies whose half-light radii are too small (<1") comparing to the SDSS resolution. These small half-light radii either arise from misidentifications (from stars) or would result in large uncertainties. We are left with a large sample of 1022 SDSS nearby galaxies, which have distances smaller than 254 Mpc. Among these galaxies, 889 have g magnitudes that are <18 and 120 have $g = 18-20$. The cumulative number counts of SDSS galaxies (e.g., [45]) are $\sim 60 \text{ deg}^{-2}$ at $g < 18$ and $\sim 450 \text{ deg}^{-2}$ at $g = 18-20$. Given our 10" search radius, we therefore expect at most 1.3 misidentifications in our 889 $g < 18$ galaxies and additional 1.3 misidentifications in our 120 $g = 18 - 20$ galaxies. These number are sufficiently small and misidentified galaxies should not impact our analyses.

The total stellar masses of galaxies are more directly reflected by the near-infrared observations. We added the data from 2MASS to our samples. 2MASS provides J , H , and K_s -band observations of the entire sky as well as a point source catalog and an extended source catalog. We use the 2MASS All-Sky Extended Source Catalog (XSC) to find the galaxies in the H I samples. To understand the quality of the identification, we first compared the 2MASS and SDSS coordinates of the 2MASS-detected galaxies. We found that

more than 90% of them have offsets between 2MASS and SDSS that are well within their half-light radii. We visually inspected all galaxies with large offsets of $> 2''$ and found small number of cases that are likely misidentifications as well as ongoing mergers. We excluded these galaxies from our samples. Because 2MASS is shallower than SDSS, we are left with 481 reliably identified galaxies in the near-infrared. To match the same aperture of the color for SDSS and 2MASS catalog, 479 galaxies are left in our near-infrared subsamples.

From the ALFALFA-released catalog 1, 2, and 3, we obtained 1796 HI data, out of which 1265 galaxies could be found in the SDSS DR7 database. There are 32 galaxies within this set that are too faint in the optical to have reliable magnitudes and luminosities. Hence we finally used the remaining 1022 galaxies in our analyses. We have also analyzed the 195 galaxies of D08 and G09 from HIPASS [46–49] and have obtained similar results. However, since the definitions of the observational variables, such as those of the rotational velocity, are not entirely consistent between the two data sets, we only report on the results from ALFALFA. These 1022 galaxies can be regarded as a blind HI-selected sample. We deduced from the data six variables, which are R_{50} (half-light radius in units of pc), R_{90} (90%-light radius in units of pc), L_r (luminosity in r band in solar units), M_{HI} (HI mass in solar units), M_d (dynamical mass in solar units), and color ($g-i$).

The variables R_{50} and R_{90} represent the radii in the Petrosian system [45, 50, 51]. In SDSS, the parameters are `petroR50` and `petroR90`, respectively. Because the Petrosian system is based on circular objects, we corrected the radii with the major-to-minor axis ratios, which are the parameters `deVAB_r` or `expAB_r` in SDSS. To do this, we follow the result in [52] and [48, 49]. The authors fitted the corrections from circular to elliptical apertures as functions of major-to-minor axis ratios. We directly adopted their formulas for our corrections. By comparing the likelihoods of the de Vaucouleur and the exponential models, we chose the one with the larger likelihood between `deVAB_r` and `expAB_r`. L_r was derived from the Petrosian system and calculated from the Petrosian magnitude, `petroMag_r`. In order to have the same aperture for 2MASS, we use Petrosian color, which is from `petroMag_g` and `petroMag_i`. The variable M_d is calculated by $(\Delta V)^2 R_{90}/G$, in which ΔV is the rotational velocity from the HI spectra and corrected for the inclination with the major-to-minor axis ratio as what we did for R_{50} and R_{90} . M_{HI} is acquired directly from the ALFALFA database and it is derived from the HI flux. This estimation is based on the assumption that the masses from optical radius are proportional to dark matter halos.

To make sure that the masses from the HI measurement can describe the dynamical mass, we compare them with the masses based on the stellar velocity dispersion, which is $M_{\text{dyn}} = K_v \rho_0^2 R_e / G$, as in [1–3]. Here, ρ_0 is the velocity dispersion, R_e is the effective radius, $K_v = 73.32 / [10.465 + (n - 0.95)^2 + 0.954]$, and n is the Sérsic index. In SDSS, the velocity dispersion can be calculated by $\rho = \rho_{\text{ap}} (8R_{\text{ap}}/R_e)^{0.066}$ [53], where $R_{\text{ap}} = 1''5$, R_e is from the best fitting circularized Sérsic profile, and ρ_{ap} is the SDSS measurement within the $3''$ fibers. The Sérsic data are from NYU Value-Added

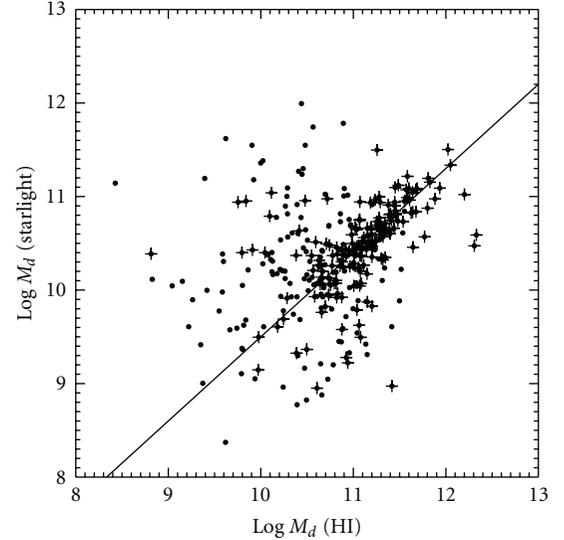


FIGURE 1: Comparison between dynamical masses from HI as D08/G09 and dynamical masses from starlight [1–3]. The dots are the 320 galaxies in ALFALFA/SDSS samples and cross symbols are the 164 galaxies from ALFALFA/SDSS/2MASS subsamples. The solid line indicates that the HI dynamical masses are on-average $\sim 4\times$ larger than the stellar dynamical masses.

Galaxy catalog (VAGC [54]). After matching, we have 320 of our 1022 main galaxy sample and 164 of our 479 2MASS subsample.

Figure 1 is the comparison between the HI and the stellar dynamical masses. There is an apparent sequence, indicating that the two masses trace similar dynamics for most of the galaxies. On the other hand, the HI dynamical masses are on-average $\sim 4\times$ larger than the stellar dynamical masses. This ratio is indicated by the solid line in Figure 1. This is not too surprising, since neutral gas (and R_{90}) can trace the dynamical mass to a larger distance in a dark matter halo. In our subsequent analyses, we adopted HI dynamical masses. This is consistent with the work of D08/G09 and provides a larger sample here.

