

Research Article

Distributions Escaping to Infinity and the Limiting Power of the Cliff-Ord Test for Autocorrelation

Kairat T. Mynbaev

*International School of Economics, Kazakh-British Technical University, Tole bi 59,
 Almaty 050000, Kazakhstan*

Correspondence should be addressed to Kairat T. Mynbaev, kairat.mynbayev@yahoo.com

Received 18 September 2012; Accepted 24 October 2012

Academic Editors: J. Hu, A. Hutt, and J. Villarroel

Copyright © 2012 Kairat T. Mynbaev. 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.

We consider a family of proper random variables which converges to an improper random variable. The limit in distribution is found and applied to obtain a closed-form expression for the limiting power of the Cliff-Ord test for autocorrelation. The applications include the theory of characteristic functions of proper random variables, the theory of almost periodic functions, and the test for spatial correlation in a linear regression model.

1. Introduction

Improper random variables do not satisfy the condition $P(X \in R) = 1$; that is, they may take values outside the real line R . They are not used much by themselves, but there are situations when they arise as limits of proper random variables. In such cases, we say that a *distribution escapes to infinity*. The main problem considered in this paper is illustrated in the following in a simplified situation.

Throughout the paper we denote by χ_A the indicator of a set A . Let $g = (1/2)\chi_{[-1,1]}$ be a uniform distribution on the segment $[-1, 1]$. Consider the family of densities $g_\lambda(t) = \lambda g(\lambda t)$, where $\lambda > 0$. The total mass is constant: $\int_R g_\lambda(t) dt = 1$ for any $\lambda > 0$. In Bayesian estimation, improper priors are obtained by letting $\lambda \rightarrow 0$. In this case, there are two effects at work: the support $\text{supp } g_\lambda = [-1/\lambda, 1/\lambda]$ stretches out indefinitely and the height of the density $\max g_\lambda = \lambda/2$ goes to zero. One might be tempted to think of the limit G of $\{g_\lambda\}$, as $\lambda \rightarrow 0$, as an infinitesimally thin layer smeared over the whole real line. This notion would be wrong because

$$\left| \int_R g_\lambda(t) \varphi(t) dt \right| \leq \frac{\lambda}{2} \int_{\text{supp } \varphi} |\varphi(t)| dt \rightarrow 0, \quad \lambda \rightarrow 0, \quad (1.1)$$

for any $\varphi \in C_0(R)$ (the set of continuous functions on R with bounded support $\text{supp } \varphi$). Thus, if the limit in distribution G of $\{g_\lambda\}$ exists, it vanishes on all elements of $C_0(R)$. By the definition of the support of a distribution [1, Chapter 1, Section 13], the support of G does not contain R . So instead of spreading the mass over R , it is more correct to say that the mass escapes to infinity.

In this paper we provide a rigorous framework for treating more complex situations. To illustrate the arising complexities, let us look at the standard normal variable X with the density $g(x) = (2\pi)^{-1/2} \exp(-x^2/2)$. Let X_λ have the density $g_\lambda(x) = \lambda g(\lambda x) = \lambda(2\pi)^{-1/2} e^{-(\lambda x)^2/2}$. Its characteristic function is

$$Ee^{itX_\lambda} = \int_R e^{i(t/\lambda)x} g(x) dx = Ee^{i(t/\lambda)X} = e^{-(t/\lambda)^2/2}. \quad (1.2)$$

The moment generating function of X_λ is $Ee^{tX_\lambda} = e^{(t/\lambda)^2/2}$. None of these expressions is good for characterizing the limit as $\lambda \rightarrow 0$. Further, let φ denote an arbitrary continuous and bounded function on R . In the expression

$$\int_R g_\lambda(t) \varphi(t) dt = \lambda \int_R g(\lambda t) \varphi(t) dt, \quad (1.3)$$

the height of the density goes to zero, so the integrand converges to zero everywhere. However, the graph of the density stretches out from the origin. Therefore, the best majorant for the integrand is $|g(0)\varphi(t)|$ which is generally not integrable. Thus, the dominated convergence theorem cannot be used to obtain convergence in distribution. While the case of *collapsing density* ($\lambda \rightarrow \infty$) is easily handled by the existing theory, the case of *stretching out* ($\lambda \rightarrow 0$) requires new tools which we develop here.

Problem 1. Describe the limit in distribution when the stretching-out is applied to a density along all or some variables.

This problem is solved in Theorem 2.2 in Section 2 (case of all variables) and Theorem 2.6 (case of some variables). In their simplest form, those results reveal the main ideas. Let g be any summable even function on R (it may change sign and not integrate to unity). Let $M\varphi$ denote the *generalized mean* of φ over R :

$$M\varphi = \lim_{r \rightarrow \infty} \frac{1}{2r} \int_{-r}^r \varphi(t) dt. \quad (1.4)$$

Then for a continuous and bounded function φ on R such that $M\varphi$ exists, one has

$$\int_R \lambda g(\lambda t) \varphi(t) dt \longrightarrow M\varphi \int_R g(t) dt, \quad \lambda \longrightarrow 0. \quad (1.5)$$

Now suppose that the function g depends on two variables and the stretching-out is applied with respect to the first of them. Let φ be continuous and bounded on R^2 . Denote $M_1\varphi$, the result of application of (1.4) to φ with respect to the first argument,

$$(M_1\varphi)(t_2) = \lim_{r \rightarrow \infty} \frac{1}{2r} \int_{-r}^r \varphi(t_1, t_2) dt_1, \quad (1.6)$$

and let $g_1(t_2) = \int_R g(t_1, t_2) dt_1$. In cases when g is not a density, we call this function a *marginal "density."* Then, (1.5) can be used to prove

$$\begin{aligned} \int_{R^2} \lambda g(\lambda t_1, t_2) \varphi(t) dt &= \int_R \left[\int_R \lambda g(\lambda t_1, t_2) \varphi(t_1, t_2) dt_1 \right] dt_2 \\ &\rightarrow \int_R (M_1\varphi)(t_2) g_1(t_2) dt_2, \quad \lambda \rightarrow 0, \end{aligned} \quad (1.7)$$

if g is integrable and spherically symmetric. As we show, the right sides of (1.5) and (1.7) determine distributions supported at infinity.

Theorem 2.2 is followed by two applications. One generalizes results from [2] on the link between jumps of a distribution function of a (proper) random variable and its characteristic function. Another application is to the fundamental theorem of H. Bohr on the Parseval identity for almost-periodic functions; see [3]. Theorem 2.6 is applied to the limiting power of the Cliff-Ord test for autocorrelation. To formulate the related problems, we need some notation.

Consider a linear regression model

$$y = X\beta + u, \quad Eu = 0, \quad \text{var}(u) = \sigma^2 \Sigma(\rho), \quad (1.8)$$

where X is an $n \times k$ matrix of rank $k < n$; a vector $\beta \in R^k$ and a number $\sigma^2 > 0$ are unknown parameters. The matrix $\Sigma(\rho)$ is assumed to be a nonnegative function of the parameter $\rho \in [0, a]$, with some $a > 0$. ρ in applications characterizes the degree of autocorrelation. Testing for autocorrelation takes the form $H_0 : \rho = 0$ versus $H_a : \rho > 0$. The case $\rho \rightarrow a$ is of special interest for determining the limiting power of tests.

Assuming that $\Sigma(\rho)$ is positive definite for $\rho \in [0, a)$, denote $\varepsilon = \Sigma^{-1/2}(\rho)u$ and let $g(\varepsilon)$ be the density of ε . The density of y is then given by [4, Equation (D.2)]

$$f_\rho(y) = \det \Sigma^{-1/2}(\rho) g\left(\Sigma^{-1/2}(\rho)(y - X\beta)\right). \quad (1.9)$$

Problem 2. Assuming that

$$\text{rank}\left(\Sigma^{-1}(a)\right) = n - 1, \quad (1.10)$$

describe the limit in distribution of f_ρ .

This problem, as an intermediate step in the proof, was considered by Martellosio [4], and his answer is reproduced as follows. Let $N(A)$ denote the null space of a matrix A and denote $A(a) = \Sigma^{-1}(a)$. Let $X\beta + N(A(a))$ be the translation by $X\beta$ of the null space of $A(a)$. Reference [4] states that in case (1.10)

$$\text{as } \rho \longrightarrow a, \quad f_\rho \text{ tends to a degenerate density supported on } X\beta + N(A(a)), \quad (1.11)$$

see page 182. Equation (1.11) arises from the confusion between the stretching-out and collapsing. On page 159, rows 10 and 11, he remarks that (1.11) can be extended to the case $\text{rank}(\Sigma^{-1}(a)) < n - 1$. He does not specify the mode of convergence, but, as we argue in Section 2, the convergence in distribution is the right one for the problem of the behavior of the limiting power of tests for autocorrelation. Unfortunately, (1.11) does not hold under the convergence in distribution. We prove this and solve Problem 2 in Theorem 2.7 for the general case $0 < \text{rank}(\Sigma^{-1}(a)) < n$.

Let Φ be a critical region for rejecting H_0 in favor of H_a . Denote

$$\beta_\rho(\Phi) = \int_{\Phi} f_\rho(y) dy, \quad (1.12)$$

the probability content of Φ under density (1.9) and define the *limiting power* $\beta(\Phi)$ as the limit of this probability content: $\beta(\Phi) = \lim_{\rho \rightarrow a} \beta_\rho(\Phi)$. Specifically, we consider critical regions that arise from the Cliff-Ord test, to be described now. With the regressor matrix X from (1.8) denote $L_X = X(X'X)^{-1}X'$, $M_X = I - L_X$. Under the spatial autocorrelation assumption, the regression disturbances follow

$$u = \rho Wu + \varepsilon, \quad E\varepsilon = 0, \quad \text{var}(\varepsilon) = \sigma_\varepsilon^2 I, \quad (1.13)$$

where ε is a new disturbance and W is some known $n \times n$ matrix. The scalar ρ , which is unknown, determines the degree of correlation among the components of u . For testing the null $H_0 : \rho = 0$ against the alternative $H_a : \rho > 0$, Cliff and Ord [5] proposed a test that rejects the null if $u'M_X W M_X u / u'M_X u > c$. Denote

$$\Phi = \left\{ u \in R^n : \frac{u'M_X W M_X u}{u'M_X u} > c \right\}, \quad \Phi_- = \left\{ u \in R^n : \frac{u'M_X W M_X u}{u'M_X u} < c \right\}. \quad (1.14)$$

Problem 3. Obtain a closed-form expression for $\beta(\Phi)$.

Theorem 2.7 is applied in Theorem 2.9 to solve this problem in case (1.10).

Problem 4. Describe the cases when *the limiting power disappears*, that is, when $\beta(\Phi) = 0$.

This problem has been the main motivation for this paper. Spatial models in general are peculiar in many respects, and the possibility of the limiting power to disappear is one of those peculiarities that has been attracting researchers' attention lately. Krämer [6] was the first to suggest that the limiting power of tests for spatial autocorrelation may vanish, for some combinations of the regressor matrix and the spatial matrix. Unfortunately, the terms in which Krämer expressed his results do not adequately reflect all the possibilities, and his

proof contains an incorrect argument; see [7, Footnote 5]. Martellosio [4] was the first to suggest that the answer is better described in terms of the geometrical relationship between the eigenvectors of $\Sigma^{-1}(a)$ and the critical region Φ . However, both of his main result [4, Theorem 1] and its proof contain errors (see Remark 2.12 in Section 2). Our answer to Problem 4 is given in Theorem 2.11 and corrects [4, Theorem 1] in that where he excludes the extreme values 0 and 1 for the limiting power, we give examples showing that the extreme values are possible. In the light of our result, several statements in [4, 7] have to be reconsidered. The analysis for the Cliff-Ord test is very involved, and it is not feasible to indicate corrections for all Martellosio's main results that depend on the wrong intermediate statement or his Theorem 2.9. The complexity of our analysis necessitates a more detailed notation than that used by Martellosio. In particular, some of his verbal definitions and statements are cast in a more formal way. As a result, our citations are not word for word.

Despite the fact that the convergence in distribution is the right one for Problem 4, it would be interesting to know what kind of convergence produces the limit indicated by Martellosio, as is stated next.

Problem 5. Design the mode of convergence that leads to (1.11).

To this end, we introduce a new convergence concept (which, given its purpose, could be called a *retrofit convergence*), which may not look intuitively appealing but allows us to prove (1.11) in Theorem 2.13. Under this alternative convergence, there is no analog of (1.12). Therefore, we did not consider Problems 3 and 4 for this convergence.

The previous plan will be implemented under conditions much more restrictive than suggested by Martellosio. All main results that are stated in Sections 2 and 3 contains all proofs.

2. Main Statements

In the multidimensional version of the generalized mean (1.4) instead of averaging over segments $[-r, r]$, we have to average over balls. The shape of those balls depends on the norm of R^n . Let $\|\cdot\|_a$ be an arbitrary norm in R^n . The balls are defined by $B_{a,n}(x, r) = \{y \in R^n : \|x - y\|_a \leq r\}$, where the indication of the space dimension will be important when dealing with more than one space. As an example, one can think of the l_p -norm defined by

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max_{1 \leq i \leq n} |x_i|, & \text{if } p = \infty. \end{cases} \quad (2.1)$$

In case of the Euclidean norm $\|\cdot\|_2$, we obtain usual balls; in cases $p = 1$ and $p = \infty$, the balls $B_{p,n}(x, r)$ are cubes. Another useful example is $\|x\|_A = (x'Ax)^{1/2}$, where A is a symmetric, positive definite matrix. We say that a function g on R^n is $\|\cdot\|_a$ -spherically symmetric if $g(x) = p(\|x\|_a)$ with some function p defined on the half-axis $[0, \infty)$. Conditions involving spherical symmetry in the following are similar to Conditions (2.1) and (2.2) from [8].

Let σ denote an element of the unit sphere $\{\sigma \in R^n : \|\sigma\|_a = 1\}$ and let $(\sigma, \rho(\sigma))$ be the representation of a point $x \in R^n$ in the polar system of coordinates such that $\sigma = x/\|x\|_a$ and $\rho(\sigma) = \|x\|_a$ for $x \neq 0$. Then, the Lebesgue measure of $B_{a,n}(x, r)$ is

$$\begin{aligned} \text{mes } B_{a,n}(x, r) &= \text{mes } B_{a,n}(0, r) = \int_{\|\sigma\|_a=1} \int_0^{r(\sigma)} \rho^{n-1} d\rho d\sigma \\ &= r^n \int_{\|\sigma\|_a=1} \int_0^{r(\sigma)/r} \rho^{n-1} d\rho d\sigma = v_a r^n, \end{aligned} \quad (2.2)$$

where $v_a = \text{mes } B_{a,n}(0, 1)$ is the volume of the unit ball. The norm $\|\cdot\|_a$ gives rise to averages

$$m(\varphi, B_{a,n}(x, r)) = \frac{1}{\text{mes } B_{a,n}(x, r)} \int_{B_{a,n}(x, r)} \varphi(t) dt, \quad (2.3)$$

and to the generalized mean

$$M_{a,n}\varphi = \lim_{r \rightarrow \infty} m(\varphi, B_{a,n}(0, r)). \quad (2.4)$$

In the one-dimensional case all balls are segments, and we write simply $M\varphi$ instead of $M_{a,1}\varphi$.

One of the basic properties of generalized means is that they do not depend on the behavior of φ in any fixed ball $B_{a,n}(0, r)$:

$$\begin{aligned} M_{a,n}\varphi &= \lim_{r \rightarrow \infty} \frac{1}{\text{mes } B_{a,n}(0, r)} \left(\int_{B_{a,n}(0, r_0)} \varphi(t) dt + \int_{B_{a,n}(0, r) \setminus B_{a,n}(0, r_0)} \varphi(t) dt \right) \\ &= \lim_{r \rightarrow \infty} \frac{1}{\text{mes } B_{a,n}(0, r)} \int_{B_{a,n}(0, r) \setminus B_{a,n}(0, r_0)} \varphi(t) dt. \end{aligned} \quad (2.5)$$

Other useful properties are (3.16), (3.37), and (3.42).

