IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 October 2023, accepted 4 November 2023, date of publication 7 November 2023,
date of current version 15 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3330903

RESEARCH ARTICLE

Minimax Density Estimation Under Radial Symmetry

KWAN-YOUNG BAK [1,2] **AND JA-YONG KOO** [3]
[1] School of Mathematics, Statistics and Data Science, Sungshin Women's University, Seoul 02844, Republic of Korea
[2] Data Science Center, Sungshin Women's University, Seoul 02844, Republic of Korea
[3] Department of Statistics, Korea University, Seoul 02841, Republic of Korea

Corresponding author: Ja-Yong Koo (jykoo@korea.ac.kr)

The work of Kwan-Young Bak was supported by the National Research Foundation (NRF) of Korea under Grant RS-2022-00165581.
The work of Ja-Yong Koo was supported by NRF of Korea under Grant RS-2023-00253020 and Grant RS-2023-00219212.

ABSTRACT This study illustrates a dimensionality reduction effect of radial symmetry in nonparametric density estimation. To deal with the class of radially symmetric functions, we adopt a generalized translation operation that preserves the symmetry structure. Radial kernel density estimators based on directly or indirectly observed random samples are proposed. For the latter case, we analyze deconvolution problems with four distinct scenarios depending on the symmetry assumptions on the signal and noise. Minimax upper and lower bounds are established for each scheme to investigate the role of the radial symmetry in determining optimal rates of convergence. The results confirm that the radial symmetry reduces the dimension of the estimation problems so that the optimal rate of convergence coincides with the univariate convergence rate except at the origin where a singularity occurs. The results also imply that the proposed estimators are rate optimal in the minimax sense for the Sobolev class of densities.

INDEX TERMS Deconvolution, Fourier analysis, Hankel transform, minimax risk, radial symmetry.

I. INTRODUCTION

Radially symmetric density functions form an important class of probability densities from both theoretical and practical viewpoints. They constitute a subclass of elliptically contoured distributions, which has received special attention in multivariate analysis; see [1] for a detailed account. Radial distributions frequently arise in practice. In physical chemistry, atomic and molecular orbitals are often modeled using spherically symmetric electron density functions; for example, see [2]. Radial distributions also regularly appear in geospatial analysis. Moreover, there is a good reason to believe that data follow a radial distribution in some applications, such as radar sea clutter data in [3] and animal motion data in [4].

From a theoretical perspective, symmetry plays a crucial role in the statistical analysis on symmetric spaces. Reference [5] considered the deconvolution problem on the Poincaré half plane, and [6] examined Wishart mixture density estimation on the space of symmetric positive matrices. The $\mathbb{SO}(2)$-invariance and $\mathbb{O}(m)$-invariance are deeply involved in their analysis, where $\mathbb{SO}(2)$ and $\mathbb{O}(m)$ denote the special orthogonal group in dimension 2 and orthogonal group in dimension m, respectively. Since these invariances correspond to the rotational invariance in $\mathbb{R}^d$, understanding the effect of the symmetry in density estimation on $\mathbb{R}^d$ can strengthen our understanding of such estimation and deconvolution in other symmetric spaces.

This study examines minimax estimation of radial densities. We first consider the standard density estimation problem based on directly observed random samples. We propose a radial density estimator that extends the standard kernel density estimator in $\mathbb{R}^2$. We obtain minimax risk upper bounds of the proposed estimator in the pointwise metric when the true density belongs to a radial Sobolev class. Corresponding lower bounds are established to determine optimal convergence rates for the radial density estimation problem, and to illustrate that the proposed estimator is rate optimal in the minimax sense. The results imply that the radial symmetry has a dimensionality reduction effect, except at the origin, where the symmetry adds no information to the estimation.

The associate editor coordinating the review of this manuscript and approving it for publication was Sajid Ali.

© 2023 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

Our analysis is extended to the deconvolution problem in which any empirical access is restricted to the data contaminated by additive random noise. Our approach, developed to deal with the radial symmetry, preserves the convolution structure under the Fourier transform. This property facilitates the development of a unified approach to the deconvolution problem with radial symmetry. We assume that the error distribution possesses the polynomial decay rate in the Fourier domain and consider four distinct cases depending on the symmetry assumptions on the signal and noise. We obtain minimax upper and lower bounds to conclude that the radial symmetry has a dimensionality reduction effect in the deconvolution problem. We find that only the radial symmetry of the signal variable has an influence on the rates of convergence. Again, we observe a singularity at the origin that results in a rate slowdown.

The main tool in our analysis is the Fourier transform, which is recognized as a constant multiple of the Hankel transform (or Bessel transform) of order zero for radial functions. The use of the Fourier calculus provides an elegant method to analyze minimaxity of the kernel density estimator. Furthermore, the Fourier transform lies at the center of the deconvolution technique since the additive contamination effect can be naturally separated in the Fourier domain; see the references in Section II. A fundamental technical difficulty of the analysis comes from a lack of a translation operation under the symmetry assumption. To resolve this issue, we adopt a generalized translation defined through the zero-order Bessel function of the first kind in the Fourier domain. We define a radial kernel density estimator with the generalized translation operation. The expected risk of the estimator for the radial Sobolev class of densities is analyzed in the Fourier domain based on the L_2-isometry of the Fourier transform. Upon obtaining risk upper bounds, we derive the corresponding minimax lower bounds using the Le Cam method [7] to examine the complexity of the estimation problem and the optimality of the proposed method. The convolution theorem enables us to apply similar lines of reasoning to the deconvolution problem under certain regularity conditions.

Our primary contribution lies in the development of a unified framework to analyze symmetry in function estimation through minimax analysis. An alternative strategy for addressing the given problem involves extracting radial coordinates from the data and constructing an estimator based on a density estimator over the positive real line. This approach would involve assuming a Hölder-type function class or employing Mellin transform-based Fourier calculus to explore the theoretical implications of the symmetry. For example, [8] considered a kernel method for the estimation of densities supported on the positive real line, which can be modified to obtain an estimator for the radial density. However, directly comparing outcomes from this approach with the established minimax result for function estimation in the Sobolev class is challenging. Consequently, it remains uncertain whether results akin to those found in our study, such as dimensionality reduction and singularity point occurrences, can be attained. Another significant aspect of our analytical methods lies in the preservation of the convolution structure under the Fourier transform. This enables us to directly analyze scenarios where the uncorrupted signal and/or contamination distribution exhibit symmetry. In contrast, focusing solely on univariate function analysis restricts us to analyzing only half of the cases, with ambiguity surrounding the concurrence of analysis with standard minimaxity results. Finally, the findings of this study can be broadly extended to K-invariant density estimation on general symmetric spaces, leveraging the group action and the Helgason-Fourier transform. This feasibility is rooted in our theory and computations, which are firmly anchored in a unified framework built upon the standard Fourier calculus.

The remainder of this paper is organized as follows. Section II presents an overview of the literature. In Section III, we collect mathematical preliminaries including the Fourier analysis and generalized translation. Section IV defines the radial kernel density estimator and investigates its minimaxity. Minimax analysis of the deconvolution problems under the radial symmetry is presented in Section V. Section VI discusses possible generalizations of the results of this paper. The proofs of the main results are deferred to Appendix VI.

II. OVERVIEW OF THE LITERATURE

This section provides an overview of the literature related to the theoretical analysis of the kernel density estimation based on the Fourier calculus. The origins of the current form of kernel density estimation method can be found in [9] and [10]. Since kernel density estimation is a method with a long history, one may refer to monographs such as [11] and [12] for a comprehensive overview. The use of Fourier calculus in kernel density estimation dates to [10] and [13]. Many studies have extended the method of minimax analysis in the Fourier domain to address the deconvolution problem. Important earlier works in this direction include [14], [15], [16], [17], [18], [19], and [20]. See also [21] for an overview of nonparametric deconvolution. The method was further extended to analyze the density deconvolution on symmetric spaces by [5] and [6]. Our study adopts a similar analysis method based on the Hankel transform. One may refer to, for example, [22], [23], and [24] for relevant mathematical backgrounds.

III. PRELIMINARIES

When dealing with radial functions, it is more convenient to work with the polar coordinates system rather than the rectangular coordinates system. Let $\mathbb{R}_+$ be the set of all positive real numbers. The polar coordinates $(r, u) \in \mathbb{R}_+ \times [0, 2\pi)$ for $x \in \mathbb{R}^2$ are defined as

$$x = rk_u, \; k_u = \begin{bmatrix} \cos u \\ \sin u \end{bmatrix}.$$

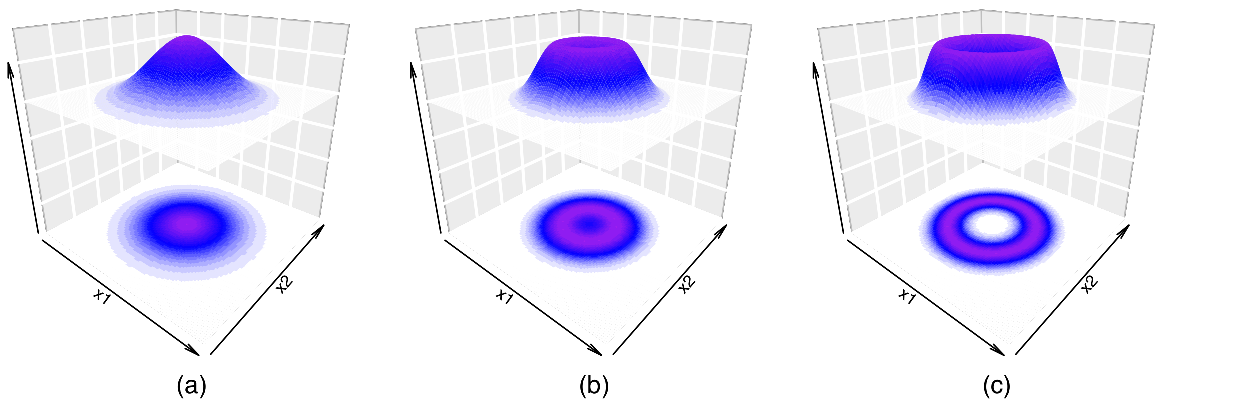

FIGURE 1. Three perspective plots with the contours illustrating the translation operation by T_s. Plot (a) is the plot of the original function g. Plots (b) and (c) present the functions translated by T_1 and $T_{\sqrt{2}}$, respectively.