Because there is high degeneracy between J , H , and K bands, we only chose J band to represent the near-infrared data. Therefore, we acquired R_j (half-light radius in J band in units of pc) and L_j (luminosity in J band in solar units) of the galaxies that overlap in the ALFALFA, SDSS, and 2MASS catalogs to gain a better insight into the stars of the galaxies. More specifically, L_j is derived from the magnitude in J band, which is the parameter `j_m_ext` in the 2MASS database. This magnitude is based on the extrapolation of the fit to the surface brightness profile. And R_j is the integrated half-flux radius of J band, which is the parameter `j_r_eff` in the 2MASS database. We adopted $i-J$ for the color of the near-infrared subsample. The aperture of `petroMag_i` in SDSS is twice of the Petrosian radius, $2\times\text{petroRad}_i$. Thus we interpolate the J -band magnitude in 2MASS at the same aperture as SDSS. In our 481 2MASS subsamples, 479 of them have sufficient data at different apertures to interpolate

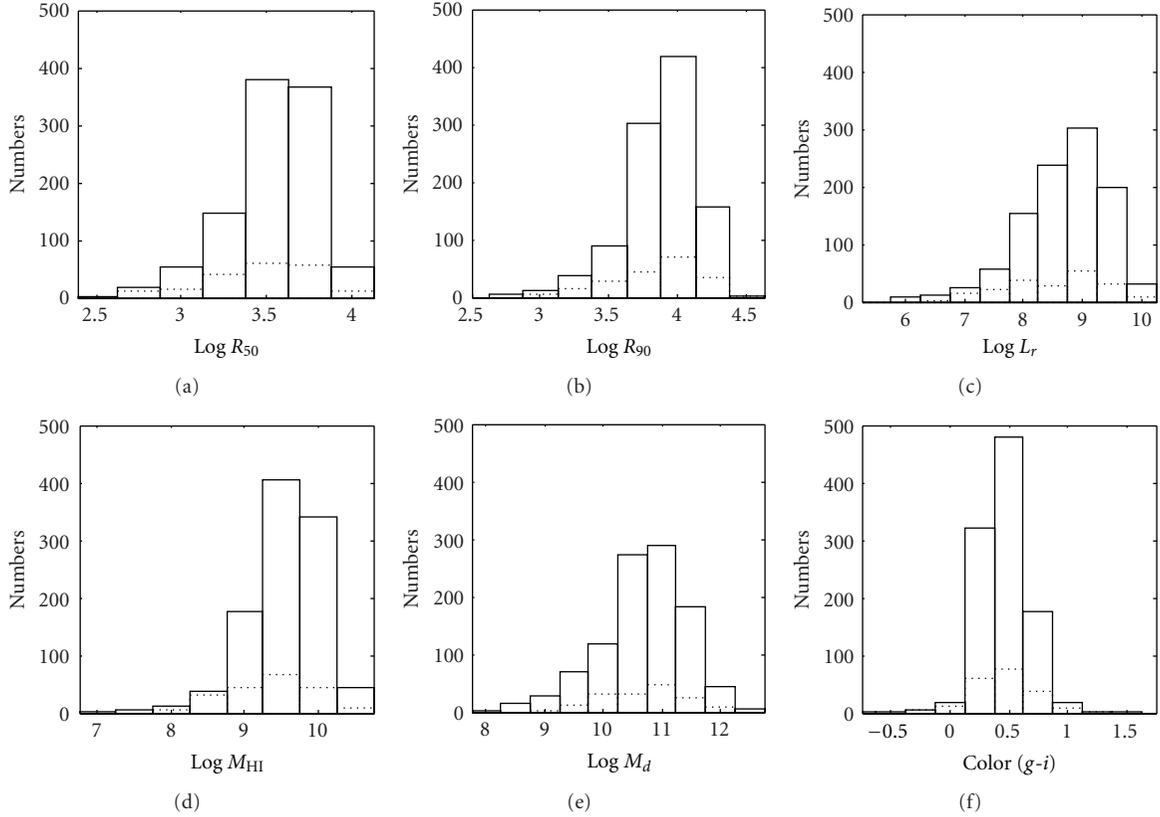


FIGURE 2: Distribution of the variables. The solid histograms are our samples and the dashed histograms are the samples of D08 and G09. We could only find 157 galaxies with rotational velocities for M_d in D08 and G09.

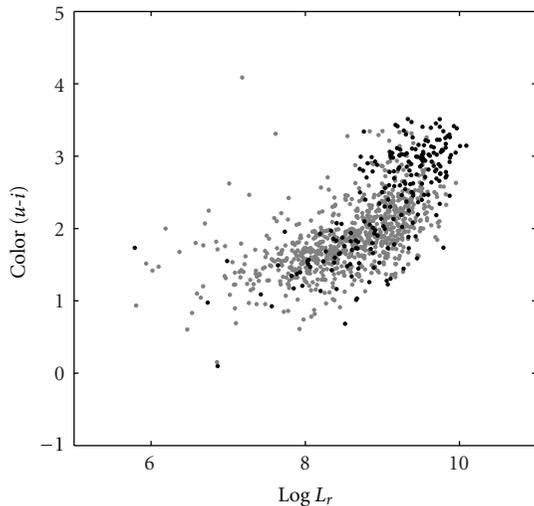


FIGURE 3: The $u-i$ versus L_r diagram. Gray dots are disk galaxies that are better described by exponential models. Black dots are spheroidal galaxies that are better described by de Vaucouleur models.

the 2MASS magnitude. Therefore our final near-infrared subsample contains 479 galaxies.

3. Methods and Results

Our sample of 1022 galaxies is not only larger than that in D08 and G09 but also covers broader ranges of size, luminosity, and mass (Figure 2). Despite that our sample is still dominated by L_* galaxies, the minimum value of L_r in our sample is much smaller than that of D08 and G09. As for the HI and the dynamical masses, although the median values of the two samples are similar, our sample contains a substantial amount of lower mass galaxies. Our sample also covers a broader range of the $g-r$ color. Because of the larger sampling and the wider range of the galactic properties, our sample is in general more representative. However since 2MASS is shallower, our 2MASS subsample of 479 galaxies is relatively speaking less representative than that from SDSS. Even so, our 2MASS subsample is still substantially larger than that of D08 and G09.

It is important to investigate whether our HI-selected sample is biased against early type galaxies, since these galaxies are usually gas poor. To do so, we identify spheroidal and disk galaxies in our sample based on the morphology with a method similar to that in [55]. In the SDSS database, there are de Vaucouleur and exponential models for each galaxy. By comparing the likelihood and the fractions of the two models for our 1022 galaxies, we found that 804 galaxies are disk like and 218 galaxies are spheroidal like.