$CL(R^n)$ denotes the set of continuous bounded functions on R^n that satisfy the Lipschitz condition

$$|\varphi(x) - \varphi(y)| \leq c_1 \|x - y\|_2, \quad x, y \in R^n; \quad (2.6)$$

By [9, Theorem 3.6.1], convergence in distribution of random elements $X_n \xrightarrow{d} X$ is equivalent to the convergence of expected values

$$E\varphi(X_n) \longrightarrow E\varphi(X), \quad \forall \varphi \in CL(R^n). \quad (2.7)$$

[10, Theorem 2.1] asserts that here $CL(R^n)$ can be replaced by the set of bounded uniformly continuous functions. $L_p(R^n)$ is the space of p -summable functions on R^n provided with the norm $\|g\|_{L_p} = (\int_{R^n} |g(t)|^p dt)^{1/p}$, $1 \leq p < \infty$.

For our applications, in the multidimensional version of (1.5), we need to allow φ to depend on the parameter λ , as in

$$\lim_{\lambda \rightarrow 0} \lambda^d \int_{\mathbb{R}^d} \varphi_\lambda(t) g(\lambda t) dt = M_{a,d} \varphi_0 \int_{\mathbb{R}^d} g(t) dt. \quad (2.8)$$

Here $\varphi \in CL(\mathbb{R}^n)$, $\varphi_\lambda(t) = \varphi(F(\lambda)t + G(\lambda))$. $F(\lambda)$, $\lambda \geq 0$, are $n \times d$ matrices and $G(\lambda)$, $\lambda \geq 0$, are $n \times 1$ vectors such that $F(\lambda)$ and $G(\lambda)$ tend to $F(0)$ and $G(0)$, respectively, sufficiently quickly, as stipulated in the next assumption.

Condition 1. (a) $g \in L_1(\mathbb{R}^d)$ is $\|\cdot\|_a$ -spherically symmetric, $g(x) = p(\|x\|_a)$.

(b) $\|F(\lambda) - F(0)\| = o(\lambda)$, $\|G(\lambda) - G(0)\|_2 = o(1)$, $\lambda \rightarrow 0$.

Remark 2.1. Using polar coordinates we see that item (a) imposes a certain integrability restriction on p :

$$\int_{\mathbb{R}^d} |g(t)| dt = \int_{\|\sigma\|_a=1} \left(\int_0^\infty |p(\rho)| \rho^{d-1} d\rho \right) d\sigma = |S|_a \int_0^\infty |p(\rho)| \rho^{d-1} d\rho < \infty, \quad (2.9)$$

where $|S|_a$ stands for the surface of the unit sphere $\{\sigma : \|\sigma\|_a = 1\}$. The class of densities satisfying (a) includes contaminated normal distribution, multivariate t -distribution, multivariate Cauchy distribution, and see; [8]. It does not matter much which norms are used in item (b).

Theorem 2.2. *Condition 1 is sufficient for (2.8) to hold on $\varphi \in CL(\mathbb{R}^n)$ such that $M_{a,d}\varphi_0$ exists.*

Remark 2.3. The triangle inequality $\|x + y\|_a \leq \|x\|_a + \|y\|_a$ is not used in the proof, and a slight generalization of Theorem 2.2 in terms of the geometry of balls is possible. The Lipschitz condition (2.6) can be omitted if F, G are constant (φ can be assumed just bounded and continuous).

The first application of Theorem 2.2 is to the theory of characteristic functions. Let F be the distribution function of a (proper, real-valued) random variable X . Denote $j(x)$ the jump of F at point x and let $\varphi(t) = Ee^{itX}$ be the characteristic function.

Corollary 2.4. *If $g \in L_1(\mathbb{R})$ is even, then*

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \int_{\mathbb{R}} e^{-itx} \varphi(t) \lambda g(\lambda t) dt &= j(x) \int_{\mathbb{R}} g(t) dt, \\ \lim_{\lambda \rightarrow 0} \int_{\mathbb{R}} |\varphi(t)|^2 \lambda g(\lambda t) dt &= \sum_k j^2(x_k) \int_{\mathbb{R}} g(t) dt, \end{aligned} \quad (2.10)$$

where the sum on the right is over all jump points x_k of F .

Theorems 3.2.3 and 3.3.4 in [2] are a special case of this corollary with $g = (1/2)\chi_{[-1,1]}$. The proof of Corollary 2.4 is obtained by combining those theorems with our Theorem 2.2.

The second application is to the theory of almost-periodic functions. A complex-valued continuous function f on R is called *almost-periodic* if for each $\varepsilon > 0$ there exists $l(\varepsilon) > 0$ such that each interval $(a, a + l(\varepsilon))$ of length $l(\varepsilon)$ contains at least one number τ for which $\sup_{t \in R} |f(t + \tau) - f(t)| < \varepsilon$. In the space of almost-periodic functions, the formula $(\varphi, \psi) = M\varphi\bar{\psi}$ defines a scalar product ($\bar{\psi}$ is a complex conjugate of ψ), and the numbers $\alpha(\lambda) = (\varphi(\cdot), e^{i\lambda\cdot})$ are called *Fourier coefficients* of φ . It is proved that for each φ at most a countable number of these coefficients are nonzero. Denoting them $\alpha(\lambda_k)$, $k = 1, 2, \dots, H$. Bohr's theorem states that $M|\varphi|^2 = \sum_k |\alpha(\lambda_k)|^2$; see [3, page 134]. This fact together with our Theorem 2.2 gives the next corollary.

Corollary 2.5. *If $g \in L_1(R)$ is even, then for any almost-periodic function φ*

$$\lim_{\lambda \rightarrow 0} \int_R |\varphi(t)|^2 \lambda g(\lambda t) dt = \sum_k |\alpha(\lambda_k)|^2 \int_R g(t) dt. \quad (2.11)$$

Further, if $\int_R g(t) dt > 0$, the formula $(\varphi, \psi)_g = \lim_{\lambda \rightarrow 0} \int_R \varphi(t) \bar{\psi}(t) \lambda g(\lambda t) dt$ defines a scalar product which is equivalent to (φ, ψ) .

Note that the first part of this corollary applies to characteristic functions of purely discrete distribution functions because from [2, Corollary 2 of Theorem 3.2.3] any such characteristic function is almost periodic. Now, we turn to the multidimensional version of (1.7). For the density, we assume a stronger condition than Condition 1(a).

Condition 2. $\bar{g} \in L_1(R^n)$ is $\|\cdot\|_2$ -spherically symmetric, $\bar{g}(x) = \bar{p}(\|x\|_2)$.

This assumption allows us to show that when some coordinates of x are fixed, $\bar{g}(x)$ as a function of the remaining coordinates satisfies Condition 1(a).

Now, we provide the intuition for the next condition. The stretching-out applied in (1.7) is described by the transformation $s = A(\lambda)t$ where

$$A(\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \quad \lambda > 0. \quad (2.12)$$

$g_\lambda(t) = \det A(\lambda) g(A(\lambda)t)$ is the analog of $g_\lambda(t) = \lambda g(\lambda t)$ because $\int_{R^2} g_\lambda(s) ds = \int_{R^2} g(t) dt$. Here, the matrix $A(\lambda)$ has two eigenvalues. The limit $\lim_{\lambda \rightarrow 0} A(\lambda) = A(0)$ is a singular matrix because one of these eigenvalues tends to zero as $\lambda \rightarrow 0$. Generalizing upon this situation and also thinking of applications to invariant tests, in the n -dimensional case we consider a symmetric nonnegative matrix $A(\rho)$ of size $n \times n$, where the parameter ρ belongs to the segment $[0, a]$. Denote its eigenvalues $0 \leq \lambda_1(\rho) \leq \dots \leq \lambda_n(\rho)$ and let $A(\rho)$ be diagonalized as $A(\rho) = P(\rho)\Lambda(\rho)P'(\rho)$, where $\Lambda(\rho) = \text{diag}[\lambda_1(\rho), \dots, \lambda_n(\rho)]$ and $P(\rho)$ is an orthogonal matrix. $A(\rho)$ degenerates at the right end of the segment $[0, a]$, owing to the following assumption.

Condition 3. The matrix $A(\rho)$ is positive definite for $0 \leq \rho < a$.

The first d eigenvalues tend to zero at the same rate, $\lambda_j(\rho) = \lambda_j(a)(1 + o(1))$, $j = 1, \dots, d$; $\lambda_1(\rho) \rightarrow 0$ as $\rho \rightarrow a$; the remaining ones have positive limits: $\lambda_j(\rho) \rightarrow \lambda_j(a) > 0$ as $\rho \rightarrow a$, $j = d + 1, \dots, n$.

The matrices $P(\rho)$ and $\Lambda(\rho)$ converge sufficiently quickly as $\rho \rightarrow a$. Namely, with the matrix

$$\tilde{\Lambda}(\rho) = \text{diag} \left[\underbrace{\lambda_1(\rho), \dots, \lambda_1(\rho)}_{d \text{ times}}, \lambda_{d+1}(a), \dots, \lambda_n(a) \right] \quad (2.13)$$

one has $\|\Lambda^{-1}(\rho)\tilde{\Lambda}(\rho) - I\| = o(\lambda_1(\rho))$, $\|P(\rho) - P(a)\| = o(\lambda_1(\rho))$, $\rho \rightarrow a$.

Note that $\Lambda^{-1}(a)$ does not exist because of part (b), but part (c) allows us to set $\Lambda^{-1}(a)\tilde{\Lambda}(a) = I$ by continuity. In line with part (b), we use the partitions

$$t = \begin{pmatrix} t^{(1)} \\ t^{(2)} \end{pmatrix}, \quad t^{(1)} = (t_1, \dots, t_d)', \quad t^{(2)} = (t_{d+1}, \dots, t_n)', \quad (2.14)$$

$$\Lambda(\rho) = \begin{pmatrix} \Lambda_1(\rho) & 0 \\ 0 & \Lambda_2(\rho) \end{pmatrix}, \quad \Lambda_1(\rho) \text{ is } d \times d, \quad \Lambda_2(\rho) \text{ is } (n-d) \times (n-d). \quad (2.15)$$

Also, partition $P(a) = (P_1(a), P_2(a))$ conformably with (2.14) ($P_1(a)$ is $n \times d$ and $P_2(a)$ is $n \times (n-d)$). Define a transformation T by $(T\varphi)(t) = \varphi(P_1(a)t^{(1)} + P_2(a)\Lambda_2^{-1}(a)t^{(2)})$. $(M_{2,d}T\varphi)(t^{(2)})$ denotes the result of application of the generalized mean operator with respect to $t^{(1)}$, keeping $t^{(2)}$ fixed:

$$(M_{2,d}T\varphi)(t^{(2)}) = \lim_{r \rightarrow \infty} m\left((T\varphi)(\cdot, t^{(2)}), B_{2,d}(0, r)\right), \quad (2.16)$$

and \bar{g}_1 denotes a marginal "density":

$$\bar{g}_1(t^{(2)}) = \int_{\mathbb{R}^d} \bar{g}(t^{(1)}, t^{(2)}) dt^{(1)}. \quad (2.17)$$

Theorem 2.6. *Let Conditions 2 and 3 hold and let $\varphi \in CL(\mathbb{R}^n)$ be such that the limit (2.16) exists for almost all $t^{(2)} \in \mathbb{R}^{n-d}$. Then,*

$$\lim_{\rho \rightarrow a} \int_{\mathbb{R}^n} \varphi(t) \det A(\rho) g(A(\rho)t) dt = \int_{\mathbb{R}^{n-d}} (M_{2,d}T\varphi)(t^{(2)}) \bar{g}_1(t^{(2)}) dt^{(2)}. \quad (2.18)$$

Next, we turn to the solution of Problem 2. Denoting $A(\rho) = \Sigma^{-1/2}(\rho)$ and assuming that $A(\rho)$ is positive definite for $\rho \in [0, a)$, we see that condition (1.10) corresponds to the case $d = 1$ of our Condition 3. The density (1.9) fits the framework of our Theorem 2.6 because the stretching-out is applied along one variable. Thus, in Theorem 2.7 we apply Theorem 2.6 to (a) characterize the limit in distribution of f_ρ in case $1 \leq \text{rank}(\Sigma^{-1}(a)) \leq n-1$ and (b) show that (1.11) does not hold under the weak convergence. By implication, (1.11) is wrong if any convergence stronger than the weak one is considered (e.g., uniform, almost sure, in probability and in L_p). Even though we use Theorem 2.6, the assumption on the density in part (b) of the next theorem is weaker than that in Theorem 2.6.

Condition 4. g is a density on R^n bounded by an integrable, $\|\cdot\|_2$ -spherically symmetric function, $g(x) \leq p(\|x\|_2)$.

An example of such a density is $g(x) = p(\|x\|_a)$, where p is a nonnegative and nonincreasing function on $[0, \infty)$, which decays at infinity sufficiently quickly for g to be integrable. By monotonicity of p inequality (3.14) in the following implies that $g(x) = p(\|x\|_a) \leq p(\|x\|_2/c_3)$.

Theorem 2.7. *Let Condition 3 hold in which $A(\rho) = \Sigma^{-1/2}(\rho)$.*

- (a) *Denote $Q = P_2(a)P_2'(a)$ the projector onto the subspace spanned by the last $n - d$ eigenvectors of $A(a)$ and let T_1 be defined by*

$$(T_1\varphi)(t) = \varphi\left(P_1(a)t^{(1)} + P_2(a)\Lambda_2^{-1}(a)t^{(2)} + QX\beta\right). \quad (2.19)$$

$M_{2,d}T_1\varphi$ is obtained by replacing T with T_1 in (2.16). If Condition 2 holds and $\varphi \in CL(R^n)$ is such that $(M_{2,d}T_1\varphi)(t^{(2)})$ exists for almost all $t^{(2)} \in R^{n-d}$, then

$$\lim_{\rho \rightarrow a} \int_{R^n} \varphi(t) f_\rho(t) dt = \int_{R^{n-d}} (M_{2,d}T_1\varphi)(t^{(2)}) \bar{g}_1(t^{(2)}) dt^{(2)}, \quad (2.20)$$

with \bar{g}_1 defined in (2.17).

- (b) *If Condition 4 holds, then (1.11) cannot be true if the convergence in distribution is understood.*

Remark 2.8. Because of the identity $X\beta + N(A(a)) = (I - Q)X\beta + QX\beta + N(A(a)) = QX\beta + N(A(a))$, (1.11) correctly captures one feature of the limit distribution: it depends on $X\beta$ only through $QX\beta$.

Before giving the solution to Problem 3, we need more notation and definitions. Obtaining a closed-form formula for $\beta(\Phi)$ involves a meticulous analysis of Φ based on the representation of Φ given in the following.

Let $\text{Im } X = \{X\beta : \beta \in R^k\}$ be the image of X . If for a given set $\Phi \subset R^n$ (a) the space R^n is represented as an orthogonal sum $R^n = E_1 \oplus E_2$ of two subspaces E_1 and E_2 and (b) there is a set $B \subset E_1$ such that $\Phi = B + E_2 = \{b + e : b \in B, e \in E_2\}$, then we say that Φ is a *cylinder* with the *base* B and *element* E_2 . For any set $S \subset R^n$, denote $\Gamma S = \{\gamma s : s \in S, \gamma \in R, \gamma \neq 0\}$. Using $s = 1 \cdot s$, we see that $S \subset \Gamma S$. We say that S is *cone-like* if $S = \Gamma S$. In particular, a cone-like set with each of its element s contains its opposite $-s$. The next representation is a stronger statement than saying that Φ defined in (1.14) is invariant with respect to transformations $y \rightarrow \gamma y + X\beta, \gamma \in R, \gamma \neq 0, \beta \in R^k$.

Representation of Φ

The rejection region Φ for the Cliff-Ord test is a cylinder

$$\Phi = \Gamma S + \text{Im } X, \quad (2.21)$$

with a cone-like base $M_X\Phi = \Gamma S$, where $S = \{s \in \text{Im } M_X : \|s\|_2 = 1, s'Ws > c\}$.

It is convenient to call an *aperture* the set S in representation (2.21). The function $s'Ws$ is continuous on the unit sphere of $\text{Im } M_X$ and the set (c, ∞) is open. By the general property of continuous mappings [11, Chapter 2, § 5, Section 5, Theorem 6], the preimage of an open set under a continuous mapping is open. Thus, the aperture is an open set (in the relative topology) of the unit sphere of $\text{Im } M_X$. We need the notations of the interior $\text{int}(\Phi)$ (defined as the set of points of Φ that belong to Φ with some neighborhood), closure $\text{cl}(\Phi)$ (the set of all limit points of Φ) and boundary $\text{bd}(\Phi) = \text{cl}(\Phi) \setminus \Phi$.