A function g is called radial if it depends only on the radial part of its argument so that

$$g(rk_u) = g(rk_0),\ (r, u) \in \mathbb{R}_+ \times [0, 2\pi).$$

For $x = (x_1, x_2) \in \mathbb{R}^2$, we write $dx = dx_1 dx_2$. Recall that the Fourier transform in rectangular coordinates is defined as

$$\mathcal{F}g(\gamma) = \int_{x\in\mathbb{R}^2} g(x) e^{i\langle\gamma,x\rangle} dx,$$

with the inversion

$$g(x) = \mathcal{F}^{-1}[\mathcal{F}g(\cdot)](x) = \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2} \mathcal{F}g(\gamma) e^{-i\langle\gamma,x\rangle} d\gamma,$$

where $x = (x_1, x_2)$, $\gamma = (\gamma_1, \gamma_2)$ and $\langle\gamma, x\rangle = \gamma_1 x_1 + \gamma_2 x_2$.

For a radial function g, the Fourier transform is recognized as the 2π times Hankel transform of order zero. When g is radial, we have

$$\mathcal{F}g(\rho k_\theta) = \mathcal{F}g(\rho k_0) = 2\pi \int_{r=0}^{\infty} g(rk_0) J(\rho r) r dr,$$

where $J(\cdot)$ denotes the zero-order Bessel function of the first kind. The inverse transform is

$$\begin{aligned} g(rk_u) = g(rk_0) &= \mathcal{F}^{-1}[\mathcal{F}g(\cdot)](\rho k_0) \\ &= \frac{1}{2\pi}\int_{\rho=0}^{\infty} \mathcal{F}g(\rho k_0) J(\rho r) \rho d\rho. \end{aligned}$$

An important property of the Fourier transform is that it extends to an L_2-isometric mapping. That is, we have the Plancherel identity

$$\int_{x\in\mathbb{R}^2} |g(x)|^2 dx = \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2} |\mathcal{F}g(\gamma)|^2 d\gamma.$$

When g is radial, it is expressed as

$$\int_{r=0}^{\infty} |g(rk_0)|^2 r dr = \frac{1}{4\pi^2}\int_{\rho=0}^{\infty} |\mathcal{F}g(\rho k_0)|^2 \rho d\rho.$$

The dilation operation for radial function can be defined in a usual way in $\mathbb{R}^2$. Let the dilation $D_h g$ of a function g (not necessarily radial) be defined as

$$D_h g(x) = g(x/h),\ h > 0.$$

Equivalently, in the polar coordinates system, we have

$$D_h g(rk_u) = g\left(\frac{r}{h} k_u\right).$$

A fundamental technical difficulty of the analysis comes from a lack of a translation operation under the radial symmetry assumption. The standard translation operation in $\mathbb{R}^2$ is not appropriate for the radial function class, since the resulting function is not radially symmetric with respect to the origin. To resolve this issue, we adopt a generalized translation operation defined as follows.

Consider the polar representation of the difference of the two vectors in $\mathbb{R}^2$. Define

$$\begin{cases} \tau(s, r, w) &= \sqrt{s^2 + r^2 - 2\, sr \cos w} \\ \eta(s, r, v, u) &= \tan^{-1}\left(\dfrac{s \sin v - r \sin u}{s \cos v - r \cos u}\right). \end{cases}$$

Then

$$sk_v - rk_u = \tau(s, r, v - u) k_{\eta(s,r,v,u)}.$$

For $s \in \mathbb{R}_+$, let T_s be the translation operator defined by

$$\begin{aligned} &T_s g(rk_u) = T_s g(rk_0) \\ &\quad = \frac{1}{2\pi}\int_{w=0}^{2\pi} g(\tau(r, s, w) k_0) dw,\ (r, u) \in \mathbb{R}_+ \times [0, 2\pi), \end{aligned}$$

where r is the radial part of x, and g is a radial function.

Recall that the standard translation on $\mathbb{R}$ can be understood as the convolution of a function g on $\mathbb{R}$ with the delta function, which, in the Fourier domain, is expressed as

$$\mathcal{F}[g(\cdot - s)](t) = e^{its}\mathcal{F}g(t),\ t, s \in \mathbb{R}.$$

The generalized translation for the radial function defined above is indeed the "right" generalization in that it results in

$$\mathcal{F}[T_s g(\cdot)](\rho k_\theta) = J(s\rho)\mathcal{F}g(\rho k_0),\ (\rho, \theta) \in \mathbb{R}_+ \times [0, 2\pi).$$

Since g is radial, the complex exponential function is replaced by the Bessel function. This result is proved in Lemma 1. Following the above discussion, the generalized translation

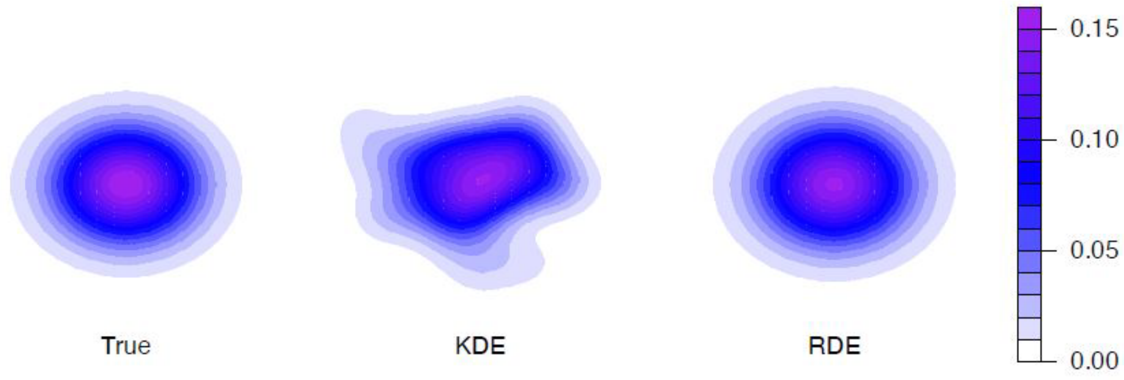

FIGURE 2. The leftmost plot is the contour plot of the standard normal density. The middle and rightmost plots illustrate the contours of the standard kernel density estimator (KDE) and the proposed radial kernel density estimator (RDE) based on a sample of size 50, respectively.

can be understood as the convolution of a radial function with the delta function. That is, the mass of the function is spread out in the vicinity of the circle of radius s with the spreading shape determined by the radial function g. Figure 1 illustrates the translation operation for radial functions.

IV. MINIMAX KERNEL ESTIMATION OF RADIAL DENSITY FUNCTIONS

Let $X_1, \ldots, X_N$ be independent copies of a random variable X with the density f with respect to the Lebesgue measure on $\mathbb{R}^2$. We consider the case in which the density f is a radial function.

Given a radial kernel function K and bandwidth $h > 0$, we define the radial kernel density estimator as

$$\begin{aligned}\hat{f}(x) &= \frac{1}{N}\sum_{n=1}^{N} T_{R_n}K_h(x)\\ &= \frac{1}{2\pi N}\sum_{n=1}^{N}\int_{w=0}^{2\pi} K_h\left(\tau(r, R_n, w)k_0\right)dw,\ x \in \mathbb{R}^2,\end{aligned}$$

where $K_h(\cdot) = h^{-2}K(\cdot/h)$, and r and R_n denote the radial parts of x and X_n, respectively. The proposed estimator is seen to be a form of the kernelization of $x - X_n$ followed by integration with respect to the angular part. Since the angular part is integrated out, the proposed estimator depends only on the radial part r of x, and is therefore radial. Lemma 2 demonstrates that the proposed estimator is a valid density provided that the kernel is given by a density function.

Our previous discussion on the generalized translation implies that the proposed estimator can be understood as follows. The impulse of $X_1, \ldots, X_N$ is first averaged around the circle of radius $R_1, \ldots, R_N$. Then, it is evenly spread out in the vicinity of the circle with the spreading shape determined by the radial kernel function K. The proposed estimator is illustrated with Figure 2. It can be seen that the radial kernel estimator closely recovers the true density with a small sample size ($N = 50$).

Consider a Sobolev class of radial densities

$$\mathcal{S}_\alpha(Q) = \Big\{ g \in L_2(\mathbb{R}^2),\ g \text{ is radial}: \int_{x\in\mathbb{R}^2} g(x)\,dx = 1,\ \|\Delta^{\alpha/2}g\|^2 \le Q^2 \Big\},$$

where $Q > 0$ and $\alpha > 1$ is a smoothness parameter. Here, $\|\cdot\|$ denotes the L^2-norm, and $\Delta^{\alpha/2}f$ is the function satisfying $\mathcal{F}\left(\Delta^{\alpha/2}g\right)(\rho) = \lambda_\rho^{\alpha/2}\mathcal{F}g(\rho)$, where $\lambda_\rho = \rho^2$ is an eigenvalue of the Laplace operator.

We choose the kernel function defined by the Fourier transform

$$\mathcal{F}K(\rho k_\theta) = \frac{1}{1+\rho^{2\alpha}},\ (\rho, \theta) \in \mathbb{R}_+ \times [0, 2\pi).$$

This choice is mainly for simplicity, and the results herein are valid for other kernel functions provided that they satisfy certain smoothness conditions. For example, the theoretical results are valid for a Pinsker-type kernel defined as

$$\mathcal{F}K(\rho k_\theta) = \Big[1 - \rho^{2\alpha}\Big]_+,\ (\rho, \theta) \in \mathbb{R}_+ \times [0, 2\pi),$$

where $[\cdot]_+$ is the plus function.