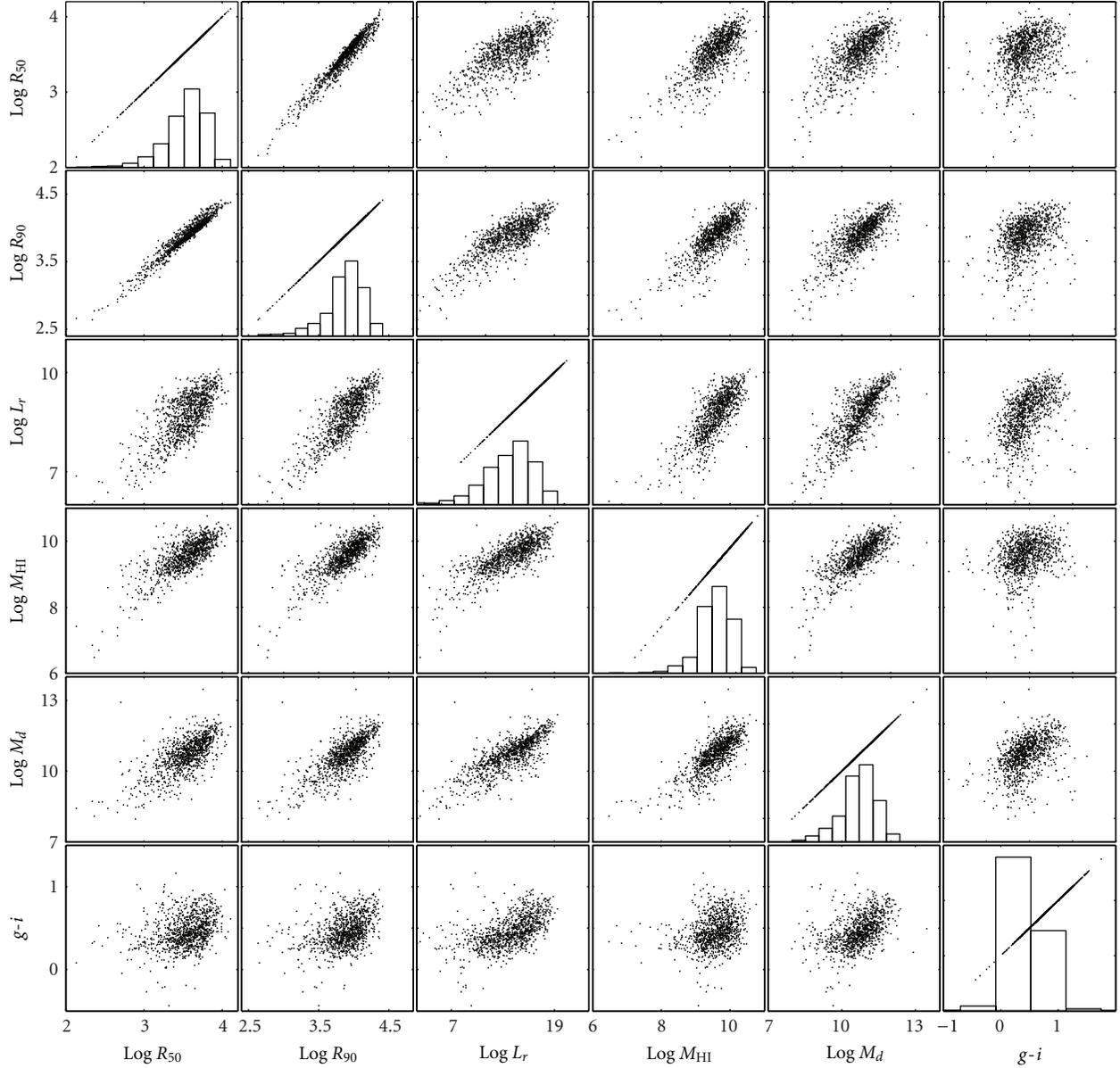


FIGURE 4: Scatter plots showing correlations between six measured variables. All the variables are in solar units and with logarithmic representation. The diagonal line is the histograms, which have vertical scales from 0 to 700.

In Figure 3, we show a color-luminosity diagram for our 1022 galaxies. The spheroidal galaxies are in general more luminous and redder than the disk galaxies. This is consistent with what we expect for elliptical and spiral galaxies. Most importantly, in the color-luminosity space, the spheroidal galaxies are redder than the blue cloud although they do not yet form a complete red sequence. Our sample thus appears to include a fair number of red and elliptical galaxies. Although the bias against extremely gas-poor galaxies can be hardly avoided here, fortunately, we found no major difference between these two types in our subsequent studies. We therefore believe that the omission of extremely gas-poor galaxies should not have caused major systematic bias in our analysis.

Relations among the key parameters can be inferred from the 1022 galaxies in both ALFALFA and SDSS. For instance, it is found that the half-light radius is proportional to the 90%-light radius ($R_{50} \propto R_{90}$; [56]); the r -band luminosity is proportional to the cubic power of the half-light radius ($L_r \propto R_{50}^3$; [57]); the H I mass is proportional to the square of the half-light radius ($M_{\text{HI}} \propto R_{50}^2$; [43, 58]); finally, the dynamical mass is proportional to the r -band luminosity ($M_d \propto L_r$; [25]). We found that all the correlations are evident even after including the near-infrared variables, except the color (Figures 4 and 5).

As a whole, the correlations between color and other variables are much weaker than other correlations. We tested various combinations of colors and found that $g-i$ gives a