Writing (1.12) in the form

$$\beta_\rho(\Phi) = \int_{R^n} \chi_\Phi(y) f_\rho(y) dy, \quad (2.22)$$

we see that Theorem 2.7(a) will be applicable if we manage to extend it from continuous Lipschitz functions to discontinuous functions of type χ_Φ . This is done in Theorem 2.9. In Theorems 2.9 and 2.11, we assume that $d = 1$. Up to the notation, this is the same assumption as (1.10). At least within our method, generalizations of the results in the following to the case $d > 1$ are hard to obtain.

Let Condition 3 hold for $A(\rho) = \Sigma^{-1/2}(\rho)$ and let $d = 1$. Then, $\lambda_1(\rho) \rightarrow 0$ as $\rho \rightarrow a$ and all other eigenvalues have positive limits. Denote f_1, \dots, f_n the orthonormal eigenvectors of $A(a) = \Sigma^{-1/2}(a)$ corresponding to the eigenvalues $\lambda_1(a) = 0, \lambda_2(a), \dots, \lambda_n(a)$. The partitions (2.14) become $t^{(1)} = t_1, t^{(2)} = (t_2, \dots, t_n)'$, $P_1(a) = f_1, P_2(a) = (f_2, \dots, f_n)$. The vector $z(t^{(2)}) = P_2(a)\Lambda_2^{-1}(a)t^{(2)} + QX\beta$ will be called a *shift* because its role is to shift the line $P_1(a)t^{(1)} + P_2(a)\Lambda_2^{-1}(a)t^{(2)} + QX\beta = f_1t_1 + z(t^{(2)})$. In the next theorem, we extend (2.20) to $\varphi = \chi_\Phi$ with Φ described by (2.21). In the notation of marginal densities, the subscript will indicate the number of integrated-out variables. For example, $\bar{g}_1(t^{(2)}) = \int_R \bar{g}(t_1, t^{(2)}) dt_1$, $\bar{g}_k(t) = \int_{R^k} \bar{p}((t^2 + \|v\|_2^2)^{1/2}) dv$.

Theorem 2.9 (representation of $\beta(\Phi)$). *Let Conditions 2 and 3 hold with $d = 1$.*

(a) *If $f_1 \in \text{Im } X$ and*

$$\int_0^1 \bar{g}_k(t) dt < \infty, \quad (2.23)$$

then

$$\beta(\Phi) = \int_{R^{n-1}} (MT_1\chi_\Phi)\left(t^{(2)}\right) \bar{g}_1\left(t^{(2)}\right) dt^{(2)}, \quad (2.24)$$

where $T_1\chi_\Phi$ is defined by $(T_1\chi_\Phi)(t) = \chi_\Phi(f_1t_1 + z(t^{(2)}))$, and the generalized mean is applied over t_1 .

(b) *If $f_1 \notin \text{Im } X$ and*

$$\int_0^1 \bar{g}_{k+1}(t) dt < \infty, \quad (2.25)$$

then (2.24) is true.

Remark 2.10. (a) As can be seen from the proof, the theorem holds for any test with the critical region satisfying (2.21). (b) Conditions (2.23) and (2.25) are technical assumptions that provide integrability in the neighborhood of the origin of marginal densities that arise in the course of the proof.

To avoid triviality, in the following theorem we assume that the inclusion $\emptyset \subset \Phi \subset R^n$ is strict. This implies that in representation (2.21) the set ΓS is a nonempty proper subset of $\text{Im } M_X$.

Theorem 2.11. *Let conditions of Theorem 2.9 hold, with (2.23) accompanying $f_1 \in \text{Im } X$ and (2.25) accompanying $f_1 \notin \text{Im } X$. Besides, we require the function \bar{p} from Condition 2 to be positive on $[0, \infty)$.*

(1) *If $f_1 \in \Phi$, then $\beta(\Phi) = 1$.*

(2) *Suppose $f_1 \in \text{bd } (\Phi)$.*

(2.1) *If $f_1 \in \text{Im } X$, then $\beta(\Phi) \in (0, 1)$.*

(2.2) *If $f_1 \notin \text{Im } X$, then examples can be presented such that $\beta(\Phi) = 0$, $\beta(\Phi) \in (0, 1)$ or $\beta(\Phi) = 1$.*

(3) *If $f_1 \in \Phi_-$, then $\beta(\Phi) = 0$.*

Remark 2.12. Here, we compare this theorem with [4, Theorem 1].

(1) Our conditions on the density and critical region are much more restrictive. Our proof reveals the distinction between the cases $f_1 \in \text{Im } X$ and $f_1 \notin \text{Im } X$. In particular, in case (2.2), Martellosio excludes the extreme values $\beta(\Phi) = 0$ and $\beta(\Phi) = 1$, while we provide counter examples showing that they are possible. Note also that we do not impose any conditions on the structure of W . What happens to the limiting power in case of W that arises in practice needs additional investigation.

(2) Martellosio's proof is based on (1.11) which we disprove in Theorem 2.7(b). A series of other propositions from the same paper (see Lemmas D.2, D.3 and E.4, Corollary 1 and Propositions 1, 2 and 5), as well as from [7] (see Lemma 3.2, Theorems 3.3, 3.5 and 4.1, and Proposition 3.6), depend on (1.11) and need a revision. In particular, his claim that his results are true for any invariant critical region and any continuous density u that is unimodal at the origin is unwarranted.

(3) Even if statement (1.11) were right, the proof of [4, Theorem 1] would be incomplete because it incorrectly uses the n -dimensional Lebesgue measure. Its use is inappropriate because for a degenerate density even one point may carry a positive mass. In our proof, we justify the use of $(n - 1)$ -dimensional Lebesgue measure.

Now, we turn to the description of the alternative approach (solution to Problem 5). A mode is not a very good characteristic of a distribution when two unimodal densities with the same modes have very different spreads. A set of points where the density is close to its maximum might be a better characteristic in this case, at least for bell-shaped distributions. Let $0 < \varepsilon < 1$ and let $m = \max f(x)$ be the maximum of a continuous density $f(x)$. We call $M_\varepsilon = \{x : f(x) \geq (1 - \varepsilon)m\}$ an ε -maximizing set of f . The idea of this density-maximizing approach is close to the maximum likelihood principle.

Suppose a decision d is taken if the statistic x belongs to a set D . In case of a favorable decision, D is chosen in such a way that the probability $P(x \in D)$ is high. The use of probability in this decision rule presumes that the statistic can be calculated repeatedly. However, in

practice, especially in economics, the decision is based on just one value of the statistic in question. In such a case it may be preferable to choose D so that $\min_{x \in D} f(x)$ is sufficiently close to $m = \max f(x)$. For the ε -maximizing set, we have $\min_{x \in M_\varepsilon} f(x) \geq (1 - \varepsilon)m$. Requiring ε to be close to 1 in the density-maximizing approach is similar to requiring $P(x \in D)$ to be close to 1 in the probability-maximizing approach (although normally $P(x \in M_\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$).

As before, we assume that the matrix $A(\rho) = \Sigma^{-1/2}(\rho)$ of size $n \times n$ is symmetric for $0 \leq \rho \leq a$ and positive definite for $0 \leq \rho < a$. The idea is to impose conditions ensuring that ε -maximizing sets are ellipsoids which in the limit give a set of the desired shape. This idea is realized through a delicate balance of the limit behavior of the eigenvalues and density contained in Conditions 5 through 7 and (2.32).

Condition 5. In the diagonal representation of $A(\rho)$, the orthogonal matrix P satisfies

$$P(a) = \lim_{\rho \rightarrow a} P(\rho), \quad (2.26)$$

the first d eigenvalues vanish as power functions

$$\lambda_j(\rho) = c_\lambda (a - \rho)^{\alpha_\lambda} (1 + o(1)), \quad j = 1, \dots, d, \text{ as } \lambda \rightarrow a, \quad (2.27)$$

with positive constants $c_\lambda, \alpha_\lambda$; the remaining eigenvalues tend to positive constants

$$\lambda_j(\rho) = \lambda_j(a) (1 + o(1)), \quad j = d + 1, \dots, n, \text{ as } \lambda \rightarrow a, \quad (2.28)$$

where $\lambda_j(a) > 0$.

Condition 6. The function $\varepsilon(\rho)$ in the definition of the set $M_{\varepsilon(\rho)}$ vanishes as a power function

$$\varepsilon(\rho) = c_\varepsilon (a - \rho)^{\alpha_\varepsilon} (1 + o(1)), \quad \rho \rightarrow a, \quad (2.29)$$

where $c_\varepsilon, \alpha_\varepsilon$ are positive constants.

Condition 7. The density g is $\|\cdot\|_2$ -spherically symmetric, $g(x) = p(\|x\|_2)$, where p is continuous and monotonically decreasing on $[0, \infty)$ and such that

$$p(r) = m (1 - c_p r^{\alpha_p}) (1 + o(r^{\alpha_p})), \quad r \rightarrow 0, \quad (2.30)$$

where $c_p, \alpha_p > 0$.

This assumption implies that $m = \max g(x) = p(0)$, that the inverse function p^{-1} is continuous and monotonically decreasing on $(0, m]$, and that

$$0 = p^{-1}(m) = \lim_{x \rightarrow m} p^{-1}(x). \quad (2.31)$$

Theorem 2.13. Denote $N(A(a))$ the null subspace of $A(a)$ (spanned by the eigenvectors corresponding to its zero eigenvalues). If Conditions 5–7 hold and

$$\frac{\alpha_\varepsilon}{\alpha_p - \alpha_\lambda} < 0, \quad (2.32)$$

then the $\varepsilon(\rho)$ -maximizing set $M_{\varepsilon(\rho)} = \{x : f_\rho(x) \geq (1 - \varepsilon(\rho))m\}$ converges to $X\beta + N(A(a))$ as $\rho \rightarrow a$.

3. Proofs

The idea of the proof is to approximate g with a step function h , prove the statement for h and then pass to the limit to obtain the statement for g . A *step function*, by definition, is a finite linear combination of indicators of measurable sets. Due to the spherical symmetry of g , these sets turn out to be balls. For the method to work, the radii of the balls should be positive and finite. Approximation of p by a continuous function in Step 1 is a trick to make sure that h takes a finite number of values.

Step 1 (p can be assumed continuous on $(0, \infty)$). Let $\Omega = (0, \infty)$ and denote $Z(\Omega)$ the space of measurable functions f on Ω such that $\|f\|_{Z(\Omega)} = \int_0^\infty |f(\rho)|\rho^{d-1}d\rho < \infty$. $Z(\Omega)$ satisfies conditions of [12, Theorem 1]:

- (1) $C_0^\infty(\Omega) \subset Z(\Omega) \subset L_1^{\text{loc}}(\Omega)$.
- (2) Minkowsky inequality: if $A \subset \mathbb{R}^m$ is a measurable set and $\Phi(x, y)$ is a measurable on $\Omega \times A$ function, then $\|\int_A \Phi(\cdot, y)dy\|_{Z(\Omega)} \leq \int_A \|\Phi(\cdot, y)\|_{Z(\Omega)}dy$.
- (3) A multiplication operator by a function $\varphi \in C_0^\infty(\Omega)$ is bounded in $Z(\Omega)$: $\|f\varphi\|_{Z(\Omega)} \leq \max |\varphi| \|f\|_{Z(\Omega)}$.
- (4) Any finite in Ω function $f \in Z(\Omega)$ is translation-continuous: $\lim_{h \rightarrow 0} \|f(\cdot + h) - f(\cdot)\|_{Z(\Omega)} = 0$.

The first three properties are standard facts of the theory of L_p spaces; the last one follows from the fact that if $\text{supp } f$ is a compact subset of Ω , then the weight ρ^{d-1} in the definition of the norm of $Z(\Omega)$ satisfies $c_1 \leq \rho^{d-1} \leq c_2$ on $\text{supp } f$ and therefore for such an f the norms $\|f\|_{Z(\Omega)}$ and $\|f\|_{L_1}$ are equivalent. Functions from L_1 are known to be translation-continuous.

By Burenkov's theorem there exists a sequence $\{p_s\} \subset C^\infty(\Omega) \cap Z(\Omega)$ such that $\|p_s - p\|_{Z(\Omega)} \rightarrow 0$. Defining $g_s(x) = p_s(\|x\|_a)$, from (2.9) we have $\|g_s - g\|_{L_1} = |S|_a \|p_s - p\|_{Z(\Omega)} \rightarrow 0$. Hence,

$$\begin{aligned} \left| \lambda^d \int_{\mathbb{R}^d} \varphi_\lambda(t) g_s(\lambda t) dt - \lambda^d \int_{\mathbb{R}^d} \varphi_\lambda(t) g(\lambda t) dt \right| &\leq \sup |\varphi| \lambda^d \int_{\mathbb{R}^d} |g_s(\lambda t) - g(\lambda t)| dt \\ &= \sup |\varphi| \|g_s - g\|_{L_1} \rightarrow 0 \end{aligned} \quad (3.1)$$

uniformly in λ . Thus, if we establish

$$\lim_{\lambda \rightarrow 0} \lambda^d \int_{\mathbb{R}^d} \varphi_\lambda(t) g_s(\lambda t) dt = M_{a,d} \varphi_0 \int_{\mathbb{R}^d} g_s(t) dt \quad \forall s, \quad (3.2)$$

then (2.8) will follow.

Step 2 (approximating g with a Step Function). Take an arbitrary $\varepsilon > 0$. We can assume that p is continuous on $(0, \infty)$.

- (a) Suppose $\|g\|_{L_1} \geq \varepsilon$. We approximate $g(t)$ by a step function $h(t)$, which will vanish where t is large or small. By summability of g , there exist $0 < N_1 = N_1(\varepsilon) < N_2 = N_2(\varepsilon) < \infty$ such that

$$\int_{\|t\|_a < N_1} |g(t)| dt \leq \frac{\varepsilon}{3}, \quad \int_{\|t\|_a > N_2} |g(t)| dt \leq \frac{\varepsilon}{3} \quad (3.3)$$

g is uniformly continuous on the ring $\{N_1 \leq \|t\|_a \leq N_2\}$. By Condition 1(a), for any natural m we can find $\delta > 0$ and split this ring into smaller rings

$$A_{m,l} = \{t : N_1 + l\delta \leq \|t\|_a < N_1 + (l+1)\delta\}, \quad l = 0, \dots, L \equiv \frac{N_2 - N_1}{\delta} - 1, \quad (3.4)$$

in such a way that g in each ring is close to its value on the inner boundary:

$$\sup_{t \in A_{m,l}} \left| g(t) - p\left(\frac{l}{m}\right) \right| \leq \frac{1}{m}. \quad (3.5)$$

Put $h(t) = 0$ if $\|t\|_a < N_1$ or $\|t\|_a > N_2$; $h(t) = p(l/m)$ for $t \in A_{m,l}$, $l = 0, \dots, L$. Combination of (3.3), (3.4), and (3.5) leads to

$$\begin{aligned} \|h - g\|_{L_1} &= \left(\int_{\|t\|_a < N_1} + \int_{\|t\|_a > N_2} \right) |g(t)| dt + \sum_{l=0}^L \int_{A_{m,l}} |h(t) - g(t)| dt \\ &\leq \frac{2\varepsilon}{3} + \frac{1}{m} \text{mes} B_{a,d}(0, N_2) = \frac{2\varepsilon}{3} + \frac{v_a N_2^d(\varepsilon)}{m}. \end{aligned} \quad (3.6)$$

Consequently, we can fix $m = m(\varepsilon)$ so that

$$\|h - g\|_{L_1} \leq \varepsilon. \quad (3.7)$$

- (b) If $\|g\|_{L_1} < \varepsilon$, we just put $h \equiv 0$ to get (3.7).

Step 3 (replacing rings by balls in the representation of h). Suppose h is not identically zero. The sets $A_{m,l}$ in (3.4) are concentric rings with finite positive radii of the inner and outer boundaries, and by construction, h is a finite linear combination of indicators of such rings.