Here and throughout this paper, let $M_1, M_2, \ldots$ and $C_1, C_2, \ldots$ denote positive constants independent of the sample size N, which may differ at various places. The maximum risk for the radial Sobolev class of densities in the pointwise metric is upper bounded as follows.

Theorem 1: Choose $h = M_1 N^{-\frac{1}{2\alpha}}$. Then, we have

$$\sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}(x_0) - f(x_0)\Big]^2 \le M_2 N^{-\frac{2\alpha-1}{2\alpha}},\ x_0 \ne 0,$$

and

$$\sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}(0) - f(0)\Big]^2 \le M_3 N^{-\frac{2\alpha-2}{2\alpha}}.$$

The optimal pointwise rate of convergence for the Sobolev class without the symmetry can be obtained using the renormalization argument of [25]. The optimal rate in this case is of order $N^{-\frac{2\alpha-d}{2\alpha}}$ on $\mathbb{R}^d$. It can be seen that the upper bound at $x_0 \neq 0$ for the radial Sobolev class $\mathcal{S}_\alpha(Q)$ in Theorem 1 corresponds to the univariate convergence rate. This can be understood as a dimensionality reduction effect that comes from the radial symmetry. However, a singularity occurs at the origin since the symmetry adds no information to the estimation at this point. Thus, the upper bound coincides with the bivariate convergence rate at the origin.

For a complete understanding of the dimensionality reduction effect due to the radial symmetry, we obtain the corresponding minimax lower bounds using the Le Cam method [7] (see also Lemma 1 of [26] and Theorem 2.2 of [27]). We provide separate analyses when $x_0 = 0$ and $x_0 \neq 0$. The generalized translation proves useful for the case in which $x_0 \neq 0$, since we need to construct two densities in $\mathcal{S}_\alpha(Q)$ with an appropriate separation rate. The standard translation operation is not applicable because the resulting function no longer belongs to the class $\mathcal{S}_\alpha(Q)$.

When $x_0 = 0$, the evaluation functional is homogeneous [25]. Thus, the renormalization argument can be applied to conclude that the optimal rate is of order $N^{-\frac{2\alpha-2}{2\alpha}}$ without appealing to the Le Cam method. However, we present the proof based on explicit construction of a two point subfamily for later references. The results are summarized in Theorem 2.

Theorem 2: As $N \to \infty$, we have

$$\inf_T \sup_{f \in \mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-1}{2\alpha}} |T(x_0) - f(x_0)|^2\right] \geq M_4, \ x_0 \neq 0,$$

and

$$\inf_T \sup_{f \in \mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-2}{2\alpha}} |T(0) - f(0)|^2\right] \geq M_5,$$

where the infimum is taken over all estimators of f.

Theorem 1 and 2 imply that imposing the radial symmetry reduces the difficulty of the density estimation problem by one dimension except at the origin. The results also suggest that the proposed estimator is rate optimal in the minimax sense for the radial Sobolev class.

V. DECONVOLUTION UNDER THE RADIAL SYMMETRY

This section deals with the deconvolution problems when the signal and/or error variables follow a radial distribution. Let $Y_1, \ldots, Y_N$ be a random sample of size N where

$$Y_n = X_n + \varepsilon_n, \ n = 1, \ldots, N.$$

Suppose that $X_1, \ldots, X_N$ are incorrupted i.i.d. random variables with the density f_X, and $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. random variables with the density f_ε, representing the contamination of the data. We assume that X_n and ε_n are independent for $n = 1, \ldots, N$ so that the density f_Y of contaminated data is given by $f_Y = f_X * f_\varepsilon$. Here, $*$ denotes the convolution operation defined as

$$(g_1 * g_2)(x) = \int_{y \in \mathbb{R}^2} g_1(x-y) g_2(y) dy,$$

for the functions g_1 and g_2 defined on $\mathbb{R}^2$. Furthermore, we assume that the error distribution is known in advance. Our goal is to estimate the unknown density f_X from a set of contaminated data $Y_1, \ldots, Y_N$. Four distinct cases can be considered depending on the symmetry assumptions on X and ε. We analyze the minimaxity in each case and examine the effect of the radial symmetry in the deconvolution problem.

The smoothness of convolution kernels is usually characterized as ordinary smooth and supersmooth depending on the decay rate of the characteristic function; see [16] for example. In the supersmooth case, it is known that the bias term dominates the variance term, and hence the dimension does not appear in the optimal convergence rate when we consider a Sobolev class of functions. Our main focus is on the dimensionality reduction effect of the radial symmetry on the convergence rates. Thus, we confine our attention to the ordinary smooth error distributions whose characteristic functions satisfy, for $\gamma \in \mathbb{R}^2$,

$$M_6 |\gamma|^{-\beta} \leq |\mathcal{F} f_\varepsilon(\gamma)| \leq M_7 |\gamma|^{-\beta} \quad \text{as} \quad |\gamma| \to \infty,$$

where $|\cdot|$ denotes the usual ℓ_2 norm of a vector. In the sequel, we assume the following:

(D1) $\mathcal{F} f_\varepsilon(\gamma) \neq 0, \ \gamma \in \mathbb{R}^2$
(D2) $|\mathcal{F} f_\varepsilon(\gamma)| |\gamma|^\beta \geq M_6$ as $|\gamma| \to \infty$
(D3) $\int_{\gamma \in \mathbb{R}^2} |\mathcal{F} K(\gamma)| |\gamma|^\beta d\gamma \leq M_8$

A. CASE 1: NEITHER OF X AND ε IS RADIAL

Suppose that f_X and f_ε are not radial. This case corresponds to a classical deconvolution problem; see, for example, [16] and [21]. In particular, we assume that $f_X \in S^*_\alpha(Q)$ where $S^*_\alpha(Q)$ denotes the Sobolev class of densities without the radial symmetry defined as

$$\mathcal{S}^*_\alpha(Q) = \left\{ g \in L_2(\mathbb{R}^2) : \int_{x \in \mathbb{R}^2} g(x)\, dx = 1, \ \|\Delta^{\alpha/2} g\|^2 \leq Q^2 \right\}.$$

We establish a minimax risk upper bound of the standard kernel deconvolution estimator for the Sobolev class of densities. Since no symmetry assumption is imposed, we work with the rectangular coordinates system in this case.

We consider a kernel deconvolution estimator defined as

$$\hat{f}_X(x) = \frac{1}{4\pi^2} \int_{\gamma \in \mathbb{R}^2} \frac{\mathcal{F}K(h\gamma) \widehat{\Psi}_Y(\gamma)}{\mathcal{F} f_\varepsilon(\gamma)} e^{-i\langle \gamma, x \rangle} d\gamma, \ x \in \mathbb{R}^2, \tag{1}$$

where $\widehat{\Psi}_Y$ is the empirical characteristic function

$$\widehat{\Psi}_Y(\gamma) = \frac{1}{N} \sum_{n=1}^N e^{i\langle \gamma, Y_n \rangle}.$$

This estimator admits the following upper bound for the maximum risk over the Sobolev class of densities.

Theorem 3: Choose $h = M_9 N^{-\frac{1}{2\alpha+2\beta}}$*. Then, we have, as* $N \to \infty$*,*

$$\sup_{f \in \mathcal{S}^*_\alpha(Q)} \mathbb{E}\Big[\hat{f}_X(x_0) - f_X(x_0)\Big]^2 \leq M_{10} N^{-\frac{2\alpha-2}{2\alpha+2\beta}},\ x_0 \in \mathbb{R}^2.$$

The renormalization argument of [25] can be applied when neither of X and ε is radial. As explained in Section 5 of [25], we may use the Riesz transform to replace the inhomogeneous convolution operator in the ordinary smooth case. This leads to a reformulation of the problem that renormalizes in an asymptotic sense. Then, it is not difficult to calculate that the optimal rate of convergence is of order $N^{-\frac{2\alpha-d}{2\alpha+2\beta}}$ on $\mathbb{R}^d$. This implies that the kernel deconvolution estimator $\hat{f}_X$ achieves the optimal rate of convergence.

B. CONVOLUTION BETWEEN RADIAL FUNCTIONS

Before we analyze the deconvolution problem under the radial symmetry, we give some justifications for our approach to deal with the radial function. We may view a radial functions as a univariate function defined on $\mathbb{R}_+$, and consider a standard convolution operation on $\mathbb{R}_+$. However, this approach does not preserve the convolution structure of the Fourier analysis on $\mathbb{R}^2$. Instead, we handle the radial function in a way that coincides with the standard Fourier transform method in $\mathbb{R}^2$ based on the generalized translation. This property enables us to investigate all possible cases under the radial symmetry in a unified way as we will see in what follows.

To illustrate this idea further, we provide a result for the concordance between two convolution operations. Let G_1 and G_2 be functions defined on $\mathbb{R}_+$. The convolution operation on $\mathbb{R}_+$ in our approach has the form

$$(G_1 \circ G_2)(r) = \int_{s=0}^{\infty} T_s G_1(r) G_2(s) s ds,\ r \in \mathbb{R}_+,$$

where, with a slight abuse of notation,

$$T_s G_1(r) = \frac{1}{2\pi} \int_{w=0}^{2\pi} G_1(\tau(r,s,w)) dw.$$

For a radial function g on $\mathbb{R}^2$, we define a lift-down function $g^{\downarrow}$ on $\mathbb{R}_+$ by

$$g^{\downarrow}(r) = 2\pi g(rk_0),\ r \in \mathbb{R}_+.$$

Then, we have

$$\mathcal{F}g(\rho k_\theta) = \mathcal{B}g^{\downarrow}(\rho),\ (\rho,\theta) \in \mathbb{R}_+ \times [0, 2\pi),$$

where $\mathcal{B}$ denotes the order zero Hankel transform defined as

$$\mathcal{B}g^{\downarrow}(\rho) = \int_{r=0}^{\infty} g^{\downarrow}(r) J(\rho r) r dr,\ \rho \in \mathbb{R}_+.$$

On the other hand, we may start with a univariate function on $\mathbb{R}_+$. For a function G defined on $\mathbb{R}_+$, we define a lift-up function $G^{\uparrow}$ on $\mathbb{R}^2$ by

$$G^{\uparrow}(rk_u) = \frac{1}{2\pi} G(r),\ (r,u) \in \mathbb{R}_+ \times [0, 2\pi),$$

which gives

$$\mathcal{F}G^{\uparrow}(\rho k_\theta) = \mathcal{B}G(\rho),\ (\rho,\theta) \in \mathbb{R}_+ \times [0, 2\pi).$$

With this formulation, we have the following proposition.