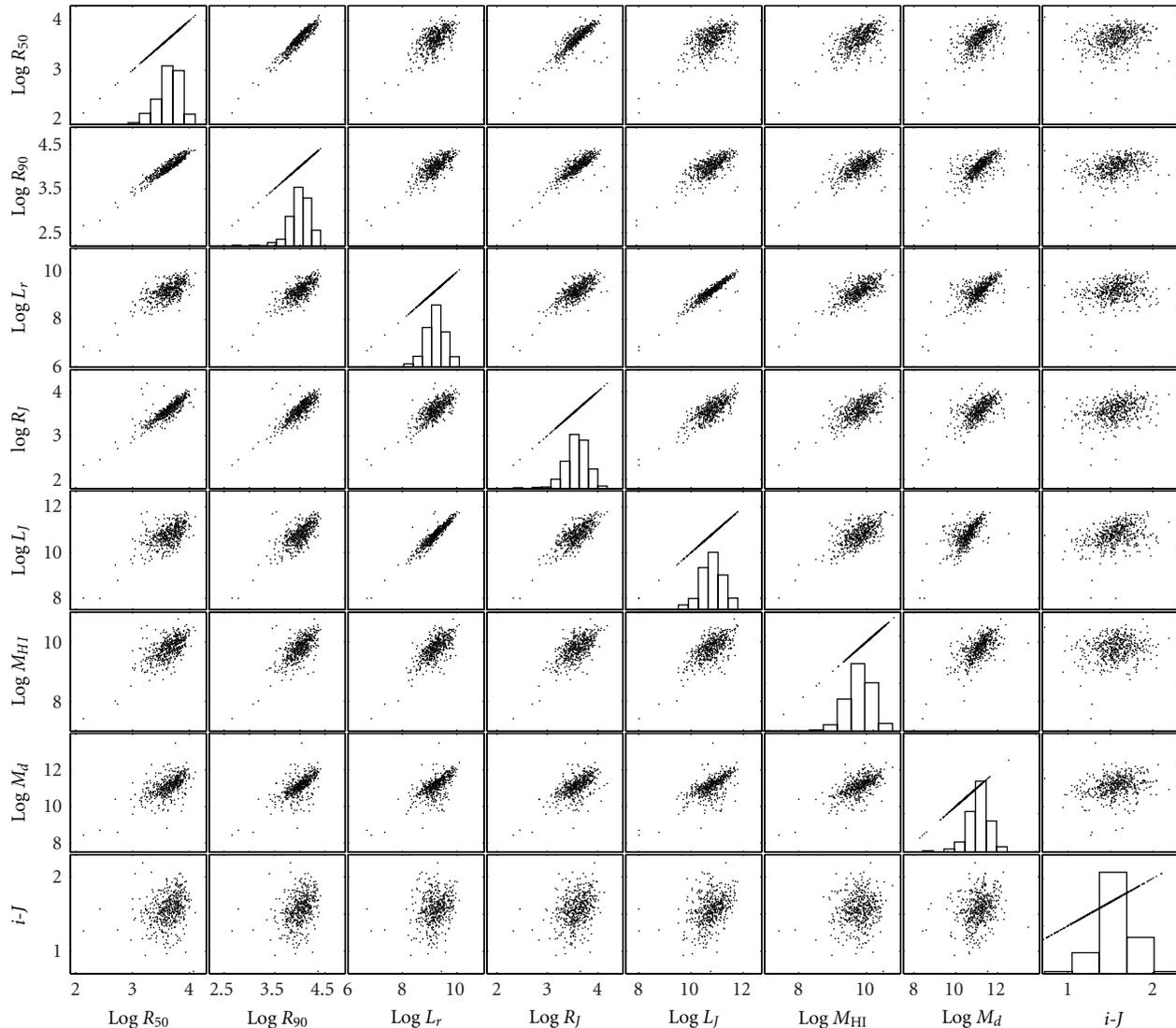


FIGURE 5: Scatter plots showing correlations between eight measured variables, including 2MASS data and reducing to 479 galaxies. All the variables are in solar units and with logarithmic representation. The diagonal line is the histograms, which have vertical scales from 0 to 350. There are small numbers of outliers in many of the plots. They are likely misidentifications or bad photometry, and they do not impact our analyses.

larger PCA correlation coefficient than other colors (e.g., $g-r$, adopted by D08). The reason could be the larger wavelength difference between g and i . Among the three 2MASS bands, the result based on J is somewhat better in the PCA, possibly because of the better signal-to-noise ratio. Hence, we adopted $i-J$ for the color when we analyzed the 479 galaxies in 2MASS. Nevertheless, our analyses show that all the correlation coefficients are smaller than 0.7, indicating that they are not so highly correlated with other parameters. Intuitively, more luminous galaxies tend to be redder because their colors are dominated by older stars. In fact, the color is more complex than any other variable because of the bias introduced by the very luminous young stars.

We conducted PCA to find correlations among the variables. PCA typically produces a series of new variables

called the principal components, namely, PC1, PC2, and so forth. The correlations between these principal components and the original variables then reveal the general correlations between the particular variable and the others. In our case we found that the first principal component, PC1, is highly correlated with the six observational variables. We notice that while the color is less correlated to PC1, possibly because of the recent star formation, it is much more correlated with the second principal component, PC2. In addition to the investigation into the diagram of correlations, the eigenvalues of the correlation matrices of the original variables give quantitative information of the degree of correlations. For the 1022 galaxies based on SDSS, the eigenvalues of PC1 through PC6 are 4.29, 0.92, 0.39, 0.20, 0.18, and 0.02 (Figure 6), respectively, where the maximum possible value

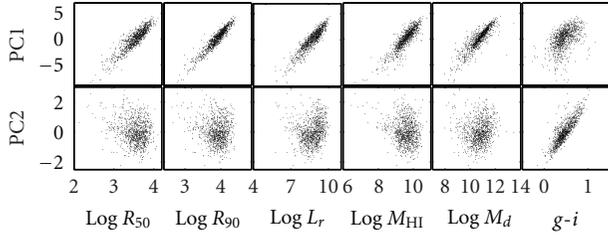


FIGURE 6: The PCA results for 1022 galaxies from ALFALFA and SDSS with colors. Here we only show the strongest ones, PC1 and PC2, because other principal components are not significant by PCA criteria. PC1 is well correlated with all the variables. In the first row, the color is still correlated with the other five variables and PC1. In the second row, the rightmost plot shows that the color is even more strongly correlated with a new principal component, PC2, which is not correlated with other variables. In this plot, $PC1 = 0.44 \log R_{50} + 0.46 \log R_{90} + 0.44 \log L_r + 0.43 \log M_{HI} + 0.42 \log M_d + 0.20(g-i)$ and $PC2 = -0.21 \log R_{50} - 0.13 \log R_{90} + 0.10 \log L_r - 0.22 \log M_{HI} + 0.04 \log M_d + 0.94(g-i)$.

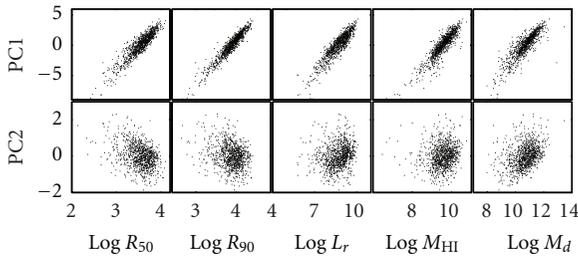


FIGURE 7: The PCA results for 1022 galaxies from ALFALFA and SDSS without colors. Here we only show the strongest ones, PC1 and PC2, because other principal components are not significant by PCA criteria. PC1 is well correlated with all the variables. In this plot, $PC1 = 0.45 \log R_{50} + 0.47 \log R_{90} + 0.45 \log L_r + 0.44 \log M_{HI} + 0.43 \log M_d$ and $PC2 = -0.59 \log R_{50} - 0.43 \log R_{90} + 0.28 \log L_r + 0.18 \log M_{HI} + 0.60 \log M_d$.

is 6. Based on common PCA criteria, eigenvalues larger than 1 are considered significant. We therefore only plot PC1 and PC2. Next we conducted PCA without color, and we found the eigenvalues of PC1 through PC5 to be 4.15, 0.41, 0.22, 0.20, and 0.02 (Figure 7), respectively, where the maximum possible value is 5.