Hence, it can be written as $h = \sum_{l=0}^L a_l (\chi_{B_{a,d}(0,r_l)} - \chi_{B_{a,d}(0,r_{l-1})})$, where a_l are some real numbers and the radii satisfy

$$N_1 = r_0 < \dots < r_L = N_2. \quad (3.8)$$

Therefore, with some new constants b_l , we can write h as a linear combination of indicators of balls:

$$h = \sum_{l=0}^L h_l, \quad \text{where } h_l = b_l \chi_{B_{a,d}(0,r_l)}. \quad (3.9)$$

If $h \equiv 0$, we put formally $r_0 = 1$, $b_0 = 0$, $L = 0$. Note that L , b_l and r_l all depend on ε and that the constants b_l may deviate from the values of h significantly.

Step 4 (introducing residuals for generalized means). Define the residual $R_1(r)$ by

$$R_1(r) = \begin{cases} m(\varphi_0, B_{a,d}(0,r)) / M_{a,d}\varphi_0 - 1, & \text{if } M_{a,d}\varphi_0 \neq 0, \\ m(\varphi_0, B_{a,d}(0,r)), & \text{if } M_{a,d}\varphi_0 = 0. \end{cases} \quad (3.10)$$

Then,

$$m(\varphi_0, B_{a,d}(0,r)) = \begin{cases} M_{a,d}\varphi_0(R_1(r) + 1), & \text{if } M_{a,d}\varphi_0 \neq 0, \\ R_1(r), & \text{if } M_{a,d}\varphi_0 = 0, \end{cases} \quad (3.11)$$

and in both cases

$$\lim_{r \rightarrow \infty} R_1(r) = 0. \quad (3.12)$$

By the Lipschitz condition (2.6) and Condition 1(b),

$$\begin{aligned} & |\varphi(F(\lambda)t + G(\lambda)) - \varphi(F(0)t + G(0))| \\ & \leq c_1 \|(F(\lambda) - F(0))t + G(\lambda) - G(0)\|_2 \\ & = o(\lambda) \|t\|_a + o(1) = o(\lambda)r + o(1) \quad \text{for } \|t\|_a \leq r. \end{aligned} \quad (3.13)$$

Here, we have used the fact that on R^d any two norms are equivalent, so

$$c_2 \|t\|_a \leq \|t\|_2 \leq c_3 \|t\|_a, \quad (3.14)$$

with some $c_2, c_3 > 0$. Equations (2.3) and (3.13) imply that

$$|m(\varphi(F(\lambda) \cdot + G(\lambda)) - \varphi_0(\cdot), B_{a,d}(0,r))| = o(\lambda)r + o(1) \quad \forall r > 0, \text{ as } \lambda \rightarrow 0, \quad (3.15)$$

where $o(\lambda)$ and $o(1)$ do not depend on r . From this bound and (3.11), it follows that

$$m(\varphi(F(\lambda) \cdot + G(\lambda)), B_{a,d}(0, r)) = \begin{cases} M_{a,d}\varphi_0(R_1(r) + 1) + R_2(\lambda, r), & \text{if } M_{a,d}\varphi_0 \neq 0, \\ R_1(r) + R_2(\lambda, r), & \text{if } M_{a,d}\varphi_0 = 0, \end{cases} \quad (3.16)$$

where $R_2(\lambda, r)$ is a new residual satisfying

$$R_2(\lambda, r) = o(\lambda)r + o(1) \quad \forall r > 0, \text{ as } \lambda \rightarrow 0. \quad (3.17)$$

Step 5 (proving (2.8) for the approximating function). Let $M_{a,d}\varphi_0 \neq 0$. For one term in (3.9) by (2.2), and the first equation in (3.16) we have

$$\begin{aligned} & \lambda^d \int_{\mathbb{R}^d} \varphi(F(\lambda)t + G(\lambda))h_l(\lambda t)dt \\ &= \lambda^d b_l \int_{B_{a,d}(0, r_l/\lambda)} \varphi(F(\lambda)t + G(\lambda))dt \\ &= \frac{v_a r_l^d b_l}{v_a (r_l/\lambda)^d} \int_{B_{a,d}(0, r_l/\lambda)} \varphi(F(\lambda)t + G(\lambda))dt \\ &= \int_{\mathbb{R}^d} h_l(s)ds \left\{ M_{a,d}\varphi_0 \left[R_1\left(\frac{r_l}{\lambda}\right) + 1 \right] + R_2\left(\lambda, \frac{r_l}{\lambda}\right) \right\}. \end{aligned} \quad (3.18)$$

Summation of these equations produces

$$\begin{aligned} & \lambda^d \int_{\mathbb{R}^d} \varphi(F(\lambda)t + G(\lambda))h(\lambda t)dt \\ &= \sum_{l=0}^L \int_{\mathbb{R}^d} h_l(s)ds \left\{ M_{a,d}\varphi_0 \left[R_1\left(\frac{r_l}{\lambda}\right) + 1 \right] + R_2\left(\lambda, \frac{r_l}{\lambda}\right) \right\} \\ &= M_{a,d}\varphi_0 \int_{\mathbb{R}^d} h(s)ds \\ &+ \sum_{l=0}^L \int_{\mathbb{R}^d} h_l(s)ds \left\{ M_{a,d}\varphi_0 R_1\left(\frac{r_l}{\lambda}\right) + R_2\left(\lambda, \frac{r_l}{\lambda}\right) \right\}. \end{aligned} \quad (3.19)$$

Here, by (3.8), (3.12), and (3.17), $R_1(r_l/\lambda) \rightarrow 0$, $R_2(\lambda, (r_l/\lambda)) = o(r_l) + o(1) = o(1)$, $l = 0, \dots, L$, as $\lambda \rightarrow 0$. From these relations and (3.19), we see that for the given ε there exists $\lambda(\varepsilon)$ such that

$$\left| \lambda^d \int_{\mathbb{R}^d} \varphi_\lambda(t)h(\lambda t)dt - M_{a,d}\varphi_0 \int_{\mathbb{R}^d} h(s)ds \right| \leq \varepsilon \quad \text{for } 0 < \lambda \leq \lambda(\varepsilon). \quad (3.20)$$

In case $M_{a,d}\varphi_0 = 0$, the only difference consists in application of the second equation in (3.16). The conclusion is (3.20) with $M_{a,d}\varphi_0 = 0$.

Step 6 (proving (2.8)). Using (3.7) and (3.20), we have

$$\begin{aligned}
& \left| \lambda^d \int_{R^d} \varphi_\lambda(t) g(\lambda t) dt - M_{a,d} \varphi_0 \int_{R^d} g(t) dt \right| \\
& \leq \left| \lambda^d \int_{R^d} \varphi_\lambda(t) [g(\lambda t) - h(\lambda t)] dt \right| \\
& \quad + \left| \lambda^d \int_{R^d} \varphi_\lambda(t) h(\lambda t) dt - M_{a,d} \varphi_0 \int_{R^d} h(t) dt \right| \\
& \quad + |M_{a,d} \varphi_0| \left| \int_{R^d} [h(t) - g(t)] dt \right| \\
& \leq \|\varphi\|_C \|g - h\|_{L_1} + \varepsilon + |M_{a,d} \varphi_0| \|g - h\|_{L_1} \\
& \leq (\|\varphi\|_C + 1 + |M_{a,d} \varphi_0|) \varepsilon, \quad 0 < \lambda \leq \lambda(\varepsilon).
\end{aligned} \tag{3.21}$$

Equation (2.8) follows because $\varepsilon > 0$ is arbitrary.

3.1. Proof of Theorem 2.6

Denote $J_\rho(\varphi) = \int_{R^n} \varphi(t) \det A(\rho) g(A(\rho)t) dt$.

Step 1 (replacing $\lim J_\rho(\varphi)$ by an equivalent limit). Using constancy of \bar{g} on spheres (Condition 2), write $J_\rho(\varphi)$ as

$$\begin{aligned}
J_\rho(\varphi) &= \int_{R^n} \varphi(y) \det \Lambda(\rho) \bar{g}(P(\rho) \Lambda(\rho) P'(\rho) y) dy \\
&= \det \Lambda(\rho) \int_{R^n} \varphi(y) \bar{g}(\Lambda(\rho) P'(\rho) y) dy \\
&= \det \Lambda(\rho) \int_{R^n} \varphi(y) \bar{g}(\tilde{\Lambda}(\rho) \tilde{\Lambda}^{-1}(\rho) \Lambda(\rho) P'(\rho) y) dy.
\end{aligned} \tag{3.22}$$

Denote $H^{-1}(\rho) = \tilde{\Lambda}^{-1}(\rho) \Lambda(\rho) P'(\rho)$ and replace $H^{-1}(\rho) y = t$ to get

$$J_\rho(\varphi) = \det \Lambda(\rho) \det H(\rho) \int_{R^n} \varphi(H(\rho)t) \bar{g}(\tilde{\Lambda}(\rho)t) dt. \tag{3.23}$$

By Condition 3(b,c),

$$\begin{aligned}
\det \Lambda(\rho) &= \lambda_1^d(\rho) \det \Lambda_2(a) (1 + o(1)), \\
\det H(\rho) &= \det(\tilde{\Lambda}^{-1}(\rho) \Lambda(\rho)) \det P'(\rho) = 1 + o(1).
\end{aligned} \tag{3.24}$$

From (3.23) and (3.24) we deduce

$$\begin{aligned} J_\rho(\varphi) &= \lambda_1^d(\rho) \det \Lambda_2(a) (1 + o(1)) \int_{R^n} \varphi(H(\rho)t) \bar{g}(\tilde{\Lambda}(\rho)t) dt = (1 + o(1)) \\ &\quad \times \det \Lambda_2(a) \int_{R^{n-d}} \left[\lambda_1^d(\rho) \int_{R^d} \varphi(H(\rho)t) \bar{g}(\lambda_1(\rho)t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)} \right] dt^{(2)}. \end{aligned} \quad (3.25)$$

Step 2 (changing the limit variable). Let us have a closer look at the expression in the square brackets in (3.25):

$$J_\rho(\varphi, t^{(2)}) \equiv \lambda_1^d(\rho) \int_{R^d} \varphi(H(\rho)t) \bar{g}(\lambda_1(\rho)t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)}. \quad (3.26)$$

Since $\Lambda^{-1}(a)\tilde{\Lambda}(a) = I$, here $H(a) = P(a)$, and by Condition 3(c),

$$\begin{aligned} \|H(\rho) - P(a)\| &= \|P(\rho)\Lambda^{-1}(\rho)\tilde{\Lambda}(\rho) - P(\rho) + P(\rho) - P(a)\| \\ &\leq \|P(\rho)\| \|\Lambda^{-1}(\rho)\tilde{\Lambda}(\rho) - I\| + \|P(\rho) - P(a)\| \\ &= o(\lambda_1(\rho)), \quad \rho \rightarrow a. \end{aligned} \quad (3.27)$$

Considering $\lambda = \lambda_1$ the independent variable, take $\rho = \rho(\lambda)$ as its function. The limit relation $\rho \rightarrow a$ translates to $\lambda \rightarrow 0$, and (3.26) and (3.27) become,

$$J_{\rho(\lambda)}(\varphi, t^{(2)}) = \lambda^d \int_{R^d} \varphi(H(\rho(\lambda))t) \bar{g}(\lambda t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)}, \quad (3.28)$$

$$\|H(\rho(\lambda)) - P(a)\| = o(\lambda), \quad \lambda \rightarrow 0, \quad (3.29)$$

respectively.

Step 3 (proving a preliminary version of (2.18)). Now we check that Theorem 2.2 applies to (3.28) for any fixed $t^{(2)}$. Letting

$$p(z) = \bar{p} \left(\left(z^2 + \|\Lambda_2(a)t^{(2)}\|_2^2 \right)^{1/2} \right), \quad z \geq 0, \quad g(t^{(1)}) = \bar{g}(t^{(1)}, \Lambda_2(a)t^{(2)}), \quad (3.30)$$

we see that $g \in L_1(R^d)$ is $\|\cdot\|_2$ -spherically symmetric:

$$g(t^{(1)}) = \bar{g}(t^{(1)}, \Lambda_2(a)t^{(2)}) = \bar{p} \left(\left(\|t^{(1)}\|_2^2 + \|\Lambda_2(a)t^{(2)}\|_2^2 \right)^{1/2} \right) = p(\|t^{(1)}\|_2). \quad (3.31)$$

Condition 1(a) is satisfied. As for Condition 1(b), partition $H = (H_1, H_2)$ conformably with (2.14) and put $F(\lambda) = H_1(\rho(\lambda))$, $G(\lambda) = H_2(\rho(\lambda))t^{(2)}$, $F(0) = H_1(a) = P_1(a)$, $G(0) = P_2(a)t^{(2)}$. Then, $H(\rho(\lambda))t = H_1(\rho(\lambda))t^{(1)} + H_2(\rho(\lambda))t^{(2)} = F(\lambda)t^{(1)} + G(\lambda)$, and Condition 1(b) follows from (3.29).

By Theorem 2.2,

$$J_{\rho(\lambda)}(\varphi, t^{(2)}) \longrightarrow J_{\rho(0)}(\varphi, t^{(2)}) \equiv (M_{2,d}\bar{\varphi})(t^{(2)}) \int_{R^d} \bar{g}(t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)}, \quad t^{(2)} \in R^{n-d}, \quad (3.32)$$

where $(M_{2,d}\bar{\varphi})(t^{(2)})$ is obtained by applying $M_{2,d}$ to $\bar{\varphi}(t) \equiv \varphi(P_1(a)t^{(1)} + P_2(a)t^{(2)})$ with respect to $t^{(1)}$. Functions (3.28) are bounded by an integrable function of $t^{(2)}$:

$$\begin{aligned} |J_{\rho(\lambda)}(\varphi, t^{(2)})| &\leq \lambda^d \|\varphi\|_C \int_{R^d} \bar{g}(\lambda t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)} \\ &= \|\varphi\|_C \int_{R^d} \bar{g}(t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)}. \end{aligned} \quad (3.33)$$

By the dominated convergence theorem, (3.32) and (3.33) imply that

$$\int_{R^{n-d}} J_{\rho(\lambda)}(\varphi, t^{(2)}) dt^{(2)} \longrightarrow \int_{R^{n-d}} J_{\rho(0)}(\varphi, t^{(2)}) dt^{(2)}, \quad \lambda \rightarrow 0. \quad (3.34)$$

Now, we combine (3.25), (3.28), (3.32), and (3.34) to get

$$\begin{aligned} \lim_{\rho \rightarrow a} J_{\rho}(\varphi) &= \det \Lambda_2(a) \int_{R^{n-d}} \left[(M_{2,d}\bar{\varphi})(t^{(2)}) \int_{R^d} \bar{g}(t^{(1)}, \Lambda_2(a)t^{(2)}) dt^{(1)} \right] dt^{(2)} \\ &= \int_{R^{n-d}} \left[(M_{2,d}\bar{\varphi})(\Lambda_2^{-1}(a)s^{(2)}) \int_{R^d} \bar{g}(t^{(1)}, s^{(2)}) dt^{(1)} \right] ds^{(2)} \\ &= \int_{R^{n-d}} (M_{2,d}T\varphi)(s^{(2)}) \bar{g}_1(s^{(2)}) ds^{(2)} \end{aligned} \quad (3.35)$$

because

$$(M_{2,d}\bar{\varphi})(\Lambda_2^{-1}(a)s^{(2)}) = \lim_{r \rightarrow \infty} m\left(\varphi\left(P_1(a) \cdot + P_2(a)\Lambda_2^{-1}(a)s^{(2)}, B_{2,d}(0, r)\right)\right) = (M_{2,d}T\varphi)(s^{(2)}). \quad (3.36)$$

3.2. Proof of Theorem 2.7

3.2.1. Part (a)

Step 1 ($M_{a,d}\varphi$ does not depend on the center of balls). Let us make definition (2.4) more precise: $(M_{a,d}\varphi)(x) = \lim_{r \rightarrow \infty} m(\varphi, B_{a,d}(x, r))$, $x \in R^d$. Here, we prove that

$$(M_{a,d}\varphi)(x) = (M_{a,d}\varphi)(0) \quad \text{for any } x \in R^d. \quad (3.37)$$

For $r \geq \|x\|_a$, one has implications $\|y\|_a \leq r - \|x\|_a \Rightarrow \|y - x\|_a \leq r \Rightarrow \|y\|_a \leq \|y - x\|_a + \|x\|_a \leq r + \|x\|_a$. Geometrically, they mean the inclusion $B_{a,d}(0, r - \|x\|_a) \subset B_{a,d}(x, r) \subset B_{a,d}(0, r + \|x\|_a)$. This inclusion implies that

$$\begin{aligned} & \left(\frac{r - \|x\|_a}{r} \right)^d m(\varphi, B_{a,d}(0, r - \|x\|_a)) \\ &= \left(\frac{r - \|x\|_a}{r} \right)^d \frac{1}{v_a(r - \|x\|_a)^d} \int_{B_{a,d}(0, r - \|x\|_a)} \varphi(y) dy \\ &\leq \frac{1}{v_a r^d} \int_{B_{a,d}(x, r)} \varphi(y) dy = m(\varphi, B_{a,d}(x, r)) \\ &\leq \left(\frac{r + \|x\|_a}{r} \right)^d \frac{1}{v_a(r + \|x\|_a)^d} \int_{B_{a,d}(0, r + \|x\|_a)} \varphi(y) dy \\ &= \left(\frac{r + \|x\|_a}{r} \right)^d m(\varphi, B_{a,d}(0, r + \|x\|_a)). \end{aligned} \quad (3.38)$$

It remains to let $r \rightarrow \infty$ to prove (3.37).