Proposition 1: Suppose f_1 *and* f_2 *are radial functions on* $\mathbb{R}^2$*. Then,*

$$(f_1 * f_2)^{\downarrow} = f_1^{\downarrow} \circ f_2^{\downarrow}.$$

Moreover, for functions g_1 *and* g_2 *on* $\mathbb{R}_+$*,*

$$(g_1 \circ g_2)^{\uparrow} = g_1^{\uparrow} * g_2^{\uparrow}.$$

Proposition 1 with the convolution theorem implies Proposition 2.

Proposition 2: For functions g_1 *and* g_2 *defined on* $\mathbb{R}_+$*,*

$$\mathcal{B}(g_1 \circ g_2)(\rho) = \mathcal{B}g_1(\rho)\mathcal{B}g_2(\rho),\ \rho \in \mathbb{R}_+$$

and

$$\mathcal{F}(g_1^{\uparrow} * g_2^{\uparrow})(\rho k_\theta) = \mathcal{F}g_1^{\uparrow}(\rho k_\theta)\mathcal{F}g_2^{\uparrow}(\rho k_\theta),\quad (\rho,\theta) \in \mathbb{R}_+ \times [0, 2\pi).$$

Moreover, for radial functions f_1 *and* f_2 *on* $\mathbb{R}^2$*,*

$$\mathcal{B}(f_1^{\downarrow} \circ f_2^{\downarrow})(\rho) = \mathcal{B}f_1^{\downarrow}(\rho)\mathcal{B}f_2^{\downarrow}(\rho),\ \rho \in \mathbb{R}_+.$$

The preceding arguments imply that our approach identify a radial function on $\mathbb{R}^2$ with a function on $\mathbb{R}_+$ in a way that preserves the convolution structure with respect to the Fourier calculus. It will prove useful in subsequent analyses of the deconvolution problem.

C. CASE 2: BOTH X AND ε ARE RADIAL

Suppose that f_X and f_ε are radial densities. As in Section IV, we determine the minimax optimal rate of convergence to identify the dimensionality reduction effect of the radial symmetry in the deconvolution problem. In the following, we adopt the polar coordinates system and assume that $f_X \in \mathcal{S}_\alpha(Q)$.

Motivated by the standard kernel deconvolution estimator (1), we consider a radial kernel deconvolution estimator defined as

$$\hat{f}_X(x) = \frac{1}{2\pi} \int_{\rho=0}^{\infty} \frac{\mathcal{F}K(h\rho k_0)\widehat{\Phi}_Y(\rho)}{\mathcal{F}f_\varepsilon(\rho k_0)} J(\rho r) \rho d\rho$$

for $x = rk_u \in \mathbb{R}^2$, where

$$\widehat{\Phi}_Y(\rho) = \frac{1}{N} \sum_{n=1}^{N} J(\rho S_n),$$

with S_n being the radial part of Y_n for $n = 1, \ldots, N$. A risk upper bound is given by Theorem 4.

Theorem 4: Choose $h = M_{11} N^{-\frac{1}{2\alpha+2\beta}}$*. Then, we have, as* $N \to \infty$*,*

$$\sup_{f \in \mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}_X(x_0) - f_X(x_0)\Big]^2 \leq M_{12} N^{-\frac{2\alpha-1}{2\alpha+2\beta}},\ x_0 \neq 0,$$

and

$$\sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}_X(0)-f_X(0)\Big]^2 \le M_{13}N^{-\frac{2\alpha-2}{2\alpha+2\beta}}.$$

Theorem 3 and 4 show that the information of the radial symmetry improves the risk upper bounds by reducing the dimension. Again, from the upper bounds, it is expected that a singularity occurs at the origin. We derive the corresponding lower bounds to conclude that the rates given in Theorem 4 are indeed optimal in the minimax sense.

As in Section IV, we make use of the Le Cam method to obtain lower bounds for $x_0 \neq 0$. To this end, we construct a pair of radial densities f_0 and f_1 in $\mathcal{S}_\alpha(Q)$ so that the separation rate $|f_0(x_0)-f_1(x_0)|$ is as large as possible under the constraint that $\chi^2(f_0 * f_\varepsilon, f_1 * f_\varepsilon)$ is of order N^{-1}. When $x_0 = 0$, the evaluation functional is homogeneous, and thus the renormalization argument provides a desired lower bound.

Theorem 5: As $N\to\infty$, we have

$$\inf_T \sup_{f_X\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[N^{\frac{2\alpha-1}{2\alpha+2\beta}}|T(x_0)-f_X(x_0)|^2\Big] \ge M_{14},\ x_0\neq 0,$$

and

$$\inf_T \sup_{f_X\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[N^{\frac{2\alpha-2}{2\alpha+2\beta}}|T(0)-f_X(0)|^2\Big] \ge M_{15},$$

where the infimum is taken over all estimators of f.

D. CASE 3: X IS NON-RADIAL AND ε IS RADIAL

This case is rather trivial. Since f_X is not radial, we can construct a pair of densities $f_0, f_1 \in \mathcal{S}^*_\alpha(Q)$ in a usual way. The function f_1 is constructed by adding a suitable perturbation function to f_0. No symmetry is involved in the constraint $f_0, f_1 \in \mathcal{S}^*_\alpha(Q)$ and the calculation of the separation rate $|f_0(x_0)-f_1(x_0)|$. Moreover, as can be seen in the proof of Theorem 5, the calculation of the chi-squared divergence $\chi^2(f_0 * f_\varepsilon, f_1 * f_\varepsilon)$ is not affected by the symmetry of f_ε. The preceding arguments imply that the optimal rate of convergence in this case is given by the bivariate convergence rate of order $N^{-\frac{2\alpha-2}{2\alpha+2\beta}}$ as in Case 1. The proof of this result is immediate 2-dimensional generalization of the proof presented in [16], and hence will be omitted.

E. CASE 4: X IS RADIAL AND ε IS NON-RADIAL

Obtaining a minimax estimator is more involved in this case, since the radial symmetry of f_X is masked by the convolution with f_ε. It can be seen in Lemma 12 that the bias of the estimator (1) does not depend on the error distribution. Thus, we may conjecture that the optimal rate in this case is given by the univariate convergence rate.

Lemma 11 implies that the deconvolution estimator (1) can be re-expressed as

$$\hat{f}_X(x) = \frac{1}{N}\sum_{n=1}^N L_h(x-Y_n),\ x\in\mathbb{R}^2,$$

where

$$L(x) = \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2} \frac{\mathcal{F}K(\gamma)}{\mathcal{F}f_\varepsilon(\gamma/h)} e^{-i\langle\gamma,x\rangle}d\gamma,\ x\in\mathbb{R}^2.$$

Suppose that we can choose a kernel function K that cancels out the angular part of f_ε so that

$$L(rk_u) = \frac{1}{2\pi}\int_{\rho=0}^\infty \frac{\mathcal{F}K(\rho k_0)}{\mathcal{F}f_\varepsilon((\rho/h)k_0)} J(\rho r)\rho d\rho$$

for $(r,u)\in\mathbb{R}_+\times[0,2\pi)$. With this choice of the kernel, we define an estimator

$$\hat{f}_X(x) = \frac{1}{N}\sum_{n=1}^N T_{S_n}L_h(rk_0),\ x = rk_u\in\mathbb{R}^2,$$

see also Lemma 14. Then, Lemma 15 and Lemma 16 imply the following risk upper bound.

Theorem 6: Choose $h = M_{16}N^{-\frac{1}{2\alpha+2\beta}}$. Then, we have, as $N\to\infty$,

$$\sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}_X(x_0)-f_X(x_0)\Big]^2 \le M_{17}N^{-\frac{2\alpha-1}{2\alpha+2\beta}},\ x_0\neq 0,$$

and

$$\sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[\hat{f}_X(0)-f_X(0)\Big]^2 \le M_{18}N^{-\frac{2\alpha-2}{2\alpha+2\beta}}.$$

It can be easily confirmed that the radial symmetry of f_ε is not required in the proof of Theorem 5. Thus, we conclude that the optimal rate of convergence in this case is the univariate rate except at the origin. That is, the minimax rate of convergence depends only on the radial symmetry of f_X in the deconvolution problem.

VI. DISCUSSION

This study investigated the dimensionality reduction effect of the radial symmetry in nonparametric density estimation. To address technical difficulties arising from a lack of translation, we adopted the generalized translation operation that preserves the radial symmetry. This results in a radial kernel density estimator analogous to the standard kernel estimator in the Fourier domain except that we have a Bessel function of the first kind in place of the character defined by the complex exponential function. We established minimax upper and lower bounds depending on whether the data are observed directly or indirectly. It was verified that the radial symmetry reduces the dimension of the estimation problems so that the optimal rate of convergence coincides with the univariate convergence rate except at the origin where a singularity occurs. The results also imply that the proposed estimators are rate optimal in the minimax sense for the Sobolev class of densities.

Our results can be generalized and extended in several ways. One obvious generalization is toward the radial density estimation in $\mathbb{R}^d$. As in this study, we may adopt the d-dimensional spherical coordinates system and integrate out the angular part to obtain a radial estimator in $\mathbb{R}^d$. The Fourier transform of a radial function g on $\mathbb{R}^d$ is characterized by the

Hankel transform of order $(d-2)/2$. Then, we may apply the approaches developed herein to analyze the dimensionality reduction effect of the radial symmetry in $\mathbb{R}^d$.