The aforementioned observations confirm the finding of D08. All the observed parameters are tightly correlated with PC1. The eigenvalue of PC1 indicates that it can explain 83% of the variance in the data (when color is not included). Color itself forms a second principal component. This might be explained by the fact that the optical color tends to be strongly affected by recent star formation activities and thus carries extra information that is unrelated to the global formation history of the galaxies. We will come back to the issue of color in Section 4. D08 claimed that the strong dominance of PC1 implies a single physical parameter to govern the structure of galaxies. However, this may be an oversimplified view of galaxies due to limited observational

parameters. To test this, we extend the PCA to the near-infrared.

We included the 2MASS J -band radius and the luminosity to investigate the role of stars, which dominate the baryonic matter in galaxies. For the 479 galaxies detected by 2MASS, we found the eigenvalues of PC1 through PC8 to be 5.40, 1.01, 0.61, 0.45, 0.32, 0.16, 0.03, and 0.02 (Figure 8), respectively, where the maximum possible value is 8. Conducting PCA without color again, we found that the eigenvalues of PC1 through PC7 are 5.35, 0.62, 0.45, 0.31, 0.19, 0.05, and 0.02 (Figure 9), respectively, where the maximum possible value is 7.

The overall trends in the previous near-infrared PCA are similar to those in the optical PCA. When color is not included, PC1 dominates and can explain 76% of the variance in the data. The importance of PC2 slightly increases from 8% in the optical case to 9% in the near-infrared case here. When color is included, it forms another principal component by itself. These again confirm the observations of D08 that only one physical parameter governs the dynamics of galaxies.

A subtle but surprising difference between the optical and near-infrared PCAs is the behavior of color. In the optical PCA in Figure 6, although the $g-i$ color forms a second principal component, it still weakly correlates with other parameters and it is part of the first principal component. On the other hand, in the near-infrared PCA in Figure 8, the $i-J$ color almost does not involve in the first principal component and itself forms a second component that is almost independent of other parameters. A potential issue here is the combination of SDSS i and 2MASS J . It is the reason we choose consistent apertures in SDSS and 2MASS to calculate the color. To test whether other minor differences between the two photometric system hamper the correlation, we replaced the $i-J$ color with the pure 2MASS $J-K$ color, and we found consistent results. In addition, the median errors in g and i for the 1022 SDSS galaxies are 0.015 and 0.018, respectively, and the median errors in i and J for the 479 2MASS galaxies are 0.011 and 0.075, respectively. These translate to typical color errors of 0.023 in $g-i$ and 0.076 in $i-J$. Both values are significantly smaller than the color dynamical ranges shown in Figures 4 and 5, meaning that the distribution of the data is not dominated by measurement errors. We therefore believe that the lack of correlation between the near-infrared color and other parameters is real. We will discuss more on this result in Section 4.

4. Discussion and Conclusion

For the 195 galaxies analyzed in D08 and G09, the correlations among the parameters R_{50} , R_{90} , L_g , M_{HI} , and M_d are obvious. By selecting significant 5 times larger, 1022 overlapping samples from SDSS and ALFALFA, we also performed the PCA and confirmed through the high eigenvalues of the correlation matrices that the correlations are similarly strong. It follows that the radius, luminosity, HI mass, and dynamical mass of those chosen galaxies are tightly correlated. D08 shows that the basic parameters of galaxies

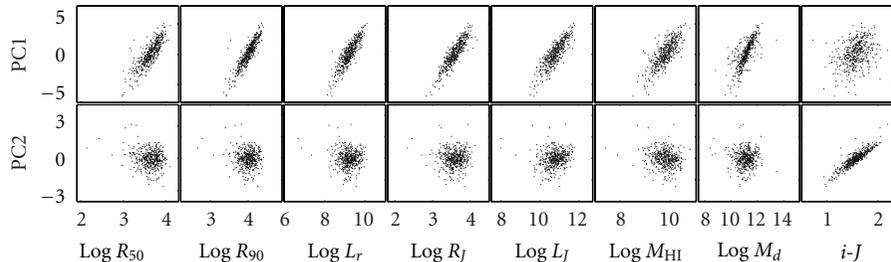


FIGURE 8: The PCA results for 479 galaxies from ALFALFA, SDSS, and 2MASS with colors. Here we only show the strongest ones, PC1 and PC2, because other principal components are not significant by PCA criterions. PC1 is well correlated with all the variables. The color is correlated with other variables and PC1 in the first row as well as strongly correlated with PC2 as in Figure 6. In this plot, $PC1 = 0.37 \log R_{50} + 0.39 \log R_{90} + 0.39 \log L_r + 0.39 \log R_j + 0.38 \log L_j + 0.36 \log M_{HI} + 0.34 \log M_d + 0.12(i-J)$ and $PC2 = -0.22 \log R_{50} - 0.18 \log R_{90} - 0.05 \log L_r + 0.08 \log R_j + 0.18 \log L_j - 0.08 \log M_{HI} - 0.04 \log M_d + 0.93(i-J)$.

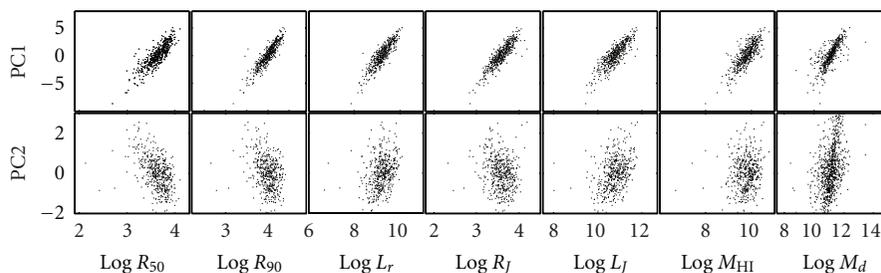


FIGURE 9: The PCA results for 479 galaxies from ALFALFA, SDSS, and 2MASS without colors. Here we only show the strongest ones, PC1 and PC2, because other principal components are not significant by PCA criterions. PC1 is well correlated with all the variables, as in Figure 7. In this plot, $PC1 = 0.38 \log R_{50} + 0.40 \log R_{90} + 0.49 \log L_r + 0.49 \log R_j + 0.38 \log L_j + 0.36 \log M_{HI} + 0.34 \log M_d$ and $PC2 = -0.59 \log R_{50} - 0.39 \log R_{90} + 0.36 \log L_r - 0.22 \log R_j + 0.46 \log L_j + 0.14 \log M_{HI} + 0.29 \log M_d$.