Step 2 (proving (2.20)). If $\varphi \in CL(R^n)$, then $\varphi(\cdot + X\beta) \in CL(R^n)$. By (2.16), we need to look at

$$\begin{aligned} & (M_{2,d}T\varphi(\cdot + X\beta)) \left(t^{(2)} \right) \\ &= \lim_{r \rightarrow \infty} \frac{1}{v_a r^d} \int_{B_{2,d}(0, r)} \varphi \left(P_1(a)t^{(1)} + P_2(a)\Lambda_2^{-1}(a)t^{(2)} + X\beta \right) dt^{(1)}. \end{aligned} \quad (3.39)$$

Using $I = P(a)P'(a) = P_1(a)P'_1(a) + Q$, we see that

$$\begin{aligned} P_1(a)t^{(1)} + X\beta &= P_1(a)t^{(1)} + P_1(a)P'_1(a)X\beta + QX\beta \\ &= P_1(a) \left(t^{(1)} + P'_1(a)X\beta \right) + QX\beta. \end{aligned} \quad (3.40)$$

Therefore, replacing $t^{(1)} + P'_1(a)X\beta = s^{(1)}$, we can continue (3.39) as

$$\begin{aligned} & (M_{2,d}T\varphi(\cdot + X\beta))(t^{(2)}) \\ &= \lim_{r \rightarrow \infty} \frac{1}{v_a r^d} \int_{B_{2,d}(P'_1(a)X\beta, r)} \varphi(P_1(a)s^{(1)} + P_2(a)\Lambda_2^{-1}(a)t^{(2)} + QX\beta) ds^{(1)}. \end{aligned} \quad (3.41)$$

Using the definition of T_1 and (3.37), we see that this is the same as $(M_{2,d}T_1\varphi)(t^{(2)})$. Hence, the existence of $(M_{2,d}T\varphi(\cdot + X\beta))(t^{(2)})$ is equivalent to that of $(M_{2,d}T_1\varphi)(t^{(2)})$. Application of Theorem 2.6 proves (2.20).

3.2.2. Part (b)

Step 1 (the generalized mean vanishes on p -integrable functions). If $\varphi \in L_p(R^d)$ with some $p \in [1, \infty)$, then by Hölder's inequality

$$\begin{aligned} |m(\varphi, B_{a,d}(0, r))| &\leq \frac{1}{v_a r^d} \int_{B_{a,d}(0, r)} |\varphi(t)| dt \\ &\leq \frac{1}{v_a r^d} (\text{mes} B_{a,d}(0, r))^{1-1/p} \left(\int_{B_{a,d}(0, r)} |\varphi(t)|^p dt \right)^{1/p} \\ &= (v_a r^d)^{-1/p} \|\varphi\|_{L_p} \rightarrow 0, \quad r \rightarrow \infty. \end{aligned} \quad (3.42)$$

Step 2 ($\lim J_\rho(\varphi)$ vanishes on functions with compact support). By Condition 4,

$$\left| \int_{R^n} \varphi(t) f_\rho(t) dt \right| \leq \int_{R^n} \tilde{\varphi}(t) \tilde{f}_\rho(t) dt, \quad (3.43)$$

where $\tilde{\varphi}(t) = |\varphi(t)|$ and \tilde{f}_ρ is obtained from (1.9) by replacing $g(x)$ with $\tilde{g}(x) = p(\|x\|_2)$. The function $\tilde{\varphi}$ satisfies the Lipschitz condition, $|\tilde{\varphi}(x) - \tilde{\varphi}(y)| = ||\varphi(x)| - |\varphi(y)|| \leq c_1 \|x - y\|_2$, and \tilde{g} satisfies Condition 2. Part (a) of this theorem allows us to use (2.20) with $\tilde{\varphi}$ in place of φ and \tilde{g} instead of g whenever $(M_{2,d}T_1\tilde{\varphi})(t^{(2)})$ exists for almost all $t^{(2)} \in R^{n-d}$.

Let $CL_0(R^n)$ denote the set of functions $\varphi \in CL(R^n)$ with compact support $\text{supp } \varphi$. Suppose $\varphi \in CL_0(R^n)$. By Step 1, $(M_{2,d}T_1\tilde{\varphi})(t^{(2)}) \equiv 0$. Equations (2.20) and (3.43) imply that

$$\lim_{\rho \rightarrow a} \int_{R^n} \varphi(t) f_\rho(t) dt = 0, \quad \varphi \in CL_0(R^n). \quad (3.44)$$

Suppose (1.11) is true in the weak sense or, more generally, the family $\{f_\rho : 0 \leq \rho < a\}$ converges in distribution, as $\rho \rightarrow a$, to a probabilistic measure μ supported on some set $S \subset R^n$. Then by (2.7) $\int_{R^n} \varphi(t) f_\rho(t) dt \rightarrow \int_S \varphi d\mu$, $\varphi \in CL(R^n)$. This equation and (3.44) imply $\int_S \varphi d\mu = 0$, $\varphi \in CL_0(R^n)$. Taking here $\varphi \in CL_0(R^n)$ such that $\varphi \geq 0$ and $\varphi = 1$ on $B_{2,n}(0, r)$ we have $\int_{S \cap B_{2,n}(0, r)} d\mu = 0$ for any $r > 0$. Thus, $\mu(S) = 0$ which is impossible.

3.3. Proof of Theorem 2.9

The proof of Theorem 2.9 will be preceded by a series of lemmas. The main difficulty consists in estimating integrals of the marginal density \bar{g}_1 over cylindrical sets of an infinite Lebesgue measure.

3.3.1. Preliminary Results

Lemma 3.1. *For the rejection region Φ defined in (1.14), representation (2.21) holds.*

Proof. Since L_X, M_X are complementary orthoprojectors, the images $\text{Im } M_X$ and $\text{Im } L_X = \text{Im } X$ are orthogonal to one another and $R^n = \text{Im } M_X \oplus \text{Im } X$. Representing $u \in \Phi$ as $u = L_X u + M_X u$ and using $M_X' = M_X, M_X^2 = M_X$ we get

$$\frac{u' M_X W M_X u}{u' M_X u} = \frac{(M_X u)' W M_X u}{(M_X u)' M_X u} = \frac{v' W v}{v' v}, \quad (3.45)$$

where we have denoted $v = M_X u$. Hence, only the component $v \in M_X \Phi$ determines whether u belongs to Φ or not, while the component $L_X u \in \text{Im } X$ is arbitrary. This proves the equation $\Phi = M_X \Phi + \text{Im } X$. Finally, the equation $M_X \Phi = \Gamma S$ follows by homogeneity from (3.45). \square

Now we review a few facts about the Moore-Penrose inverse in a form adapted to our needs. Let A be a matrix of size $m \times n$, not necessarily square, and denote its rank $r = \text{rank } A$. Its Moore-Penrose inverse A^+ is of size $n \times m$.

Lemma 3.2 (see, [13, pages 24-25]). *A^+ is unique and has the following properties.*

- (a) $AA^+A = A$ (A^+ is a generalized inverse of A).
- (b) $A^+AA^+ = A^+$ (A is a generalized inverse of A^+).
- (c) The matrix $H = A^+A$ is an orthoprojector, $H^2 = H, H' = H$, and $\text{rank } H = \text{rank } A^+ = r$.
- (d) For a given $y \in \text{Im } A$, a general solution of the equation $Ax = y$ is given by

$$x = A^+y + (I - H)z, \quad (3.46)$$

where $z \in R^n$ is arbitrary. (Rao writes $(H - I)z$ instead of $(I - H)z$, but our form is more convenient).

Regarding (3.46), we can remark the following.

- (i) The two terms on the right are orthogonal because by (b) and (c)

$$(A^+y, (I - H)z) = (HA^+y, (I - H)z) = (A^+y, (H - H^2)z) = 0. \quad (3.47)$$

Here and in the sequel, (\cdot, \cdot) stands for the scalar product.

- (ii) Instead of writing the second term in the form $(I-H)z$, $z \in R^n$, we can equivalently write it as $v \in \text{Im}(I-H)$.
- (iii) $A^+y = 0$ implies that $y = 0$. This is because if $A^+y = 0$ and $y = Ax$, then by (a) $y = Ax = AA^+(Ax) = A(A^+y) = 0$.

Recalling that a general solution of the equation $Ax = y$ is, by definition, the preimage $A^{-1}y$, we state our conclusion in a lemma.

Lemma 3.3. *Let $Y \subseteq \text{Im } A$. The preimage $A^{-1}Y$ is a cylinder $A^+Y + \text{Im}(I-H)$ with the base $A^+Y \subseteq \text{Im } H$ and element $\text{Im}(I-H)$. Further, $A^+y = 0$ for $y \in \text{Im } A$ implies that $y = 0$, that is, A^+ on $\text{Im } H$ is one-to-one.*

Lemma 3.4. *Suppose $f_1 \in \text{Im } X$. Fix a set $Y \subseteq \text{Im } M_X$ and a vector $\tau \in \text{Im } M_X$. Denote $A = M_X P_2(a) \Lambda_2^{-1}(a)$ and let A^+ and H be the accompanying matrices from Lemma 3.2. Then, the set $T^{(2)}$ defined by $T^{(2)} = \{t^{(2)} \in R^{n-1} : At^{(2)} + \tau \in Y\}$ is a cylinder:*

$$T^{(2)} = A^+(Y - \tau) + \text{Im}(I-H). \quad (3.48)$$

Suppose $f_1 \notin \text{Im } X$ and denote $p = M_X f_1 / \|M_X f_1\|_2$. Then $M_X - pp'$ is an orthoprojector with the image

$$\text{Im}(M_X - pp') = \{x \in \text{Im } M_X : x \perp p\}. \quad (3.49)$$

In part (a) replace Y, τ, A by $Y \subseteq \text{Im}(M_X - pp')$, $\tau \in \text{Im}(M_X - pp')$, and $A = (M_X - pp')P_2(a)\Lambda_2^{-1}(a)$, respectively. Then the set $T^{(2)}$ has representation (3.48).

Proof. Let us show that $f_1 \in \text{Im } X$ implies that

$$\text{Im } A = \text{Im } M_X. \quad (3.50)$$

The inclusion $\text{Im } A \subset \text{Im } M_X$ is obvious. Conversely, let $x \in \text{Im } M_X$. Then, $x = M_X x$. Since f_1, \dots, f_n is an orthonormal basis, we have $x = \sum_{i=1}^n (x, f_i) f_i$. Denoting $t^{(2)} = ((x, f_2), \dots, (x, f_n))'$, it follows that

$$\begin{aligned} x = M_X x &= \sum_{i=1}^n (x, f_i) M_X f_i = \sum_{i=2}^n (x, f_i) M_X f_i \\ &= (M_X f_2, \dots, M_X f_n) t^{(2)} = M_X P_2(a) t^{(2)} = A \Lambda_2(a) t^{(2)} \in \text{Im } A. \end{aligned} \quad (3.51)$$

We have proved (3.50) which implies $Y - \tau \subset \text{Im } A$. Application of Lemma 3.3 proves (3.48).

It is easy to check that $M_X - pp'$ is an orthoprojector and its image satisfies (3.49). From (3.49), we conclude that M_X and $M_X - pp'$ commute and

$$(M_X - pp')M_X = M_X - pp', \quad (3.52)$$

see [14, Chapter VI, Section 3, Lemma 2(a)]. Applying $M_X - pp'$ to both sides of (3.50) and using the last equation we get

$$\begin{aligned} \operatorname{Im} \left[(M_X - pp') P_2(a) \Lambda_2^{-1}(a) \right] &= \operatorname{Im} \left[(M_X - pp') M_X P_2(a) \Lambda_2^{-1}(a) \right] \\ &= (M_X - pp') \left[\operatorname{Im} M_X P_2(a) \Lambda_2^{-1}(a) \right] \\ &= (M_X - pp') \operatorname{Im} M_X = \operatorname{Im} (M_X - pp'). \end{aligned} \quad (3.53)$$

We have proved an analog of (3.50) for the matrix under consideration:

$$\operatorname{Im} A = \operatorname{Im} (M_X - pp'). \quad (3.54)$$

As previously mentioned, application of Lemma 3.3 proves (3.48). \square

Lemma 3.5. *Let f_1, τ, A, A^+ and H be the same as in Lemma 3.4(a).*

- (a) *Suppose $Y \subseteq \operatorname{Im} M_X$ is of the form $Y = \Gamma \Delta$ where $\Delta \subseteq \{s \in \operatorname{Im} M_X : \|s\|_2 = 1\}$. If (2.23) holds, then for the set $T^{(2)}$ from (3.48) one has*

$$\alpha(Y, \tau) \equiv \int_{T^{(2)}} \bar{g}_1(t^{(2)}) dt^{(2)} \leq c(\|\tau\|_2) \operatorname{mes}(\Delta). \quad (3.55)$$

- (b) *If $Y \subseteq \{s \in \operatorname{Im} M_X : \|s\|_2 \leq 1\}$, then $\alpha(Y, \tau) \rightarrow 0$, as $\operatorname{mes}(Y) \rightarrow 0$, uniformly in τ .*

Proof. Part (a)

Step 1 (proving an auxiliary bound). Denote

$$\beta(Y, \tau) = \int_{A^+(Y-\tau)} \bar{g}_k(\|u\|_2) du. \quad (3.56)$$

We intend to prove that

$$\beta(\Gamma \Delta, \tau) \leq c(\|\tau\|_2) \operatorname{mes}(\Delta). \quad (3.57)$$

By Lemma 3.2, (3.50) implies that

$$r \equiv \operatorname{rank} A = n - k. \quad (3.58)$$

By Lemma 3.3, A^+ is a one-to-one mapping from $\operatorname{Im} A$ to $\operatorname{Im} A^+$, both of which are of dimension r . We use this fact twice.