The results of this study can also be generalized to invariant density estimation on general symmetric spaces. The minimax analyses presented herein are performed mostly in the Fourier domain. Thus, the results are expected to be extended based on the group action and Helgason-Fourier transform on a given symmetric space. One may refer to [23] and [28] for the Fourier analysis on symmetric spaces. In line with our study, we may analyze the minimax convergence rates for density estimation under the K-invariance assumptions. For example, we may consider the $\mathbb{SO}(2)$-invariant density estimation on the Poincaré upper half-plane model, or the $\mathbb{O}(m)$-invariant density estimation on the space of positive matrices.

APPENDIX
PROOFS
A. PROOF OF THEOREM 1

Lemma 1: Let g be a radial function. For $\rho \geq 0$, $h > 0$, and $\theta \in [0, 2\pi)$,

$$\mathcal{F}\big[T_{r_0}g(\cdot)\big](\rho k_\theta) = J(r_0\rho)\mathcal{F}g(\rho k_0), \tag{2}$$

$$\mathcal{F}[D_h g(\cdot)](\rho k_\theta) = h^2\mathcal{F}g(h\rho k_0), \tag{3}$$

and thus

$$\mathcal{F}\big[T_{r_0}D_h g(\cdot)\big](\rho k_\theta) = h^2 J(r_0\rho)\mathcal{F}g(h\rho k_0).$$

Proof: Note

$$\begin{aligned}
&\mathcal{F}\big[T_{r_0}g(\cdot)\big](\rho k_\theta)\\
&= \int_{r=0}^{\infty}\left[\frac{1}{2\pi}\int_{w=0}^{2\pi} g(\tau(r, r_0, w)k_0)dw\right]J(\rho r)rdr\\
&= \frac{1}{4\pi^2}\int_{w=0}^{2\pi}\int_{r=0}^{\infty}\int_{z=0}^{2\pi} g\big(\tau(r, r_0, z-w)k_{\eta(r,r_0,z,w)}\big)\\
&\quad\times e^{i\rho r\cos(0-z)}dzrdrdw\\
&= \frac{1}{4\pi^2}\int_{w=0}^{2\pi}\int_{r=0}^{\infty}\int_{z=0}^{2\pi} g(rk_z - r_0k_w)e^{i\langle\rho k_0, rk_z\rangle}dzrdrdw\\
&= \frac{1}{4\pi^2}\int_{w=0}^{2\pi}\int_{s=0}^{\infty}\int_{v=0}^{2\pi} g(sk_v)e^{i\langle\rho k_0, sk_v + r_0k_w\rangle}dvsdsdw\\
&= \frac{1}{4\pi^2}\int_{w=0}^{2\pi} e^{i\rho r_0\cos w}\\
&\quad\times\left[\int_{s=0}^{\infty}\int_{v=0}^{2\pi} g(sk_v)e^{i\rho s\cos v}dvsds\right]dw\\
&= \frac{1}{2\pi}\int_{w=0}^{2\pi} e^{i\rho r_0\cos w}dw\int_{s=0}^{\infty} g(sk_v)J(\rho s)sds\\
&= J(r_0\rho)\mathcal{F}g(\rho k_0).
\end{aligned}$$

For the dilation, we have

$$\begin{aligned}
\mathcal{F}[D_h g(\cdot)](\rho k_\theta) &= \int_{r=0}^{\infty}\int_{u=0}^{2\pi} g\Big(\frac{r}{h}k_u\Big)e^{i\rho r\cos(\theta-u)}durdr\\
&= h^2\int_{r=0}^{\infty}\int_{u=0}^{2\pi} g(sk_u)e^{ih\rho r\cos(\theta-u)}dusds\\
&= h^2\mathcal{F}g(h\rho k_\theta) = h^2\mathcal{F}g(h\rho k_0),
\end{aligned}$$

where we have used the symmetry assumption for the last equality. □

Lemma 2 shows that the proposed methods provide a valid density estimator provided the kernel is given by a density function.

Lemma 2: Let K be a radial kernel such that $\int_{x\in\mathbb{R}^2} K(x)dx = 1$. We have

$$\int_{x\in\mathbb{R}^2}\hat{f}(x)dx = 1.$$

Proof: Note

$$\begin{aligned}
&\int_{x\in\mathbb{R}^2}\hat{f}(x)dx\\
&= \frac{1}{2\pi N}\sum_{n=1}^{N}\int_{w=0}^{2\pi}\int_{u=0}^{2\pi}\int_{r=0}^{\infty} K_h\left(\tau(r, R_n, w)k_0\right)rdrdudw\\
&= \frac{1}{N}\sum_{n=1}^{N}\int_{u=0}^{2\pi}\int_{r=0}^{\infty} K_h\big(\tau(r, R_n, u-U_n)k_{\eta(r,R_n,u,U_n)}\big)\\
&\quad\times rdrdu\\
&= \frac{1}{N}\sum_{n=1}^{N}\int_{x\in\mathbb{R}^2} K_h\left(x - X_n\right)dx = 1.
\end{aligned}$$

□

Define a radial empirical characteristic function

$$\widehat{\Phi}(\rho) = \frac{1}{N}\sum_{n=1}^{N} J(\rho R_n),\ \ \rho\in\mathbb{R}_+.$$

Observe that $\mathbb{E}\widehat{\Phi}(\rho) = \mathcal{F}f(\rho k_0)$. Lemma 3 enables us to use the Fourier analysis technique.

Lemma 3: We have

$$\mathcal{F}\hat{f}(\rho k_\theta) = \widehat{\Phi}(\rho)\mathcal{F}K(h\rho k_0),\ \ \theta\in[0, 2\pi).$$

Proof: Since

$$\begin{aligned}
&\int_{u=0}^{2\pi}\int_{r=0}^{\infty}\int_{z=0}^{2\pi} K_h\left(\tau(r, R_n, z-u)k_0\right)e^{i\rho r\cos(-z)}dzrdrdu\\
&= \int_{u=0}^{2\pi}\int_{s=0}^{\infty}\int_{v=0}^{2\pi} K_h\left(sk_v\right)e^{i\rho k_0\cdot(sk_v + R_nk_u)}dvsdsdu\\
&= \int_{s=0}^{\infty} K_h\left(sk_0\right)\left[\int_{v=0}^{2\pi} e^{i\rho s\cos v}dv\right]\\
&\quad\times\left[\int_{u=0}^{2\pi} e^{i\rho R_n\cos u}du\right]sds,
\end{aligned}$$

we have

$$\begin{aligned}\mathcal{F}\hat{f}(\rho k_\theta) &= \int_{r=0}^{\infty}\left[\frac{1}{2\pi N}\sum_{n=1}^{N}\int_{u=0}^{2\pi}K_h(\tau(r,R_n,u)k_0)\,du\right]\\&\quad\times\left[\int_{z=0}^{2\pi}e^{i\rho r\cos z}dz\right]rdr\\&=\frac{1}{2\pi N}\sum_{n=1}^{N}\int_{u=0}^{2\pi}\int_{r=0}^{\infty}\int_{z=0}^{2\pi}K_h(\tau(r,R_n,z-u)k_0)\\&\quad\times e^{i\rho r\cos(-z)}dzrdrdu\\&=\left[\frac{1}{N}\sum_{n=1}^{N}J(\rho R_n)\right]2\pi\int_{s=0}^{\infty}K_h(sk_0)\,J(\rho s)\,sds\\&=\widehat{\Phi}(\rho)\mathcal{F}K_h(\rho k_0)=\widehat{\Phi}(\rho)\mathcal{F}K(h\rho k_0),\end{aligned}$$

where the last inequality follows from (3). □

Recall that the mean squared error is decomposed as

$$\mathrm{MSE}(x_0)=\mathbb{E}\Big[\hat{f}(x_0)-f(x_0)\Big]^2=B^2(x_0)+V(x_0),$$

where $B(x_0)$ and $V(x_0)$ are, respectively, the bias and variance term defined by

$$B(x_0)=\left|\mathbb{E}\hat{f}(x_0)-f(x_0)\right|$$

and

$$V(x_0)=\mathbb{E}\Big[\hat{f}(x_0)-\mathbb{E}\hat{f}(x_0)\Big]^2.$$

We find an upper bound for each term in Lemma 4 and Lemma 5.

Lemma 4: We have

$$B(x_0)\le M_{19}h^{\alpha-\frac{1}{2}},\ x_0\neq 0\quad \textit{and}\quad B(0)\le M_{20}h^{\alpha-1}.$$

Proof: Observe

$$\begin{aligned}f(x_0)&=\mathcal{F}^{-1}[\mathcal{F}f(\cdot)](r_0k_0)\\&=\frac{1}{2\pi}\int_{\rho=0}^{\infty}\mathcal{F}f(\rho k_0)J(\rho r_0)\,\rho d\rho.\end{aligned}\tag{4}$$

Lemma 3 implies

$$\mathbb{E}\hat{f}(x_0)=\frac{1}{2\pi}\int_{\rho=0}^{\infty}\mathcal{F}f(\rho k_0)\mathcal{F}K(h\rho k_0)J(r_0\rho)\,\rho d\rho.\tag{5}$$

Combining (4) and (5), we have

$$B(x_0)=\frac{1}{2\pi}\left|\int_{\rho=0}^{\infty}\mathcal{F}f(\rho k_0)J(r_0\rho)[\mathcal{F}K(h\rho k_0)-1]\,\rho d\rho\right|.$$