are highly correlated and that there exists only one dominant principal component. From our studies, we know that this is true in both optical and near-infrared bands. Based on the hierarchical galaxy formation assumption, the dark matter halo formation has been well studied through the merger tree process (e.g., [20, 21, 59–64]). D08 believes that this scenario may not be consistent with the simple relation between all basic galactic parameters because the processes of merger would break the original galactic structures. Although our analysis has confirmed D08’s finding, we do not find it compelling or sufficient reason to reject Λ CDM. On the other hand, the dynamical mass-to-light ratio enclosed by the optical radius that strongly anticorrelates with surface brightness has been studied [65–67]. It is not shown in our data. It may be caused by the selection effects from L_* galaxies and insufficient dynamic range in surface brightness with SDSS and 2MASS data. Nevertheless, the strong correlations between parameter in D09 and our gas-selected samples are obvious in both optical and near-infrared data.

However, the color appears to be much less correlated, as T11 mentioned. This indeed complicates the situation and may be explained by a more sophisticated theory that would include, for example, the influence of recent star formation and very luminous young stars. Indeed these studies have already been pursued by many authors (e.g., [55, 59, 68–71]). D08 suggests that the optical color ($g-r$ in their case) consists of two components: the systematic component that

correlates with other parameters (and therefore involves in the first principal component), and the rogue component that is more or less random (and forms the second principal component). It is tempting to assume that the systematic component comes from the established stellar populations and is related to the global formation history of the galaxies, while the random component is related to the ongoing star formation activities and can be short-lived events in the formation history. Surprisingly, our near-infrared analyses suggest a different story. In our near-infrared PCA, the color is even more uncorrelated with other parameters, comparing to the optical color.

One may believe that optical colors like $g-i$ can be more strongly affected by recent star formation, whereas the $i-J$ color may be a better tracer of integrated star formation history because it is less affected by ongoing star formation and dust extinction. If this is correct, the result that $g-i$ is more related to other dynamical properties suggests that the recent star formation history is more controlled by the current dynamical structure galaxies. On the other hand, our result that $i-J$ is unrelated to other dynamical properties suggests that the formation history of the old stellar components is a more chaotic process. In other words, the stronger second component may be an indication of complex formation and evolution history of galaxies. If this is the correct interpretation of our data, then the evidence of complex merger history predicted by the CDM hierarchical

formation model had been hidden in the near-infrared colors and was not revealed by D08. However, we believe that this has to be tested by more observations and simulations, and it is premature to suggest the failure of the CDM model.

To sum up, we believe that the 5 times larger sample size strengthens the claim in D08 that there is one dominant parameter in galactic structure. In addition, our near-infrared analyses also provide additional insight into the color of galaxies. The tight correlations and the uncorrelated color part between the parameters in the optical and near-infrared data provide potentially powerful observational constraints on the hierarchical structure formation theory and any other cosmology models.

Acknowledgments

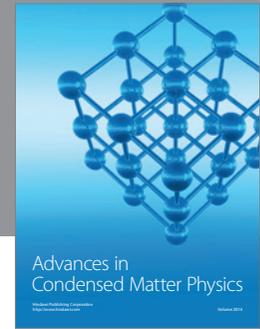
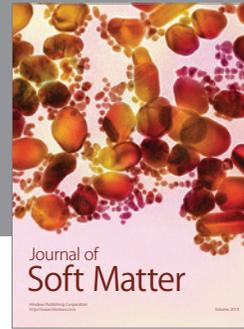
This work is supported by Taiwan National Science Council under Project no. NSC97-2112-M-002-026-MY3 (P. Chen and Y.-Y. Chang), NSC98-2112-M-001-003-MY2 and NSC99-2112-M-001-012-MY3 (W.-H. Wang and Y.-Y. Chang), and by US Department of Energy under Contract no. DE-AC03-76SF00515. The authors also thank the support of the National Center for Theoretical Sciences of Taiwan. They are grateful to L. Lin, H. Hirashita, M. Disney, P. Kroupa, and H.-W. Rix for useful comments and discussion. The Arecibo Observatory is part of the National Astronomy and Ionosphere Center, which is operated by Cornell University under a cooperative agreement with the National Science Foundation. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS website is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium for the participating institutions. The participating institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. This paper 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.

References

- [1] E. N. Taylor, M. Franx, J. Brinchmann, A. Van Der Wel, and P. G. Van Dokkum, "On the masses of galaxies in the local universe," *Astrophysical Journal Letters*, vol. 722, no. 1, pp. 1–19, 2010.
- [2] D. A. Wake, M. Franx, and P. G. van Dokkum, "Which galaxy property is the best indicator of its host dark matter halo properties?" <http://arxiv.org/abs/1201.1913>.
- [3] D. A. Wake, P. G. van Dokkum, and M. Franx, "Revealing velocity dispersion as the best indicator of a galaxy's color, compared to stellar mass, surface mass density or morphology," <http://arxiv.org/abs/1201.4998>.
- [4] G. F. Smoot, C. L. Bennett, A. Kogut et al., "Structure in the COBE differential microwave radiometer first-year maps," *Astrophysical Journal Letters*, vol. 396, no. 1, pp. L1–L5, 1992.
- [5] C. L. Bennett, A. Kogut, G. Hinshaw et al., "Cosmic temperature fluctuations from two years of COBE differential microwave radiometers observations," *Astrophysical Journal Letters*, vol. 436, no. 2, pp. 423–442, 1994.
- [6] C. L. Bennett, M. Halpern, G. Hinshaw et al., "First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: preliminary maps and basic results," *Astrophysical Journal*, vol. 148, no. 1, pp. 1–27, 2003.
- [7] C. L. Bennett, R. S. Hill, G. Hinshaw et al., "First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: foreground emission," *Astrophysical Journal*, vol. 148, no. 1, pp. 97–117, 2003.
- [8] E. Komatsu, J. Dunkley, M. R. Nolte et al., "Five-year wilkinson microwave anisotropy probe observations: cosmological interpretation," *Astrophysical Journal*, vol. 180, no. 2, pp. 330–376, 2009.
- [9] J. Dunkley, D. N. Spergel, E. Komatsu et al., "Five-year wilkinson microwave anisotropy probe (WMAP) observations: bayesian estimation of cosmic microwave background polarization maps," *Astrophysical Journal Letters*, vol. 701, no. 2, pp. 1804–1813, 2009.
- [10] A. G. Riess, A. V. Filippenko, P. Challis et al., "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astronomical Journal*, vol. 116, no. 3, pp. 1009–1038, 1998.
- [11] F. Zwicky, "On the masses of nebulae and of clusters of nebulae," *Astrophysical Journal*, vol. 86, p. 217, 1937.
- [12] B. Moore, S. Ghigna, F. Governato et al., "Dark matter substructure within galactic halos," *Astrophysical Journal Letters*, vol. 524, no. 1, pp. L19–L22, 1999.
- [13] A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, "Where are the missing galactic satellites?" *Astrophysical Journal Letters*, vol. 522, no. 1, pp. 82–92, 1999.
- [14] J. R. Primack, "Cosmology: small-scale issues," *New Journal of Physics*, vol. 11, no. 10, Article ID 105029, 2009.
- [15] A. Kravtsov, "Dark matter substructure and dwarf galactic satellites," *Advances in Astronomy*, vol. 2010, Article ID 281913, 2010.
- [16] J. Sommer-Larsen and A. Dolgov, "Formation of disk galaxies: warm dark matter and the angular momentum problem," *Astrophysical Journal Letters*, vol. 551, no. 2, pp. 608–623, 2001.
- [17] P. J. E. Peebles, "The void phenomenon," *Astrophysical Journal Letters*, vol. 557, no. 2, pp. 495–504, 2001.
- [18] W. J. G. De Blok and S. S. McGaugh, "The dark and visible matter content of low surface brightness disc galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 290, no. 3, pp. 533–552, 1997.