Firstly, selecting two bases, one in $\text{Im } A$ and another in $\text{Im } A^+$, we can identify A^+ with its matrix representation in these bases. This matrix is square of order r and rank r , so the Jacobian c_1 of the transformation $u = A^+v$, $v \in \Gamma\Delta - \tau$, is positive. Therefore,

$$\beta(\Gamma\Delta, \tau) = c_1 \int_{\Gamma\Delta - \tau} \bar{g}_k(\|A^+v\|_2) dv = c_1 \int_{\Gamma\Delta} \bar{g}_k(\|A^+(w - \tau)\|_2) dw. \quad (3.59)$$

Secondly, $\|A^+v\|_2 = 0$ implies that $A^+v = 0$ which for $v \in \text{Im } A$ implies that $v = 0$. Thus, $\|A^+v\|_2$ is a norm on $\text{Im } A$ and by (3.14),

$$\|A^+(w - \tau)\|_2 \geq c_2\|w - \tau\|_2 \geq \max\{c_2(\|w\|_2 - \|\tau\|_2), 0\}, \quad w \in \Gamma\Delta. \quad (3.60)$$

As \bar{g} is assumed nonincreasing, \bar{g}_k is also nonincreasing, and (3.59) leads to

$$\beta(\Gamma\Delta, \tau) \leq c_1 \int_{\Gamma\Delta} \bar{g}_k(\max\{c_2(\|w\|_2 - \|\tau\|_2), 0\}) dw. \quad (3.61)$$

We bound this using polar coordinates:

$$\beta(\Gamma\Delta, \tau) \leq c_1 \int_{\Delta} \left[\int_{\|\tau\|_2}^{\infty} \bar{g}_k(c_2(\rho - \|\tau\|_2)) \rho^{r-1} d\rho \right] d\sigma = c_3 \text{mes}(\Delta). \quad (3.62)$$

Here, the quantity

$$\int_{\|\tau\|_2}^{\infty} \bar{g}_k(c_2(\rho - \|\tau\|_2)) \rho^{r-1} d\rho = \frac{1}{c_2} \int_0^{\infty} \bar{g}_k(\rho) \left(\frac{\rho}{c_2} + \|\tau\|_2 \right)^{r-1} d\rho \quad (3.63)$$

is finite because by condition (2.23), \bar{g}_k is integrable in the neighborhood of zero and $(\rho/c_2 + \|\tau\|_2)^{r-1}$ is of order ρ^{r-1} at infinity.

Step 2 (proving the main bound). The matrix A from Lemma 3.4(a) is of size $n \times (n-1)$ and A^+ is of size $(n-1) \times n$. Equation (3.50) and $R^{n-1} = \text{Im } H \oplus \text{Im}(I - H)$ imply that

$$\text{rank}(I - H) = n - 1 - r = k - 1. \quad (3.64)$$

Equation (3.48) written point-wise is $t^{(2)} = u + v$, $t^{(2)} \in T^{(2)}$, $u \in A^+(\Gamma\Delta - \tau)$, $v \in \text{Im}(I - H)$. Since u, v here are orthogonal, we have

$$\bar{g}_1(t^{(2)}) = \int_R \bar{p} \left(\left(t_1^2 + \|u\|_2^2 + \|v\|_2^2 \right)^{1/2} \right) dt_1 \equiv h(\|u\|_2, \|v\|_2), \quad (3.65)$$

where the right side defines the function h .

The transition from $t^{(2)}$ to the coordinates (u, v) is described by an orthogonal matrix. This can be shown as follows. Take in $\text{Im } H$, an orthonormal basis and identify u with the

set of its coordinates in that basis. Do the same with $\text{Im}(I - H)$ and v . Join the two bases. The equation $t^{(2)} = u + v$ induces a transformation between the original coordinates of $t^{(2)}$ and the ones in the joint basis. Since $\text{Im } H \perp \text{Im}(I - H)$, the Jacobian of the resulting transformation is unity. Hence, by (3.48), (3.64), (3.65), and (3.57),

$$\begin{aligned}\alpha(Y, \tau) &= \int_{A^+(\Gamma\Delta - \tau) + \text{Im}(I - H)} \bar{g}_1(u + v) du dv \\ &= \int_{A^+(\Gamma\Delta - \tau)} \left[\int_{R^{k-1}} h(\|u\|_2, \|v\|_2) dv \right] du \\ &= \int_{A^+(\Gamma\Delta - \tau)} \bar{g}_k(\|u\|_2) du = \beta(\Gamma\Delta, \tau) \\ &\leq c(\|\tau\|_2) \text{mes}(\Delta).\end{aligned}\tag{3.66}$$

Part (b). In the bound (see (3.59) and (3.60)),

$$\beta(Y, \tau) = c_1 \int_{Y - \tau} \bar{g}_k(\|A^+ v\|_2) dv \leq c_1 \int_{Y - \tau} \bar{g}_k(c_2 \|v\|_2) dv,\tag{3.67}$$

we have $\text{mes}(Y - \tau) = \text{mes}(Y) \rightarrow 0$, so $\beta(Y, \tau) \rightarrow 0$ uniformly in τ . As in (3.66), it follows that

$$\alpha(Y, \tau) = \int_{A^+(Y - \tau) + \text{Im}(I - H)} \bar{g}_1(u + v) du dv = \beta(Y, \tau) \rightarrow 0.\tag{3.68}$$

□

Lemma 3.6. Let f_1, p, τ, A, A^+ , and H be the same as in Lemma 3.4(b).

- (a) Suppose $Y = \Gamma\Delta$ where $\Delta \subseteq \{s \in \text{Im}(M_X - pp') : \|s\|_2 = 1\}$. If (2.25) holds, then (3.55) is true with $T^{(2)}$ from Lemma 3.4(b).
- (b) If $Y \subseteq \{s \in \text{Im}(M_X - pp') : \|s\|_2 \leq 1\}$, then $\alpha(Y, \tau) \rightarrow 0$, as $\text{mes}(Y) \rightarrow 0$, uniformly in τ .

Proof. The proof essentially follows that of Lemma 3.5(a). From (3.49), and (3.54) we see that instead of (3.58) we have $r = n - k - 1$. Instead of (3.64) we have $\text{rank}(I - H) = k$. This leads to the definition of the marginal density used in (2.25). With these changes, (3.62), and (3.66) are true and prove part (a).

With the same changes the proof of Lemma 3.5(b) goes through. □

In the course of the proof, we approximate χ_Φ by functions $\varphi_\varepsilon^\pm \in CL(R^n)$ from below and above. Then, we show that the rate of approximation improves as $\varepsilon \rightarrow 0$ and $\rho \rightarrow a$.

Step 1 (constructing the approximating functions). Denote $\rho(x, A) = \inf_{y \in A} \|x - y\|_2$ the distance from a point x to a set A . For a given $\varepsilon > 0$, set $\Psi_\varepsilon^+ = \{x \in R^n : \rho(x, \Phi) \leq \varepsilon\}$, $\Psi_\varepsilon^- = \{x \in \Phi : \rho(x, \text{bd}(\Phi)) > \varepsilon\}$.

Obviously, $\Psi_\varepsilon^- \subset \Phi \subset \Psi_\varepsilon^+$. As in [10, page 12 proof of (ii) \Rightarrow (iii)], we can construct functions φ_ε^\pm that are uniformly continuous on R^n and satisfy

$$\begin{aligned} \varphi_\varepsilon^+ &= 1 \text{ on } \Phi, & \varphi_\varepsilon^+ &= 0 \text{ outside } \Psi_\varepsilon^+, & 0 \leq \varphi_\varepsilon^+ &\leq 1 \text{ on } R^n, \\ \varphi_\varepsilon^- &= 1 \text{ on } \Psi_\varepsilon^-, & \varphi_\varepsilon^- &= 0 \text{ outside } \Phi, & 0 \leq \varphi_\varepsilon^- &\leq 1 \text{ on } R^n. \end{aligned} \quad (3.69)$$

Step 2 (reducing the convergence statement to a special case). Denote

$$I(\varphi) = \int_{R^{n-1}} (MT_1\varphi)(t^{(2)}) \bar{g}_1(t^{(2)}) dt^{(2)}. \quad (3.70)$$

Here, we show that

$$\lim_{\varepsilon \rightarrow 0} I(\varphi_\varepsilon^\pm) = I(\chi_\Phi) \quad (3.71)$$

implies (2.24). From the inequality $\varphi_\varepsilon^- \leq \chi_\Phi \leq \varphi_\varepsilon^+$ we have

$$J_\rho(\varphi_\varepsilon^-) \leq J_\rho(\chi_\Phi) \leq J_\rho(\varphi_\varepsilon^+). \quad (3.72)$$

By Theorem 2.7,

$$\lim_{\rho \rightarrow a} J_\rho(\varphi_\varepsilon^\pm) = I(\varphi_\varepsilon^\pm), \quad \text{for any } \varepsilon > 0. \quad (3.73)$$

By (3.71), for any $\delta > 0$, there exists $\varepsilon(\delta) > 0$ such that

$$I(\varphi_\varepsilon^+) - \delta \leq I(\chi_\Phi) \leq I(\varphi_\varepsilon^-) + \delta, \quad 0 < \varepsilon \leq \varepsilon(\delta). \quad (3.74)$$

By (3.73), for the same δ , there is $\rho(\delta) \in (0, a)$ such that

$$|J_\rho(\varphi_\varepsilon^\pm) - I(\varphi_\varepsilon^\pm)| \leq \delta, \quad \rho(\delta) \leq \rho < a. \quad (3.75)$$

Equations (3.72), (3.74), and (3.75) imply that

$$\begin{aligned} I(\chi_\Phi) - 2\delta &\leq I(\varphi_\varepsilon^-) - \delta \leq J_\rho(\varphi_\varepsilon^-) \leq J_\rho(\chi_\Phi) \leq J_\rho(\varphi_\varepsilon^+) \\ &\leq I(\varphi_\varepsilon^+) + \delta \leq I(\chi_\Phi) + 2\delta, \quad \rho(\delta) \leq \rho < a. \end{aligned} \quad (3.76)$$

Since $\delta > 0$ is arbitrary, this proves (2.24).

Step 3 (further reduction to an indicator function). Let us call $B(\varepsilon) = \Psi_\varepsilon^+ \setminus \Psi_\varepsilon^-$ a band. Here, we show that (3.71) will follow if we prove that

$$\lim_{\varepsilon \rightarrow 0} I(\chi_{B(\varepsilon)}) = 0. \quad (3.77)$$

Note the identity

$$\begin{aligned} I(\chi_{B(\varepsilon)}) &= I(\chi_{\Psi_\varepsilon^+}) - I(\chi_{\Psi_\varepsilon^-}) \\ &= [I(\chi_{\Psi_\varepsilon^+}) - I(\chi_\Phi)] + [I(\chi_\Phi) - I(\chi_{\Psi_\varepsilon^-})]. \end{aligned} \quad (3.78)$$

Since each of the square brackets on the right is nonnegative, (3.77) gives $\lim_{\varepsilon \rightarrow 0} I(\chi_{\Psi_\varepsilon^+}) = I(\chi_\Phi)$. Now, it suffices to take into account the inequality $I(\chi_{\Psi_\varepsilon^-}) \leq I(\varphi_\varepsilon^-) \leq I(\chi_\Phi) \leq I(\varphi_\varepsilon^+) \leq I(\chi_{\Psi_\varepsilon^+})$ to see that (3.71) follows.

Step 4 (geometric description of the task). Denote $l(s, i) = \{st + i : t \in R\}$ a *straight line* with a *slope* $s \in R^n$ and *intercept* $i \in R^n$ and put

$$b_\varepsilon(t^{(2)}) = (MT_1 \chi_{B(\varepsilon)})(t^{(2)}) = \lim_{r \rightarrow \infty} \frac{1}{2r} \int_{-r}^r \chi_{B(\varepsilon)}(f_1 t_1 + z(t^{(2)})) dt_1. \quad (3.79)$$

This expression shows that we need to study the set $\{t_1 \in R : f_1 t_1 + z(t^{(2)}) \in B(\varepsilon)\}$ of points common to the straight line $l(f_1, z(t^{(2)}))$ and the band $B(\varepsilon)$. The cylindrical nature of the sets considered implies that the band admits an equivalent definition as a cylinder with the base

$$M_X B(\varepsilon) = \{x \in \text{Im } M_X : \rho(x, \text{bd}(M_X \Phi)) \leq \varepsilon\} \quad (3.80)$$

and element $\text{Im } X$. From this definition, we see that the condition $f_1 t_1 + z(t^{(2)}) \in B(\varepsilon)$ is equivalent to

$$(M_X f_1) t_1 + M_X z(t^{(2)}) \in M_X B(\varepsilon). \quad (3.81)$$

We have to sort out points $t^{(2)} \in R^{n-1}$ into (a) those for which the set of points t_1 satisfying (3.81) is so small that (3.79) is zero and (b) that set is large enough for (3.79) to be positive.

Step 5 (case $f_1 \in \text{Im } X$). In this case, $M_X f_1 = 0$, and (3.81) reduces to

$$M_X P_2(a) \Lambda_2^{-1}(a) t^{(2)} + M_X Q X \beta \in M_X B(\varepsilon). \quad (3.82)$$

This condition does not contain t_1 . Therefore, if it is not satisfied, the integrand in (3.79) is identically zero and $b_\varepsilon(t^{(2)}) = 0$ for such $t^{(2)}$. If it is satisfied, the integrand in (3.79) is identically 1 and $b_\varepsilon(t^{(2)}) = 1$ for such $t^{(2)}$. Hence,

$$I(\chi_{B(\varepsilon)}) = \int_{(3.82) \text{ is true}} \bar{g}_1(t^{(2)}) dt^{(2)}. \quad (3.83)$$

The set of $t^{(2)}$ satisfying (3.82) has an infinite Lebesgue measure which precludes us from application of the continuity property of the Lebesgue measure (for an integrable $h, \int_A h(t) dt \rightarrow 0$ as $\text{mes}(A) \rightarrow 0$).

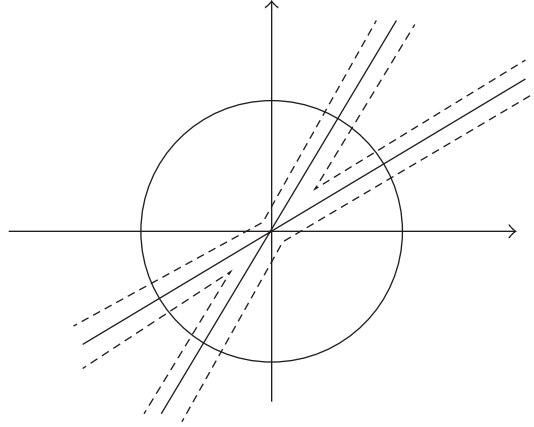


Figure 1: $M_X\Phi$ consists of two cones bounded by solid lines. $M_X B(\varepsilon)$ is bounded by dashed lines.

As one can see from Figure 1, $M_X B(\varepsilon)$ is not a cone-like set. We divide it into two parts: $\alpha_\varepsilon^{\text{in}} = \{s \in M_X B(\varepsilon) : \|s\|_2 \leq 1\}$ includes the points inside the unit ball and on the unit sphere and $\alpha_\varepsilon^{\text{out}} = \{s \in M_X B(\varepsilon) : \|s\|_2 > 1\}$ includes the points of $M_X B(\varepsilon)$ outside the unit ball. Put $\Delta_\varepsilon = \{s \in \text{Im } M_X : \|s\|_2 = 1, \rho(s, \text{bd}(M_X\Phi)) \leq \varepsilon\}$ and let us show that

$$\alpha_\varepsilon^{\text{out}} \subset \Gamma \Delta_\varepsilon. \quad (3.84)$$

Suppose $x \in \alpha_\varepsilon^{\text{out}}$. Then, $x \in \text{Im } M_X$, $\|x\|_2 > 1$ and by (3.80), there exists a sequence $\{y^{(k)}\} \subset \text{bd}(M_X\Phi)$ such that $\limsup \|x - y^{(k)}\|_2 \leq \varepsilon$. The sequence $\{y^{(k)}\}$ is bounded and therefore precompact, while $\text{bd}(M_X\Phi) = M_X(\text{cl}(\Phi) \setminus \Phi)$ is closed by construction. Hence, we can assume that $y^{(k)} \rightarrow y \in \text{bd}(M_X\Phi)$ (passing to a subsequence, if necessary). Then, $\|x - y\|_2 \leq \varepsilon$. This implies $\|x/\|x\|_2 - y/\|x\|_2\|_2 \leq \varepsilon/\|x\|_2 < \varepsilon$. Let $s = x/\|x\|_2$. Since $\text{bd}(M_X\Phi)$ is cone-like, it follows that $y/\|x\|_2 \in \text{bd}(M_X\Phi)$ and $\rho(s, \text{bd}(M_X\Phi)) < \varepsilon$. This proves that $s \in \Delta_\varepsilon$, $x = \|x\|_2 s \in \Gamma \Delta_\varepsilon$ and (3.84) is true.

From (3.84), we conclude that

$$M_X B(\varepsilon) \subset \alpha_\varepsilon^{\text{in}} \cup \Gamma \Delta_\varepsilon. \quad (3.85)$$

This means that the set of solutions of (3.82) is contained in the union of two sets: $T_1^{(2)} = \{t^{(2)} \in R^{n-1} : At^{(2)} + \tau \in \Gamma \Delta_\varepsilon\}$ and $T_2^{(2)} = \{t^{(2)} \in R^{n-1} : At^{(2)} + \tau \in \alpha_\varepsilon^{\text{in}}\}$ where $A = M_X P_2(a) \Lambda_2^{-1}(a)$ and $\tau = M_X Q X \beta$.