It follows from the definition of $\mathcal{S}_\alpha(Q)$ and the Cauchy-Schwarz inequality that

$$\begin{aligned}B(0)&=\frac{1}{2\pi}\left|\int_{\rho=0}^{\infty}\mathcal{F}f(\rho k_0)J(0)[\mathcal{F}K(h\rho k_0)-1]\,\rho d\rho\right|\\&\le\frac{1}{2\pi}\int_{\rho=0}^{\infty}|\mathcal{F}f(\rho k_0)||\mathcal{F}K(h\rho k_0)-1|\,\rho d\rho\\&\le\frac{1}{\sqrt{2\pi}}\left(\int_{\rho=0}^{\infty}\frac{1}{2\pi}|\mathcal{F}f(\rho k_0)|^2\rho^{2\alpha}\rho d\rho\right)^{1/2}\\&\quad\times\left(\int_{\rho=0}^{\infty}\frac{h^{2\alpha}(h\rho)^{2\alpha}}{(1+(h\rho)^{2\alpha})^2}\rho d\rho\right)^{1/2}\\&=\frac{Q}{\sqrt{2\pi}}h^{\alpha-1}\left(\frac{1}{2\alpha}\int_{z=0}^{1}\left(\frac{1}{z}-1\right)^{1/\alpha}dz\right)^{1/2}\\&=M_{21}h^{\alpha-1},\ M_{21}=\frac{Q\mathcal{B}(1+\alpha^{-1},1-\alpha^{-1})}{\sqrt{4\pi\alpha}},\end{aligned}$$

where $\mathcal{B}(\cdot,\cdot)$ denotes the beta function. For $x_0=r_0k_{u_0}\neq 0$, using the inequality $|J(a)|\le\sqrt{\frac{2}{\pi a}}$ for $a>0$ (see Chapter 7 of [24]), we have

$$\begin{aligned}B(x_0)&\le\frac{1}{2\pi}\int_{\rho=0}^{\infty}|\mathcal{F}f(\rho k_0)||J(r_0\rho)||\mathcal{F}K(h\rho k_0)-1|\,d\rho\\&\le\frac{Q}{\sqrt{2\pi}}\left(\int_{\rho=0}^{\infty}|\mathcal{F}f(\rho k_0)|^2\rho^{2\alpha+1}d\rho\right)^{1/2}\\&\quad\times\left(\int_{\rho=0}^{\infty}\frac{h^{2\alpha}(h\rho)^{2\alpha}}{(1+(h\rho)^{2\alpha})^2}|J(r_0\rho)|^2\rho d\rho\right)^{1/2}\\&\le M_{22}h^{\alpha-\frac{1}{2}},\end{aligned}$$

for $M_{22}=\frac{Q}{\pi\sqrt{r_0}}\sqrt{\frac{1}{4}+\frac{\mathcal{B}(1+\alpha^{-1},1-\alpha^{-1})}{2\alpha}}$. □

Lemma 5: We have

$$V(x_0)\le\frac{M_{23}}{Nh},\ x_0\neq 0\quad \textit{and}\quad V(0)\le\frac{M_{24}}{Nh^2}.$$

Proof: Following the argument in Lemma 3.1 of [29], we have $\|f\|_\infty\le M_{25}$, where $\|\cdot\|_\infty$ denotes the usual sup-norm of a function on $\mathbb{R}^2$. It follows from Lemma 1, the Plancherel theorem, and the inequality $|J(a)|\le\sqrt{\frac{2}{\pi a}}$ for $a>0$ that, for $x_0=r_0k_{u_0}\neq 0$,

$$\begin{aligned}V(x_0)&\le\frac{1}{4\pi^2N}\int_{x\in\mathbb{R}^2}\left[\int_{w=0}^{2\pi}K_h\left(\frac{\tau(r_0,r,w)}{h}k_0\right)dw\right]^2\\&\quad\times f(x)\,dx,\quad x=rk_u\\&\le\frac{2\pi M_{25}}{N}\int_{r=0}^{\infty}\left[\frac{1}{2\pi}\int_{w=0}^{2\pi}K_h\left(\frac{\tau(r_0,r,w)}{h}k_0\right)dw\right]^2rdr\\&=\frac{2\pi M_{25}}{N}\int_{r=0}^{\infty}\left[T_{r_0}K_h(rk_0)\right]^2rdr\\&=\frac{M_{25}}{2\pi N}\int_{\rho=0}^{\infty}|J(r_0\rho)\mathcal{F}K_h(\rho k_0)|^2\rho d\rho\\&\le\frac{M_{25}}{\pi^2r_0N}\int_{\rho=0}^{\infty}|\mathcal{F}K(h\rho k_0)|^2d\rho\\&\le\frac{M_{23}}{Nh},\ M_{23}=\frac{M_{25}}{\pi^2r_0}\int_{\xi=0}^{\infty}|\mathcal{F}K(\xi k_0)|^2d\xi.\end{aligned}$$

At the origin, we have

$$\begin{aligned}V(0)&\le\frac{M_{25}}{2\pi N}\int_{\rho=0}^{\infty}|\mathcal{F}K(h\rho k_0)|^2\rho d\rho\\&=\frac{M_{25}}{2\pi Nh^2}\int_{\xi=0}^{\infty}|\mathcal{F}K(\xi k_0)|^2\xi d\xi\\&\le\frac{M_{24}}{Nh^2},\ M_{24}=\frac{M_{25}\|K\|^2}{4\pi^2}.\end{aligned}$$

□

1) PROOF OF THEOREM 1

Lemma 4 and 5 imply that

$$\begin{aligned}\mathbb{E}\Big[\hat{f}(x_0)-f(x_0)\Big]^2 &= B^2(x_0)+V(x_0)\\ &\le M_{22}^2 h^{2\alpha-1}+\frac{M_{23}}{Nh}\\ &\le C_1 N^{-\frac{2\alpha-1}{2\alpha}}+C_2 N^{-1}N^{\frac{1}{2\alpha}}\\ &\le M_2 N^{-\frac{2\alpha-1}{2\alpha}},\ x_0\neq 0,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\Big[\hat{f}(0)-f(0)\Big]^2 &= B^2(0)+V(0)\\ &\le M_{21}^2 h^{2\alpha-2}+\frac{M_{24}}{Nh^2}\\ &\le C_3 N^{-\frac{2\alpha-2}{2\alpha}}+C_4 N^{-1}N^{\frac{2}{2\alpha}}\\ &\le M_3 N^{-\frac{2\alpha-2}{2\alpha}}.\end{aligned}$$

B. PROOF OF THEOREM 2

We provide proofs for the case $x_0 = 0$ and $x_0 \neq 0$ in Proposition 3 and 4, respectively. The χ^2 divergence between two probability measures P and Q is defined as $\chi^2(p,q) = \int (p-q)^2 q^{-1}$ when they admit densities p and q, respectively.

1) LOWER BOUND AT THE ORIGIN

We construct a pair of densities f_0 and f_1 to obtain a minimax lower bound. Let ψ be a radial function satisfying the following:

(L1) $\psi(0) > 0$
(L2) $\psi(rk_0) = 0$ for $r \notin [0,1]$
(L3) $\int_{r=0}^{1} \psi(rk_0)\,rdr = 0$
(L4) $\|\Delta^{\alpha/2}\psi\|^2 \le Q^2/4$

Let f_0 be a radial density in $\mathcal{S}_\alpha(Q/2)$ such that $f_0(rk_0) \ge M_{26}$ for $r \in [0,1]$, and define

$$f_1(rk_u) = f_0(rk_0) + M_{27}\delta^{\alpha-1}\psi\left(\frac{r}{\delta}k_0\right)$$

for $(r,u) \in [0,1]\times[0,2\pi)$, where M_{27} will be determined later. For $M_{28} > 0$, which will be specified later, we choose

$$\delta = \delta_N = M_{28}N^{-\frac{1}{2\alpha}}.$$

Observe, for , $M_{29} = M_{27}M_{28}^{\alpha-1}|\psi(0)|$,

$$|f_0(0)-f_1(0)| = M_{27}\delta^{\alpha-1}|\psi(0)| \ge M_{29}N^{-\frac{\alpha-1}{2\alpha}}. \quad (6)$$

Lemma 6: We have $f_0, f_1 \in \mathcal{S}_\alpha(Q)$.

Proof: By the assumptions $f_0 \in \mathcal{S}_\alpha(Q/2)$, (L4), and equation (3) in Lemma 1, we have

$$\begin{aligned}\|\Delta^{\alpha/2}f_1\|^2 &= \frac{1}{2\pi}\int_{\rho=0}^{\infty}|\mathcal{F}f_1(\rho k_0)|^2\rho^{2\alpha+1}d\rho\\ &= \frac{1}{2\pi}\int_{\rho=0}^{\infty}\Big|\mathcal{F}f_0(\rho k_0)+M_{27}\delta^{\alpha-1}\mathcal{F}[\psi(\cdot/\delta)](\rho k_0)\Big|^2\\ &\quad\times\rho^{2\alpha+1}d\rho\\ &\le \frac{Q^2}{2}+\frac{M_{27}^2\delta^{2\alpha-2}}{\pi}\int_{\rho=0}^{\infty}|\mathcal{F}[\psi(\cdot/\delta)](\rho k_0)|^2\rho^{2\alpha+1}d\rho\\ &= \frac{Q^2}{2}+\frac{M_{27}^2\delta^{2\alpha-2}}{\pi}\int_{\rho=0}^{\infty}\Big|\delta^2\mathcal{F}\psi(\delta\rho k_0)\Big|^2\rho^{2\alpha+1}d\rho\\ &\le \frac{Q^2}{2}+\frac{M_{27}^2}{\pi}\frac{\pi Q^2}{2} \le Q^2,\end{aligned}$$

provided that we choose $M_{27} \le 1$. □

Lemma 7: Let $\varrho > 0$ be given. Then, for sufficiently large N, we have

$$\chi^2(f_1,f_0) \le \frac{\varrho}{N}.$$

Proof: Let $\varrho > 0$ be given and N be large enough for $1/h$ to be greater than 1. Note

$$\begin{aligned}\chi^2(f_1,f_0) &= 2\pi\int_{r\in[0,1]}\frac{(f_1(rk_0)-f_0(rk_0))^2}{f_0(rk_0)}rdr\\ &\le \frac{2\pi M_{27}^2}{M_{26}}\delta^{2\alpha-2}\int_{r\in[0,1]}\psi^2\left(\frac{r}{\delta}k_0\right)rdr\\ &= \frac{1}{4M_{26}}\delta^{2\alpha}\int_{s=0}^{1}\psi^2(sk_0)\,sds\\ &\le M_{30}\delta^{2\alpha} \le \frac{\varrho}{N},\end{aligned}$$

provided that we choose a sufficiently small M_{28}. □

Combining (6), Lemma 6, and Lemma 7, we apply the Le Cam method to obtain the following lower bound.