- [19] J. J. Dalcanton and C. J. Hogan, "Halo cores and phase-space densities: observational constraints on dark matter physics and structure formation," *Astrophysical Journal Letters*, vol. 561, no. 1, pp. 35–45, 2001.
- [20] S. D. M. White and C. S. Frenk, "Galaxy formation through hierarchical clustering," *Astrophysical Journal Letters*, vol. 379, no. 1, pp. 52–79, 1991.
- [21] S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk, "Hierarchical galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 319, no. 1, pp. 168–204, 2000.
- [22] C. M. Baugh, "A primer on hierarchical galaxy formation: the semi-analytical approach," *Reports on Progress in Physics*, vol. 69, no. 12, article R02, pp. 3101–3156, 2006.
- [23] S. M. Faber and R. E. Jackson, "Velocity dispersions and mass-to-light ratios for elliptical galaxies," *Astrophysical Journal*, vol. 204, pp. 668–683, 1976.
- [24] R. B. Tully and J. R. Fisher, "A new method of determining distances to galaxies," *Astronomy and Astrophysics*, vol. 54, no. 3, pp. 661–673, 1977.
- [25] G. Gavazzi, D. Pierini, and A. Boselli, "The phenomenology of disk galaxies," *Astronomy and Astrophysics*, vol. 312, no. 2, pp. 397–408, 1996.
- [26] M. R. Blanton, D. W. Hogg, N. A. Bahcall et al., "The broadband optical properties of galaxies with redshifts $0.02 < Z < 0.22$," *Astrophysical Journal Letters*, vol. 594, no. 1 I, pp. 186–207, 2003.
- [27] M. Bernardi, R. K. Sheth, J. Annis et al., "Early-type galaxies in the sloan digital sky survey. II. Correlations between observables," *Astronomical Journal*, vol. 125, no. 4, pp. 1849–1865, 2003.
- [28] G. Kauffmann, T. M. Heckman, S. D. M. White et al., "Stellar masses and star formation histories for 105 galaxies from the Sloan Digital Sky Survey," *Monthly Notices of the Royal Astronomical Society*, vol. 341, no. 1, pp. 33–53, 2003.
- [29] G. Kauffmann, S. D. M. White, T. M. Heckman et al., "The environmental dependence of the relations between stellar mass, structure, star formation and nuclear activity in galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 353, no. 3, pp. 713–731, 2004.
- [30] M. R. Blanton, M. Geha, and A. A. West, "Testing cold dark matter with the low-mass tully-fisher relation," *Astrophysical Journal Letters*, vol. 682, no. 2, pp. 861–873, 2008.
- [31] M. J. Disney, J. D. Romano, D. A. Garcia-Appadoo, A. A. West, J. J. Dalcanton, and L. Cortese, "Galaxies appear simpler than expected," *Nature*, vol. 455, no. 7216, pp. 1082–1084, 2008.
- [32] D. A. Garcia-Appadoo, A. A. West, J. J. Dalcanton, L. Cortese, and M. J. Disney, "Correlations among the properties of galaxies found in a blind H I survey, which also have SDSS optical data," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 1, pp. 340–356, 2009.
- [33] L. Staveley-Smith, W. E. Wilson, T. S. Bird et al., "The Parkes 21 cm multibeam receiver," *Publications of the Astronomical Society of Australia*, vol. 13, no. 3, pp. 243–248, 1996.
- [34] R. Giovanelli, M. P. Haynes, B. R. Kent et al., "The arcibo legacy fast alfa survey. III. H I source catalog of the northern virgo cluster region," *Astronomical Journal*, vol. 133, no. 6, pp. 2569–2583, 2007.
- [35] A. Saintonge, R. Giovanelli, M. P. Haynes et al., "The arcibo legacy fast alfa survey. V. the H I source catalog of the anti-virgo region at $\delta = +27^\circ$," *Astronomical Journal*, vol. 135, no. 2, pp. 588–604, 2008.
- [36] B. R. Kent, R. Giovanelli, M. P. Haynes et al., "The Arcibo legacy fast alfa survey. VI. Second HI source catalog of the Virgo cluster region," *Astronomical Journal*, vol. 136, no. 2, pp. 713–724, 2008.
- [37] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros et al., "The seventh data release of the sloan digital sky survey," *Astrophysical Journal*, vol. 182, no. 2, pp. 543–558, 2009.
- [38] M. C. Toribio, J. M. Solanes, R. Giovanelli, M. P. Haynes, and K. L. Masters, "H I content and optical properties of field galaxies from the ALFALFA survey. I. Selection of a control sample," *Astrophysical Journal Letters*, vol. 732, no. 2, article 92, 2011.
- [39] M. C. Toribio, J. M. Solanes, R. Giovanelli, M. P. Haynes, and A. M. Martin, "H I content and optical properties of field galaxies from the ALFALFA survey. II. Multivariate analysis of a galaxy sample in low-density environments," *Astrophysical Journal Letters*, vol. 732, no. 2, article 93, 2011.
- [40] I. T. Jolliffe, *Principal Component Analysis*, vol. 1, Springer, 1986.
- [41] M. F. Skrutskie, R. M. Cutri, R. Stiening et al., "The two Micron All Sky Survey (2MASS)," *Astronomical Journal*, vol. 131, no. 2, pp. 1163–1183, 2006.
- [42] J. Brinchmann, S. Charlot, S. D. M. White et al., "The physical properties of star-forming galaxies in the low-redshift Universe," *Monthly Notices of the Royal Astronomical Society*, vol. 351, no. 4, pp. 1151–1179, 2004.
- [43] J. L. Rosenberg, S. E. Schneider, and J. Posson-Brown, "Gas and stars in an H I-selected galaxy sample," *Astronomical Journal*, vol. 129, no. 3, pp. 1311–1330, 2005.
- [44] A. Chung, J. H. van Gorkom, J. D. Kenney, H. Crowl, and B. Vollmer, "VLA imaging of virgo spirals in atomic gas (VIVA). I. The atlas and the H I properties," *Astronomical Journal*, vol. 138, no. 6, pp. 1741–816, 2009.
- [45] N. Yasuda, M. Fukugita, V. K. Narayanan et al., "Galaxy number counts from the Sloan Digital Sky Survey commissioning data," *Astronomical Journal*, vol. 122, no. 3, pp. 1104–1124, 2001.
- [46] M. J. Meyer, M. A. Zwaan, R. L. Webster et al., "The HIPASS catalogue—I. Data presentation," *Monthly Notices of the Royal Astronomical Society*, vol. 350, no. 4, pp. 1195–1209, 2004.
- [47] K. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros et al., "The second data release of the sloan digital sky survey," *Astronomical Journal*, vol. 128, no. 1, pp. 502–512, 2004.
- [48] A. A. West, D. A. Garcia-Appadoo, J. J. Dalcanton et al., "H I-Selected Galaxies in the Sloan Digital Sky Survey. II. The colors of Gas-Rich galaxies," *Astronomical Journal*, vol. 138, no. 3, pp. 796–807, 2009.
- [49] A. A. West, D. A. Garcia-Appadoo, J. J. Dalcanton et al., "H I-selected galaxies in the sloan digital sky survey. I. Optical data," *Astronomical Journal*, vol. 139, no. 2, pp. 315–328, 2010.
- [50] V. Petrosian, "Surface brightness and evolution of galaxies," *Astrophysical Journal*, vol. 209, pp. L1–L5, 1976.
- [51] M. R. Blanton, J. Dalcanton, D. Eisenstein et al., "The luminosity function of galaxies in SDSS commissioning data," *Astronomical Journal*, vol. 121, no. 5, pp. 2358–2380, 2001.
- [52] A. A. West, *H I selected galaxies in the sloan digital sky survey*, Ph.D. thesis, University of Washington, 2005.
- [53] M. Cappellari, R. Bacon, M. Bureau et al., "The SAURON project—IV. The mass-to-light ratio, the virial mass estimator and the Fundamental Plane of elliptical and lenticular galaxies," *Monthly Notices of the Royal Astronomical Society*, vol. 366, no. 4, pp. 1126–1150, 2006.
- [54] M. R. Blanton, D. J. Schlegel, M. A. Strauss et al., "New York University Value-Added Galaxy Catalog: a galaxy catalog based