By Lemma 3.4(a), $T_1^{(2)} = A^+(\Gamma \Delta_\varepsilon - \tau) + \text{Im}(I - H)$. By Lemma 3.5(a)

$$\int_{T_1^{(2)}} \bar{g}_1(t^{(2)}) dt^{(2)} \leq c(\|\tau\|_2) \text{mes}(\Delta_\varepsilon) \rightarrow 0, \quad \varepsilon \rightarrow 0. \quad (3.86)$$

By Lemma 3.4(a), $T_2^{(2)} = A^+(\alpha_\varepsilon^{\text{in}} - \tau) + \text{Im}(I - H)$. By Lemma 3.5(b)

$$\int_{T_2^{(2)}} \bar{g}_1(t^{(2)}) dt^{(2)} \rightarrow 0 \quad \text{as } \text{mes}(\alpha_\varepsilon^{\text{in}}) \rightarrow 0. \quad (3.87)$$

Equations (3.86) and (3.87) prove that $I(\chi_{B(\varepsilon)}) \rightarrow 0$, $\varepsilon \rightarrow 0$.

Step 6 (case $f_1 \notin \text{Im } X$). Equation (2.5) implies that if the straight line $l(M_X f_1, M_X z(t^{(2)}))$ has a bounded intersection with the band $M_X B(\varepsilon)$, then $(MT_1 \chi_{B(\varepsilon)})(t^{(2)}) = 0$. Therefore, we are interested in the set of slopes:

$$\text{Sl} = \left\{ M_X f_1 : l(M_X f_1, M_X z(t^{(2)})) \cap M_X B(\varepsilon) \text{ is unbounded} \right\}. \quad (3.88)$$

Let us prove that

$$M_X f_1 \in \text{Sl} \implies M_X f_1 / \|M_X f_1\|_2 \in \text{bd}(S). \quad (3.89)$$

Suppose $M_X f_1 \in \text{Sl}$. Then, there exists a sequence of points $p^{(k)}$ which lie on the straight line, $p^{(k)} = (M_X f_1)t_1^{(k)} + M_X z(t^{(2)})$, where $|t_1^{(k)}| \rightarrow \infty$, and which are 2ε -close to $\text{bd}(M_X \Phi)$: there exists a sequence $y^{(k)} \in \text{bd}(M_X \Phi)$ such that

$$\|p^{(k)} - y^{(k)}\|_2 \leq 2\varepsilon, \quad \|y^{(k)}\|_2 \rightarrow \infty. \quad (3.90)$$

Equation (2.21) implies that $M_X \Phi = \Gamma S$ and $\text{bd}(M_X \Phi) = \Gamma \text{bd}(S) \cup \{0\}$. Therefore $s^{(k)} \equiv y^{(k)} / \|y^{(k)}\|_2 \in \text{bd}(S)$. Because $\text{bd}(S)$ is compact, we can assume that $s^{(k)} \rightarrow s \in \text{bd}(S)$. From (3.90), we have $\|(M_X f_1)t_1^{(k)} + M_X z(t^{(2)}) - \|y^{(k)}\|_2 s^{(k)}\|_2 \leq 2\varepsilon$ or, dividing both sides by $\|y^{(k)}\|_2$, $\|(M_X f_1)t_1^{(k)} / \|y^{(k)}\|_2 - s^{(k)}\|_2 \rightarrow 0$. This shows that the sequence $\{t_1^{(k)} / \|y^{(k)}\|_2\}$ is bounded and can be assumed convergent to a number δ . Letting $k \rightarrow \infty$, we get $(M_X f_1)\delta = s \in \text{bd}(S)$, and (3.89) is true.

Fix $M_X f_1 \in \text{Sl}$ and denote $p = M_X f_1 / \|M_X f_1\|_2 \in \text{bd}(S) \subset \text{Im } M_X$. The matrix pp' projects $\text{Im } M_X$ onto the straight line $l(M_X f_1, 0)$. The projector $M_X - pp'$ projects $\text{Im } M_X$ onto $\text{Im}(M_X - pp')$. Condition (3.81) implies two equations in projections:

$$\begin{aligned} pp'[(M_X f_1)t_1 + M_X z(t^{(2)})] &\in pp' M_X B(\varepsilon), \\ (M_X - pp')[(M_X f_1)t_1 + M_X z(t^{(2)})] &\in (M_X - pp') M_X B(\varepsilon). \end{aligned} \quad (3.91)$$

The first one can be safely omitted because the shift $pp' M_X z(t^{(2)})$ along a line does not change the property of the line to have an unbounded intersection with $M_X B(\varepsilon)$. Denoting $A = (M_X - pp')P_2(a)\Lambda_2^{-1}(a)$, $\tau = (M_X - pp')QX\beta$, and using (3.52), we write the second equation as

$$At^{(2)} + \tau \in (M_X - pp')B(\varepsilon). \quad (3.92)$$

The set of points $t^{(2)}$ such that (3.81) holds for an unbounded set of values t_1 is clearly contained in the set $T^{(2)}$ of points $t^{(2)}$ that satisfy (3.92). For $t^{(2)} \in T^{(2)}$ we use an obvious bound $b_\varepsilon(t^{(2)}) \leq 1$ (see (3.79)), and, therefore, $I(\chi_{B(\varepsilon)}) \leq \int_{(3.92) \text{ is true}} \bar{g}_1(t^{(2)}) dt^{(2)}$. The rest of the proof closely follows that for the case $f_1 \in \text{Im } X$. An inclusion of type (3.85) holds with M_X replaced by $M_X - pp'$ in all definitions. The reference to Lemma 3.4(a) is replaced by Lemma 3.4(b). Lemma 3.6 is used instead of Lemma 3.5. The conclusion is that $I(\chi_{B(\varepsilon)}) \rightarrow 0$.

We shall use several times the fact that for the density \bar{g}_1 in representation (2.24) will, the conditions $\int_A \bar{g}_1(t^{(2)}) dt^{(2)} > 0$ and $\text{mes}(A) > 0$ are equivalent. From (2.24), we see that the task is to calculate the function (3.79) with Φ in place of $B(\varepsilon)$. Because of (2.21), the condition that $f_1 t_1 + z(t^{(2)}) \in \Phi$ is equivalent to the condition in projections:

$$(M_X f_1) t_1 + M_X z(t^{(2)}) \in \Gamma S. \quad (3.93)$$

The magnitude of $\beta(\Phi)$ depends on the set

$$A(t^{(2)}) = \{t_1 \in R : f_1 t_1 + z(t^{(2)}) \in \Phi\}. \quad (3.94)$$

The properties of $A(t^{(2)})$ depend on those of $\text{mes}\{t^{(2)} \in R^{n-1} : M_X z(t^{(2)}) \in G\}$ for certain sets $G \subseteq \text{Im } M_X$, as we will see.

Lemma 3.7. *Suppose $\text{Im } Q \subseteq \text{Im } X$ (Q is the projector from Theorem 2.7). Then, $M_X z(t^{(2)}) = 0$ for all $t^{(2)} \in R^{n-1}$.*

Let $\text{Im } Q \not\subseteq \text{Im } X$. Then, $\text{mes}\{t^{(2)} : M_X z(t^{(2)}) = 0\} = 0$.

Proof. The assumption $\text{Im } Q \subseteq \text{Im } X$ implies that $M_X f_i = 0$ for $i = 2, \dots, n$, which can be written as $M_X P_2(a) = 0$. Hence, $M_X z(t^{(2)}) = M_X P_2(a) [\Lambda_2^{-1}(a) t^{(2)} + P_2'(a) X \beta] = 0$ for any $t^{(2)} \in R^{n-1}$.

If $\text{Im } Q \not\subseteq \text{Im } X$, then $f_i \notin \text{Im } X$ and $M_X f_i \neq 0$ for at least one i . The matrix $M_X P_2(a)$ is of size $n \times (n-1)$, and of rank at least 1. From the equation $\dim N(M_X P_2(a)) + \text{rank}(M_X P_2(a)) = n - 1$ we see that the null space $N(M_X P_2(a))$ is of dimension at most $n - 2$. The set of $t^{(2)}$ satisfying $M_X z(t^{(2)}) = M_X P_2(a) \Lambda_2^{-1}(a) t^{(2)} + M_X Q X \beta = 0$ is a partial solution $t_0^{(2)}$ plus an arbitrary element from the null space $N(M_X P_2(a))$. Thus the set of all solutions is a hyperplane in R^{n-1} of dimension at most $n - 2$ and has $(n - 1)$ -dimensional Lebesgue measure zero. \square

Lemma 3.8. *Suppose we are given a set $G \subseteq R^{n-1}$ and for almost each $t^{(2)} \in G$, the set $A(t^{(2)}) \subseteq R$ is defined by (3.94).*

(a) *If $A(t^{(2)})$ is bounded for almost all $t^{(2)} \in G$ and $\text{mes}(G) > 0$, then $\beta(\Phi) < 1$. In particular, if $G = R^{n-1}$, then $\beta(\Phi) = 0$.*

(b) *If*

$$R \setminus A(t^{(2)}) \text{ is bounded for almost all } t^{(2)} \in G \quad (3.95)$$

and $\text{mes}(G) > 0$, then $\beta(\Phi) > 0$. In particular, if $G = R^{n-1}$, then $\beta(\Phi) = 1$.

(c) Suppose that $G = \mathbb{R}^{n-1}$ and for almost all $t^{(2)} \in G$,

$$\text{the set } A(t^{(2)}) \text{ is a half-axis} \quad (3.96)$$

(i.e., either $A(t^{(2)}) = (a(t^{(2)}), \infty)$ or $A(t^{(2)}) = (-\infty, a(t^{(2)}))$ with some $a(t^{(2)}) \in \mathbb{R}$). Then, $\beta(\Phi) = 1/2$.

Proof. If $A(t^{(2)})$ is bounded, then $\chi_\Phi(f_1 t_1 + z(t^{(2)})) = 0$ outside a bounded set, and by (2.5),

$$M\chi_{A(t^{(2)})} = (MT_1\chi_\Phi)(t^{(2)}) = 0 \quad (3.97)$$

for almost all $t^{(2)} \in G$. Since $0 \leq M\chi_{A(t^{(2)})} \leq 1$ outside G , we have

$$\beta(\Phi) \leq \int_{\mathbb{R}^{n-1} \setminus G} \bar{g}_1(t^{(2)}) dt^{(2)} < 1. \quad (3.98)$$

Obviously, the integral is null if $G = \mathbb{R}^{n-1}$.

By (3.97), condition (3.95) implies that

$$M\chi_{A(t^{(2)})} = M(1 - \chi_{\mathbb{R} \setminus A(t^{(2)})}) = 1 \quad \text{a.e. in } G. \quad (3.99)$$

Since $\text{mes}(G) > 0$, we have

$$\beta(\Phi) \geq \int_G \bar{g}_1(t^{(2)}) dt^{(2)} > 0. \quad (3.100)$$

In case $G = \mathbb{R}^{n-1}$, the last integral is 1.

Consider, for example, the case of the right half-axis:

$$M\chi_{A(t^{(2)})} = \lim_{r \rightarrow \infty} \frac{1}{2r} \left(\int_{a(t^{(2)})}^0 dt + \int_0^r dt \right) = \frac{1}{2}. \quad (3.101)$$

Hence, $\beta(\Phi) = (1/2) \int_{\mathbb{R}^{n-1}} \bar{g}_1(t^{(2)}) dt^{(2)} = 1/2$. □

Case 1. Let $f_1 \in \Phi$. Then, $f_1 \notin \text{Im } X$ and $M_X f_1 \neq 0$. Define $p = M_X f_1 / \|M_X f_1\|_2$. Obviously, $M_X f_1 \in \Gamma S$ and $p \in S$. In the equation

$$(M_X f_1)t_1 + M_X z(t^{(2)}) = \|M_X f_1\|_2 t_1 \left[p + \frac{M_X z(t^{(2)})}{\|M_X f_1\|_2 t_1} \right], \quad (3.102)$$

for all large $|t_1|$, the term in the square brackets is close to p and, because $M_X \Phi$ is open in $\text{Im } M_X$, this term belongs to $M_X \Phi$. Since $M_X \Phi$ is cone-like, the left side of (3.102) belongs

to $M_X\Phi$ for all large $|t_1|$. Hence, the set $A(t^{(2)})$ contains the neighborhood of infinity (more precisely, all t_1 with $|t_1| > N$ where N depends on $t^{(2)}$). By Lemma 3.8(b), where $G = R^{n-1}$, then $\beta(\Phi) = 1$.

Case 2. Suppose $f_1 \in \text{bd}(\Phi)$.

Subcase 2.1. Let $f_1 \in \text{Im } X$ or, equivalently, $M_X f_1 = 0$. Condition (3.93) reduces to

$$M_X z(t^{(2)}) = M_X P_2(a) \Lambda_2^{-1}(a) t^{(2)} + M_X Q X \beta \in \Gamma S. \quad (3.103)$$

By Lemma 3.4(a), the set of solutions of (3.103) is $T^{(2)} = A^+(\Gamma S - \tau) + \text{Im}(I - H)$ where A^+ is the Moore-Penrose inverse of $A = M_X P_2(a) \Lambda_2^{-1}(a)$, $\tau = M_X Q X \beta$, and H is the projector from Lemma 3.2. Using Lemma 3.8(b), (3.66), and (3.59) we get

$$\beta(\Phi) = \int_{T^{(2)}} \bar{g}_1(t^{(2)}) dt^{(2)} = \int_{A^+(\Gamma S - \tau)} \bar{g}_k(\|u\|_2) du. \quad (3.104)$$

The density $\bar{g}_k(\|u\|_2)$ is spherically symmetric. It is positive everywhere because

$$\bar{g}_k(t) = \int_{R^k} \bar{p}\left((t^2 + \|v\|_2^2)^{1/2}\right) dv = 0 \quad (3.105)$$

would mean $\bar{p}((t^2 + t_1^2)^{1/2}) = 0$ for any $t_1 \geq 0$, that is, $\bar{p}(t_1) = 0$ for any $t_1 \geq t$, which is impossible by assumption. A^+ is one-to-one on $\text{Im } A = \text{Im } M_X$, so $A^+(\Gamma S - \tau)$ is a nonempty proper subset of $\text{Im } A^+$. Therefore, $0 < \beta(\Phi) < 1$.

Subcase 2.2. Suppose $f_1 \notin \text{Im } X$. As in Case 1, p is defined. Let us show that it satisfies

$$p' W p = c. \quad (3.106)$$

Since f_1 is a limit point of Φ , there is a sequence $x_n \in \Phi$ such that $x_n \rightarrow f_1$. Hence, $M_X x_n \rightarrow M_X f_1$ where $\|M_X f_1\|_2 \neq 0$ because $f_1 \notin \text{Im } X$. From (2.21), it follows that $M_X x_n \in \Gamma S$, $z_n \equiv M_X x_n / \|M_X x_n\|_2 \in S$ and $z_n' W z_n > c$. Letting $n \rightarrow \infty$ and using $z_n \rightarrow p$, we get $p' W p \geq c$. Since $f_1 \notin \Phi$, we have $M_X f_1 \notin M_X \Phi$ and $p \notin S$. Thus, $p' W p > c$ is impossible and (3.106) is true.