Proposition 3: As $N \to \infty$, we have

$$\inf_T \sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\Big[N^{\frac{2\alpha-2}{2\alpha}}|T(0)-f(0)|^2\Big] \ge M_{38}.$$

2) LOWER BOUND AT $X_0 \neq 0$

To obtain a lower bound at $x_0 = r_0 k_{u_0} \neq 0$, we need to construct a perturbation around the circle with the radius r_0. Since we cannot make use of the standard translation operation on Euclidean spaces, we construct a perturbation based on the generalized translation operation. That is, we put mass around the circle with the radius r_0 and spread it by convolving with a radial function ψ so that the perturbation satisfies the desired smoothness conditions. Here, we can replace (L4) by

(L4-1) $\|\Delta^{(\alpha-1/2)/2}\psi\|^2 \le Q^2/4$

Define, for $(r,u) \in [0,1]\times[0,2\pi)$,

$$f_1(rk_u) = f_0(rk_0) + M_{31}\delta^{\alpha-3/2}T_{r_0}\psi\left(\frac{r}{\delta}k_0\right),$$

where M_{31} will be determined later. For $M_{32} > 0$, which will be specified later, we choose

$$\delta = \delta_N = M_{32}N^{-\frac{1}{2\alpha}}.$$

Lemma 8: We have $f_0, f_1 \in \mathcal{S}_\alpha(Q)$.

Proof: By the assumptions $f_0 \in \mathcal{S}_\alpha(Q/2)$, $\|\Delta^{(\alpha-1/2)/2}\psi\|^2 \le Q^2/4$, Lemma 1, and $|J(a)| \le \sqrt{\frac{2}{\pi a}}$ for $a > 0$, we have

$$\begin{aligned}
&\|\Delta^{\alpha/2} f_1\|^2 \\
&= \frac{1}{2\pi}\int_{\rho=0}^{\infty} |\mathcal{F}f_1(\rho k_0)|^2 \rho^{2\alpha+1} d\rho \\
&= \frac{1}{2\pi}\int_{\rho=0}^{\infty} \left|\mathcal{F}f_0(\rho k_0) + M_{31}\delta^{\alpha-3/2}\mathcal{F}\left[T_{r_0}\psi(\cdot/\delta)\right](\rho k_0)\right|^2 \\
&\quad \times \rho^{2\alpha+1} d\rho \\
&\le \frac{Q^2}{2} + \frac{M_{31}^2\delta^{2\alpha-3}}{\pi}\int_{\rho=0}^{\infty}\left|\mathcal{F}\left[T_{r_0}\psi(\cdot/\delta)\right](\rho k_0)\right|^2\rho^{2\alpha+1}d\rho \\
&= \frac{Q^2}{2} + \frac{M_{31}^2\delta^{2\alpha-3}}{\pi}\int_{\rho=0}^{\infty}\left|\delta^2 J(r_0\rho)\mathcal{F}\psi(\delta\rho k_0)\right|^2\rho^{2\alpha+1}d\rho \\
&\le \frac{Q^2}{2} + \frac{2M_{31}^2\delta^{2\alpha+1}}{\pi^2 r_0}\int_{\rho=0}^{\infty}|\mathcal{F}\psi(\delta\rho k_0)|^2\rho^{2\alpha}d\rho \\
&\le \frac{Q^2}{2} + \frac{M_{31}^2}{\pi r_0}Q^2 \le Q^2,
\end{aligned}$$

provided that we choose $M_{31} \le \sqrt{\pi r_0/2}$. □

Lemma 9: Let $\varrho > 0$ be given. Then, for sufficiently large N, we have

$$\chi^2(f_1, f_0) \le \frac{\varrho}{N}.$$

Proof: Let N be sufficiently large. Using the Plancherel identity and the inequality $|J(a)| \le \sqrt{\frac{2}{\pi a}}$ for $a > 0$, we obtain

$$\begin{aligned}
\chi^2(f_1, f_0) &= 2\pi\int_{r\in[0,1]}\frac{(f_1(rk_0) - f_0(rk_0))^2}{f_0(rk_0)}rdr \\
&\le \frac{2\pi M_{31}^2}{M_{26}}\delta^{2\alpha-3}\int_{r\in[0,1]}\left|T_{r_0}\psi\left(\frac{r}{\delta}k_0\right)\right|^2 rdr \\
&= \frac{M_{31}^2}{2\pi M_{26}}\delta^{2\alpha+1}\int_{\rho=0}^{\infty}|J(r_0\rho)\mathcal{F}\psi(\delta\rho k_0)|^2\rho d\rho \\
&\le \frac{M_{31}^2}{\pi^2 r_0 M_{26}}\delta^{2\alpha}\int_{\xi=0}^{\infty}|\mathcal{F}\psi(\xi k_0)|^2 d\xi \\
&\le M_{33}\delta^{2\alpha} \le \frac{\varrho}{N},
\end{aligned}$$

provided that we choose a sufficiently small M_{32}. □

Lemma 10: We have

$$|f_0(r_0k_0) - f_1(r_0k_0)| \ge M_{34}N^{-\frac{\alpha-1/2}{2\alpha}}.$$

Proof: Let N be sufficiently large. A change of variables $z = \sqrt{1-\cos w}/\delta$ implies

$$\begin{aligned}
&|f_0(r_0k_u) - f_1(r_0k_u)| \\
&\quad = \left|M_{31}\delta^{\alpha-3/2}T_{r_0}\psi\left(\frac{r}{\delta}k_0\right)\right| \\
&\quad = \frac{M_{31}\delta^{\alpha-3/2}}{\pi}\left|\int_{w=0}^{\pi}\psi\left(\frac{\tau(r_0, r_0, w)}{\delta}k_0\right)dw\right| \\
&\quad = M_{31}\delta^{\alpha-3/2}\left|\int_{z=0}^{1/\sqrt{2}r_0}\psi(\sqrt{2}r_0zk_0)\frac{2\delta dz}{\sqrt{2-\delta^2z^2}}\right| \\
&\quad \ge M_{35}\delta^{\alpha-3/2}2\delta\left|\int_{z=0}^{1/\sqrt{2}r_0}\psi(\sqrt{2}r_0zk_0)dz\right| \\
&\quad \ge M_{36}\delta^{\alpha-1/2} \ge M_{34}N^{-\frac{\alpha-1/2}{2\alpha}}.
\end{aligned}$$

□

Combining Lemma 8, Lemma 9, and Lemma 10, we make use of the Le Cam method to obtain the following lower bound.

Proposition 4: As $N \to \infty$, we have

$$\inf_T \sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-1}{2\alpha}}|T(x_0) - f(x_0)|^2\right] \ge M_{57},\ x_0 \ne 0.$$

3) PROOF OF THEOREM 2

Proposition 3 and 4 imply that, as $N \to \infty$,

$$\inf_T \sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-1}{2\alpha}}|T(x_0) - f(x_0)|^2\right] \ge M_{37},\ x_0 \ne 0,$$

and

$$\inf_T \sup_{f\in\mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-2}{2\alpha}}|T(0) - f(0)|^2\right] \ge M_{38}.$$

C. PROOF OF THEOREM 3

Lemma 11: Define

$$L(x) = \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\frac{\mathcal{F}K(\gamma)}{\mathcal{F}f_\varepsilon(\gamma/h)}e^{-i\langle\gamma,x\rangle}d\gamma,\ x\in\mathbb{R}^2.$$

Then, we have

$$\hat{f}_X(x) = \frac{1}{N}\sum_{n=1}^{N}L_h(x - Y_n),\ x\in\mathbb{R}^2.$$

Proof: Note

$$\begin{aligned}
\hat{f}_X(x) &= \frac{1}{4\pi^2}\int_{\nu\in\mathbb{R}^2}\frac{\mathcal{F}K(h\nu)\widehat{\Psi}_Y(\nu)}{\mathcal{F}f_\varepsilon(\nu)}e^{-i\langle\nu,x\rangle}d\nu \\
&= \frac{1}{Nh^2}\sum_{n=1}^{N}\frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\frac{\mathcal{F}K(\gamma)}{\mathcal{F}f_\varepsilon(\gamma/h)}e^{-i\langle\gamma/h,x\rangle}e^{i\langle\gamma/h,Y_n\rangle}d\gamma \\
&= \frac{1}{Nh^2}\sum_{n=1}^{N}\frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\frac{\mathcal{F}K(\gamma)}{\mathcal{F}f_\varepsilon(\gamma/h)}e^{-i\left\langle\gamma,\frac{x-Y_n}{h}\right\rangle}d\gamma \\
&= \frac{1}{Nh^2}\sum_{n=1}^{N}L\left(\frac{x-Y_n}{h}\right).
\end{aligned}$$

□

Lemma 12: We have

$$B(x_0) \le M_{39}h^{\alpha-1},\ x_0\in\mathbb{R}^2.$$

Proof: Observe

$$\begin{aligned}
\mathbb{E}\hat{f}_X(x) &= \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\frac{\mathcal{F}K(h\gamma)\mathcal{F}f_Y(\gamma)}{\mathcal{F}f_\varepsilon(\gamma)}e^{-i\langle\gamma,x\rangle}d\gamma \\
&= \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\mathcal{F}K(h\gamma)\mathcal{F}f_X(\gamma)e^{-i\langle\gamma,x\rangle}d\gamma.
\end{aligned}$$