- on new public surveys,” *Astronomical Journal*, vol. 129, no. 6, pp. 2562–2578, 2005.
- [55] I. Strateva, Ž. Ivezić, G. R. Knapp et al., “Color separation of galaxy types in the Sloan Digital Sky Survey imaging data,” *Astronomical Journal*, vol. 122, no. 4, pp. 1861–1874, 2001.
- [56] P. C. van der Kruit, “The three-dimensional distribution of light and mass in disks of spiral galaxies,” *Astronomy and Astrophysics*, vol. 192, no. 1-2, pp. 117–127, 1988.
- [57] M. R. Blanton, D. W. Hogg, N. A. Bahcall et al., “The galaxy luminosity function and luminosity density at redshift $z = 0.1$,” *Astrophysical Journal Letters*, vol. 592, no. 2 I, pp. 819–838, 2003.
- [58] M. P. Haynes and R. Giovanelli, “Neutral hydrogen in isolated galaxies. IV—results for the Arecibo ample,” *Astronomical Journal*, vol. 89, pp. 758–800, 1984.
- [59] J. Kormendy, “Integrated colors of bright galaxies in the u, b, V system,” *Astrophysical Journal*, vol. 218, p. 333, 1977.
- [60] M. S. Longair, *Galaxy Formation*, vol. 1, Springer, 2000.
- [61] 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.
- [62] 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.
- [63] J. Kormendy, D. B. Fisher, M. E. Cornell, and R. Bender, “Structure and formation of elliptical and spheroidal galaxies,” *Astrophysical Journal*, vol. 182, no. 1, pp. 216–309, 2009.
- [64] F. Fontanot, G. De Lucia, P. Monaco, R. S. Somerville, and P. Santini, “The many manifestations of downsizing: hierarchical galaxy formation models confront observations,” *Monthly Notices of the Royal Astronomical Society*, vol. 397, no. 4, pp. 1776–1790, 2009.
- [65] M. A. Zwaan, J. M. van der Hulst, W. J. G. de Blok, and S. S. McGaugh, “The Tully-Fisher relation for low surface brightness galaxies: implications for galaxy evolution,” *Monthly Notices of the Royal Astronomical Society*, vol. 273, pp. L35–L38, 1998.
- [66] S. S. McGaugh and W. J. G. De Blok, “Testing the hypothesis of modified dynamics with low surface brightness galaxies and other evidence,” *Astrophysical Journal Letters*, vol. 499, no. 1, pp. 66–81, 1998.
- [67] R. Reyes, R. Mandelbaum, J. E. Gunn, J. Pizagno, and C. N. Lackner, “Calibrated Tully-Fisher relations for improved estimates of disk rotation velocities,” <http://arxiv.org/abs/1106.1650>.
- [68] G. de Vaucouleurs, “Integrated colors of bright galaxies in the u, b, V system,” *Astrophysical Journal Supplement Series*, vol. 5, p. 233, 1961.
- [69] R. B. Tully, J. R. Mould, and M. Aaronson, “A colormagnitude relation for spiral galaxies,” *Astrophysical Journal*, vol. 257, pp. 527–537, 1982.
- [70] D. W. Hogg, M. R. Blanton, J. Brinchmann et al., “The dependence on environment of the color-magnitude relation of galaxies,” *Astrophysical Journal*, vol. 601, no. 1, pp. L29–L32, 2004.
- [71] Y. Chen, J. D. Lowenthal, and M. S. Yun, “Color-magnitude relation and morphology of low-redshift ulirgs in sloan digital sky survey,” *Astrophysical Journal Letters*, vol. 712, no. 2, pp. 1385–1402, 2010.

The Hindawi logo consists of two interlocking loops, one blue and one green, forming a stylized infinity symbol.

Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