In the three examples in the following we assume that $n > k + 1$. Then, $\text{Im } Q \not\subseteq \text{Im } X$, and by Lemma 3.7(b),

$$M_X z(t^{(2)}) \neq 0, \text{ for almost all } t^{(2)} \in R^{n-1}. \quad (3.107)$$

Sub-sub-case (a) (Example of $\beta(\Phi) = 1$). Let W be symmetric. Assumptions (1.13) imply that $\text{var}(u) = \sigma_\varepsilon^2 (I - \rho W)^{-2}$ whenever $\det(I - \rho W) \neq 0$, so $\Sigma^{-1}(\rho) = (I - \rho W)^2$. If μ_i denote eigenvalues of W , then the eigenvalues of $I - \rho W$ are $v_i(\rho) = 1 - \rho \mu_i$. Let the constant c in

the definition of the critical region Φ be positive. Put $a = 1/c$ and with some integer i_0 and real numbers μ_-, μ_+ such that $1 < i_0 < n$, $\mu_- < 1/a < \mu_+$ define μ_i by

$$\mu_1 = \cdots = \mu_{i_0-1} = \mu_- < \mu_{i_0} = \frac{1}{a} < \mu_{i_0+1} = \cdots = \mu_n = \mu_+. \quad (3.108)$$

Then, for $0 < \rho \leq a$,

$$\begin{aligned} v_1(\rho) = \cdots = v_{i_0-1}(\rho) &= 1 - \rho\mu_- > v_{i_0}(\rho) = 1 - \frac{\rho}{a} \\ &> v_{i_0+1}(\rho) = \cdots = v_n(\rho) = 1 - \rho\mu_+, \end{aligned} \quad (3.109)$$

and $\Sigma^{-1/2}(a)$ has eigenvalues

$$\begin{aligned} v_1(a) = \cdots = v_{i_0-1}(a) &= 1 - a\mu_- > v_{i_0}(a) = 0 \\ &< |v_{i_0+1}(a)| = \cdots = |v_n(a)| = |1 - a\mu_+|. \end{aligned} \quad (3.110)$$

Hence, in the notation of Condition 3, $\lambda_1(\rho) = v_{i_0}(\rho)$ and $\lambda_2(\rho), \dots, \lambda_n(\rho)$ are obtained from $|v_i(\rho)|$, $i \neq i_0$, by reordering. For example, if $1 - a\mu_- < |1 - a\mu_+|$, then $\lambda_i(\rho) = v_{i-1}(\rho)$, $i = 2, \dots, i_0$; $\lambda_i(\rho) = |v_i(\rho)|$, $i = i_0 + 1, \dots, n$. Assuming this ordering, denote the corresponding eigenvectors of $\Sigma^{-1/2}(a)$ by $f_1, f_2^-, \dots, f_{i_0}^-, f_{i_0+1}^+, \dots, f_n^+$ where the minuses and pluses remind us that the corresponding eigenvalues depend on μ_- and μ_+ , respectively.

Let $k = i_0 - 1$ and suppose that

$$f_2^-, \dots, f_{i_0}^- \text{ span Im } X, \quad f_1, f_{i_0+1}^+, \dots, f_n^+ \text{ span Im } M_X. \quad (3.111)$$

Denote

$$x = (M_X f_1) t_1 + M_X z(t^{(2)}) \in \text{Im } M_X. \quad (3.112)$$

Because of (3.111), here $M_X f_1 = f_1$. Besides, since $z(t^{(2)})$ is orthogonal to f_1 , the decomposition

$$M_X z(t^{(2)}) = \sum_{i=i_0+1}^n c_i f_i^+ \quad (3.113)$$

does not contain f_1 . Hence, (3.112) becomes

$$x = t_1 f_1 + \sum_{i=i_0+1}^n c_i f_i^+. \quad (3.114)$$

Remembering that the eigenvectors of $\Sigma^{-1/2}(a)$ are just reordered eigenvectors of W , we get

$$\begin{aligned} Wx &= \frac{1}{a}t_1f_1 + \mu_+ \sum_{i=i_0+1}^n c_i f_i^+, \\ x'Wx &= \frac{1}{a}t_1^2 + \mu_+ \sum_{i=i_0+1}^n c_i^2 = ct_1^2 + \mu_+ \sum_{i=i_0+1}^n c_i^2, \end{aligned} \quad (3.115)$$

for any x of form (3.112). From (3.107) and (3.113) we see that here

$$\sum_{i=i_0+1}^n c_i^2 \neq 0 \quad \text{a.e. in } R^{n-1}. \quad (3.116)$$

Hence, normalizing x , $s = x/\|x\|_2$, and assuming that $\mu_+ > c$, we get from (3.115)

$$s'Ws = \frac{ct_1^2 + \mu_+ \sum_{i=i_0+1}^n c_i^2}{t_1^2 + \sum_{i=i_0+1}^n c_i^2} > c, \quad (3.117)$$

for all $t_1 \in R$ and almost all $t^{(2)} \in R^{n-1}$. This shows that the set (3.94) coincides with R for almost all $t^{(2)} \in R^{n-1}$ and by Lemma 3.8(b) $\beta(\Phi) = 1$.

Sub-sub-case (b) Example of $\beta(\Phi) = 0$. We change (3.111) so as to obtain in (3.115) μ_- in place of μ_+ . Let $k = n - i_0$ and suppose that

$$f_{i_0+1}^+, \dots, f_n^+ \text{ span Im } X, \quad f_1, f_2^-, \dots, f_{i_0}^- \text{ span Im } M_X. \quad (3.118)$$

Instead of (3.113), we get $M_X z(t^{(2)}) = \sum_{i=2}^{i_0} c_i f_i^-$ which leads to an analog of (3.114): $x = t_1 f_1 + \sum_{i=2}^{i_0} c_i f_i^-$. Equation (3.115) is replaced by

$$x'Wx = ct_1^2 + \mu_- \sum_{i=2}^{i_0} c_i^2, \quad (3.119)$$

where $\sum_{i=2}^{i_0} c_i^2 \neq 0$ a.e. in R^{n-1} . The result is that

$$s'Ws = \frac{ct_1^2 + \mu_- \sum_{i=2}^{i_0} c_i^2}{t_1^2 + \sum_{i=2}^{i_0} c_i^2} < c \quad (3.120)$$

for all $t_1 \in R$ and almost all $t^{(2)} \in R^{n-1}$. Hence, the set (3.94) is empty for almost all $t^{(2)} \in R^{n-1}$ and by Lemma 3.8(a), $\beta(\Phi) = 0$.

Sub-sub-case (c) Example of $\beta(\Phi) = 1/2$. $M_X z(t^{(2)})$ runs over $\text{Im}(M_X P_2(a))$. If $k > 1$, then $n-1 > n-k$, and we can choose the eigenvectors f_2, \dots, f_n of $\Sigma^{-1/2}(a)$ so that $\text{Im } Q \supset \text{Im } M_X$, and then $M_X z(t^{(2)})$ runs over $\text{Im } M_X = \text{Im}(M_X P_2(a))$. Let ΓS be a union of two cones, as in Figure 2. The line $l(M_X f_1, 0)$ lies on the boundary of the cones. In Figure 2, $L_a = l(M_X f_1, 0)$

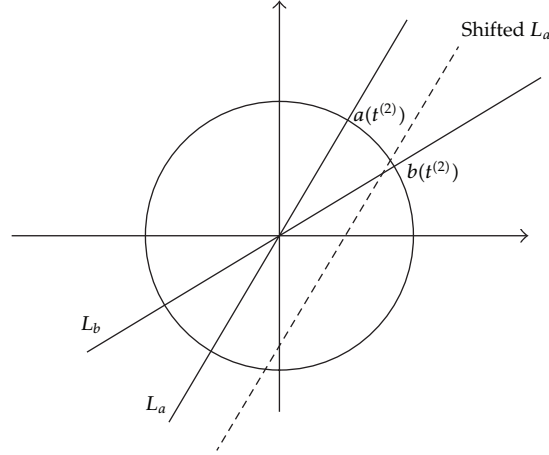


Figure 2: The base of Φ contains the area between two solid lines L_a and L_b . The dashed line is $l(M_X f_1, M_X z(t^{(2)})) = L_a + M_X z(t^{(2)})$.

and $L_b = l(b(t^{(2)}), 0)$. For any nonzero shift $M_X z(t^{(2)})$, half of the shifted line is contained in $M_X \Phi$ because L_a and L_b are not parallel. Here, “half” means “either for all large positive t_1 or for all large negative t_1 .” By Lemma 3.8(c), $\beta(\Phi) = 1/2$.

Case 3. Suppose $f_1 \in \Phi_-$. Then, as in Case 1, $M_X f_1 \neq 0$ and p is defined. It satisfies $p'Wp < c$. The term in the square brackets in (3.102) for all large $|t_1|$ belongs to an open set $M_X \Phi_-$. Since this set is cone-like, the left side of (3.102) belongs to $M_X \Phi_-$ for all large $|t_1|$. The set $A(t^{(2)})$ is bounded for all $t^{(2)} \in R^{n-1}$, and by Lemma 3.8(a), $\beta(\Phi) = 0$.

Step 1 (the $\varepsilon(\rho)$ -maximizing set is an ellipsoid). For $y \in M_{\varepsilon(\rho)}$ by orthogonality of P and Condition 7, we have

$$f_\rho(y) = \prod_{j=1}^n \lambda_j(\rho) p(\|\Lambda(\rho) P'(\rho)(y - X\beta)\|_2) \geq (1 - \varepsilon(\rho)) \prod_{j=1}^n \lambda_j(\rho) m. \quad (3.121)$$

Since p^{-1} is decreasing, this is equivalent to $\|\Lambda(\rho) P'(\rho)(y - X\beta)\|_2 \leq p^{-1}((1 - \varepsilon(\rho))m)$. Denoting $h(\rho) = p^{-1}((1 - \varepsilon(\rho))m)$, $z = P'(\rho)(y - X\beta)$, we rewrite the resulting inequality as

$$\sum_{j=1}^d \frac{z_j^2}{[h(\rho)/\lambda_j(\rho)]^2} + \sum_{j=d+1}^n \frac{z_j^2}{[h(\rho)/\lambda_j(\rho)]^2} \leq 1. \quad (3.122)$$

In z coordinates, this is an ellipsoid with half-axes $h(\rho)/\lambda_j(\rho)$.

Step 2 (characterizing the rate of decay of $h(\rho)$). First we determine the rate with which $p^{-1}(x)$ approaches 0 when $x \rightarrow m$. To approximately invert p , in the equation $x = p(r)$, we use

representation (2.30): $x = m(1 - c_p r^{\alpha_p})(1 + o(r^{\alpha_\varepsilon}))$. This is equivalent to $(x/m)(1 + o(r^{\alpha_\varepsilon})) = 1 - c_p r^{\alpha_p}$. Solving this equation for the r involved in the right-hand side produces

$$p^{-1}(x) = r = \left\{ \frac{1}{c_p} \left[1 - \frac{x}{m} (1 + o(r^{\alpha_\varepsilon})) \right] \right\}^{1/\alpha_p}, \quad r \rightarrow 0. \quad (3.123)$$

Since $\rho \rightarrow a$, we can put $r = a - \rho$ in the last equation to get

$$p^{-1}(x) = \left\{ \frac{1}{c_p} \left[1 - \frac{x}{m} (1 + o((a - \rho)^{\alpha_\varepsilon})) \right] \right\}^{1/\alpha_p}, \quad \rho \rightarrow a. \quad (3.124)$$

To obtain $h(\rho)$, let $x = (1 - \varepsilon(\rho))m$. Then, by (2.29),

$$\begin{aligned} 1 - \frac{x}{m} (1 + o((a - \rho)^{\alpha_\varepsilon})) &= 1 - (1 - \varepsilon(\rho)) (1 + o((a - \rho)^{\alpha_\varepsilon})) \\ &= 1 - 1 + o((a - \rho)^{\alpha_\varepsilon}) + \varepsilon(\rho) (1 + o((a - \rho)^{\alpha_\varepsilon})) \\ &= \varepsilon(\rho) (1 + o(1)). \end{aligned} \quad (3.125)$$

Combining this with (3.124) and (2.29), we have

$$\begin{aligned} h(\rho) &= p^{-1}(x) = \left\{ \frac{c_\varepsilon}{c_p} (a - \rho)^{\alpha_\varepsilon} (1 + o(1)) \right\}^{1/\alpha_p} \\ &= \left(\frac{c_\varepsilon}{c_p} \right)^{1/\alpha_p} (a - \rho)^{\alpha_\varepsilon/\alpha_p} (1 + o(1)). \end{aligned} \quad (3.126)$$

Step 3 (behavior of half-axes). We intend to show that for $j = d + 1, \dots, n$ the half-axes tend to zero, so that in the limit

$$z_{d+1} = \dots = z_n = 0, \quad (3.127)$$

whereas for $j = 1, \dots, d$, they tend to ∞ , and in the limit, z_1, \dots, z_d can be any real numbers.

For $j = d + 1, \dots, n$ from (2.28), we see that the condition $h(\rho)/\lambda_j(\rho) \rightarrow 0$ is equivalent to $h(\rho) \rightarrow 0$. Because of (2.31), this condition is the same as $(1 - \varepsilon(\rho))m \rightarrow m$ which holds by (2.29).

For the remaining half-axes, we get by (2.27), (3.126), and (2.32)

$$\frac{h(\rho)}{\lambda_j(\rho)} = \frac{1}{c_\lambda} \left(\frac{c_\varepsilon}{c_p} \right)^{1/\alpha_p} (a - \rho)^{\alpha_\varepsilon/\alpha_p - \alpha_\lambda} (1 + o(1)) \rightarrow \infty. \quad (3.128)$$

We have proved our statement on convergence of the half-axes. According to (2.26), (3.127) rewrites as $(P'(a)(y - X\beta))_j = 0$, $j = d + 1, \dots, n$. Since the rows of $P'(a)$ are eigenvectors

of $A(a)$, the last equation means that $y - X\beta$ is orthogonal to eigenvectors numbered $j = d + 1, \dots, n$. Hence, $y - X\beta$ belongs to the null space of $A(a)$, and the set of all such y 's is $X\beta + N(A(a))$.

In this proof, the set $X\beta + N(A(a))$ is obtained as the limit of the ridge of f_ρ . In general, this ridge can have any shape and the limit will not be $X\beta + N(A(a))$.

4. Summary

Theorem 2.2 is the key result in this paper. We tried to obtain it under minimal conditions on g . It would be interesting to relax our conditions. In particular, densities that are not spherically symmetric are of interest in applications. Note that when the dependence on λ is more complex than in $g_\lambda(x) = \lambda g(\lambda x)$, one needs very different techniques and more detailed assumptions about g to obtain an analog of Theorem 2.2. Such results can be found in [15]. Another promising direction of research is to consider tests other than the Cliff-Ord one.

Acknowledgments

The author would like to thank F. Martellosio for his remarks that helped him refine Theorem 2.11 and improve the exposition in general.

References

- [1] K. Yosida, *Functional Analysis*, Springer, 1965.
- [2] E. Lukacs, *Characteristic Functions*, Griffin, 2nd edition, 1970.
- [3] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space*, Dover Publications, 1993.
- [4] F. Martellosio, "Power properties of invariant tests for spatial autocorrelation in linear regression," *Econometric Theory*, vol. 26, no. 1, pp. 152–186, 2010.
- [5] A. D. Cliff and J. K. Ord, *Spatial Processes: Models & Applications*, Pion, 1981.
- [6] W. Krämer, "Finite sample power of Cliff-Ord-type tests for spatial disturbance correlation in linear regression," *Journal of Statistical Planning and Inference*, vol. 128, no. 2, pp. 489–496, 2005.
- [7] F. Martellosio, "Testing for spatial autocorrelation: the regressors that make the power disappear," *Econometric Reviews*, vol. 31, no. 2, pp. 215–240, 2012.
- [8] T. Kariya, "Locally robust tests for serial correlation in least squares regression," *The Annals of Statistics*, vol. 8, no. 5, pp. 1065–1070, 1980.
- [9] R. M. Dudley, *Uniform Central Limit Theorems*, vol. 63, Cambridge University Press, 1999.
- [10] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, 1968.
- [11] A. N. Kolmogorov and S. V. Fomin, *Introduction to the Theory of Functions and Functional Analysis*, Nauka, 1972.
- [12] V. I. Burenkov, "Approximation by infinitely differentiable functions with preservation of boundary values," in *Proceedings of the Steklov Mathematical Institute*, vol. 180, pp. 76–79, 1989.
- [13] C. R. Rao, *Linear Statistical Inference and Its Applications*, John Wiley & Sons, 1965.
- [14] N. Dunford and J. T. Schwartz, *Linear Operators. Part I: General Theory*, Wiley-Interscience, 1958.
- [15] B. Van Es and H.-W. Uh, "Asymptotic normality of kernel-type deconvolution estimators," *Scandinavian Journal of Statistics*, vol. 32, no. 3, pp. 467–483, 2005.