Note that the bias term does not depend on the error distribution, and it is therefore of order $h^{\alpha-1}$ following the proof of Lemma 4. To observe this, note

$$\begin{aligned}
B(x_0) &= \frac{1}{4\pi^2}\left|\int_{\gamma\in\mathbb{R}^2}\mathcal{F}f_X(\gamma)[\mathcal{F}K(h\gamma)-1]e^{-i\langle\gamma,x_0\rangle}d\gamma\right| \\
&\le \frac{1}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}|\mathcal{F}f_X(\gamma)||\mathcal{F}K(h\gamma)-1|d\gamma \\
&\le \frac{1}{2\pi}\left(\int_{\gamma\in\mathbb{R}^2}\frac{1}{4\pi^2}|\mathcal{F}f_X(\gamma)|^2|\gamma|^{2\alpha}d\gamma\right)^{1/2} \\
&\quad\times\left(\int_{\gamma\in\mathbb{R}^2}\left[\frac{(h|\gamma|)^{2\alpha}}{1+(h|\gamma|)^{2\alpha}}\right]^2|\gamma|^{-2\alpha}d\gamma\right)^{1/2} \\
&\le \frac{Q}{2\pi}\left(2\pi\int_{\rho=0}^{\infty}\frac{h^{2\alpha}(h\rho)^{2\alpha}}{\left(1+(h\rho)^{2\alpha}\right)^2}\rho d\rho\right)^{1/2} \\
&\le M_{39}h^{\alpha-1},\ M_{39}=\frac{Q\mathcal{B}(1+\alpha^{-1},1-\alpha^{-1})}{\sqrt{4\pi\alpha}}.
\end{aligned}$$

□

Lemma 13: Suppose $h\to 0$ as $N\to\infty$. We have, as $N\to\infty$,

$$V(x_0)\le\frac{M_{40}}{Nh^{2\beta+2}},\ x_0\in\mathbb{R}^2.$$

Proof: Since $Y_1,\ldots,Y_N$ are i.i.d., it follows from Lemma 11 that

$$V(x_0)\le\frac{1}{Nh^4}\mathbb{E}L^2\left(\frac{x_0-Y_1}{h}\right).$$

Observe that $\|f_Y\|_\infty\le M_{25}$ because

$$\begin{aligned}
f_Y(y) &= \int_{x\in\mathbb{R}^2}f_X(y-x)f_\varepsilon(x)dx \\
&\le M_{25}\int_{x\in\mathbb{R}^2}f_\varepsilon(x)dx = M_{25}.
\end{aligned}$$

The Plancherel identity implies

$$\begin{aligned}
\mathbb{E}L^2\left(\frac{x_0-Y_1}{h}\right) &= \int_{y\in\mathbb{R}^2}L^2\left(\frac{x_0-y}{h}\right)f_Y(y)dy \\
&= h^2\int_{z\in\mathbb{R}^2}L^2(z)f_Y(x_0-hz)dz \\
&\le M_{25}h^2\int_{z\in\mathbb{R}^2}L^2(z)dz \\
&= \frac{M_{25}h^2}{4\pi^2}\int_{\gamma\in\mathbb{R}^2}\frac{|\mathcal{F}K(\gamma)|^2}{|\mathcal{F}f_\varepsilon(\gamma/h)|^2}d\gamma.
\end{aligned}$$

When $M_{41}h\le|\gamma|$ for a large fixed constant M_{41}, (D2) implies

$$\frac{1}{|\mathcal{F}f_\varepsilon(\gamma/h)|}\le C_1h^{-\beta}|\gamma|^\beta$$

Thus, we obtain

$$\begin{aligned}
&\int_{|\gamma|\ge M_{41}h}\frac{|\mathcal{F}K(\gamma)|^2}{|\mathcal{F}f_\varepsilon(\gamma/h)|^2}d\gamma \\
&\quad\le C_1^2h^{-2\beta}\int_{|\gamma|\ge M_{41}h}|\mathcal{F}K(\gamma)|^2|\gamma|^{2\beta}d\gamma \\
&\quad\le C_2h^{-2\beta},
\end{aligned}$$

where $C_2=C_1^2M_8$. When $|\gamma|\le M_{41}h$, (D1) implies

$$|\mathcal{F}f_\varepsilon(\gamma/h)|\ge C_3>0,\ C_3=\min_{|\nu|\le M_{41}}|\mathcal{F}f_\varepsilon(\nu)|,$$

so that

$$\int_{|\gamma|\le M_{41}h}\frac{|\mathcal{F}K(\gamma)|^2}{|\mathcal{F}f_\varepsilon(\gamma/h)|^2}d\gamma\le C_4$$

for

$$C_4=\frac{1}{C_3}\int_{|\gamma|\le M_{41}h}|\mathcal{F}K(\gamma)|^2d\gamma.$$

Combining the results, we have

$$V(x_0)\le\frac{M_{40}}{Nh^{2\beta+2}}.$$

□

1) PROOF OF THEOREM 3

Lemma 12 and 13 imply that

$$\begin{aligned}
\mathbb{E}\left[\hat{f}_X(x_0)-f_X(x_0)\right]^2 &= B^2(x_0)+V(x_0) \\
&\le M_{39}^2h^{2\alpha-2}+\frac{M_{40}}{Nh^{2\beta+2}} \\
&\le C_5N^{-\frac{2\alpha-1}{2\alpha}}+C_6N^{-1}N^{\frac{2\beta+2}{2\alpha+2\beta}} \\
&\le M_{10}N^{-\frac{2\alpha-2}{2\alpha+2\beta}},\ x_0\in\mathbb{R}^2.
\end{aligned}$$

D. PROOF OF PROPOSITION 1

Note

$$\begin{aligned}
&(f_1*f_2)^{\downarrow}(s) \\
&\quad= 2\pi(f_1*f_2)(sk_0) \\
&\quad= 2\pi\int_{r=0}^{\infty}\left[\int_{w=0}^{2\pi}f_1(\tau(s,r,w)k_0)dw\right]f_2(rk_0)rdr \\
&\quad= 2\pi\int_{r=0}^{\infty}\left[\int_{w=0}^{2\pi}\frac{1}{2\pi}f_1^{\downarrow}(\tau(s,r,w))dw\right]\frac{1}{2\pi}f_2^{\downarrow}(rk)rdr \\
&\quad= \int_{r=0}^{\infty}\left[\frac{1}{2\pi}\int_{w=0}^{2\pi}f_1^{\downarrow}(\tau(s,r,w))dw\right]f_2^{\downarrow}(rk)rdr \\
&\quad= (f_1^{\downarrow}\circ f_2^{\downarrow})(s),
\end{aligned}$$

and

$$\begin{aligned}
&(g_1^{\uparrow}*g_2^{\uparrow})(sk_v) \\
&\quad= \int_{r=0}^{\infty}\int_{u=0}^{2\pi}g_1^{\uparrow}(sk_v-rk_u)g_2^{\uparrow}(rk_u)durdr \\
&\quad= \int_{r=0}^{\infty}\int_{u=0}^{2\pi}g_1^{\uparrow}(\tau(s,r,v-u)k_{\eta(s,r,v,u)})g_2^{\uparrow}(rk_u)durdr \\
&\quad= \int_{r=0}^{\infty}\int_{u=0}^{2\pi}\left[\frac{1}{2\pi}g_1(\tau(s,r,v-u))\right]\left[\frac{1}{2\pi}g_2(r)\right]durdr \\
&\quad= \frac{1}{4\pi^2}\int_{r=0}^{\infty}\left[\int_{w=0}^{2\pi}g_1\left(\sqrt{s^2+r^2-2sr\cos w}\right)dw\right] \\
&\qquad\times g_2(r)rdr
\end{aligned}$$

$$= \frac{1}{2\pi}\int_{r=0}^{\infty}\left[\frac{1}{2\pi}\int_{w=0}^{2\pi} g_1(\tau(s,r,w))dw\right] g_2(r)rdr$$
$$= \frac{1}{2\pi}(g_1 \circ g_2)(s)$$
$$= (g_1 \circ g_2)^{\uparrow}(sk_v).$$

E. PROOF OF THEOREM 4

Lemma 14: We have

$$\hat{f}_X(x) = \frac{1}{N}\sum_{n=1}^{N} T_{S_n} L_h(rk_0), \ x = rk_u \in \mathbb{R}^2.$$

Proof: We have

$$L_h(x) = \frac{1}{2\pi h^2}\int_{\rho=0}^{\infty} \frac{\mathcal{F}K(\rho k_0)}{\mathcal{F}f_\varepsilon((\rho/h)k_0)} J(\rho r/h)\rho d\rho$$
$$= \mathcal{F}^{-1}\left[\frac{\mathcal{F}K((h\cdot)k_0)}{\mathcal{F}f_\varepsilon(\cdot k_0)}\right](rk_0),$$

and, by relation (2) in Lemma 1,

$$T_{S_n}\mathcal{F}^{-1}\left[\frac{\mathcal{F}K(h\cdot)}{\mathcal{F}f_\varepsilon(\cdot)}\right](rk_0)$$
$$= \mathcal{F}^{-1}\left[J(\rho S_n)\frac{\mathcal{F}K(h\cdot)}{\mathcal{F}f_\varepsilon(\cdot)}\right]$$
$$= \frac{1}{2\pi}\int_{\rho=0}^{\infty}\frac{\mathcal{F}K(h\rho k_0)}{\mathcal{F}f_\varepsilon(\rho k_0)} J(\rho r)J(\rho S_n)\rho d\rho.$$

It follows that

$$\hat{f}_X(x) = \frac{1}{2\pi}\int_{\rho=0}^{\infty}\frac{\mathcal{F}K(h\rho k_0)\widehat{\Phi}_Y(\rho)}{\mathcal{F}f_\varepsilon(\rho k_0)} J(\rho r)\rho d\rho$$
$$= \frac{1}{N}\sum_{n=1}^{N}\frac{1}{2\pi}\int_{\rho=0}^{\infty}\frac{\mathcal{F}K(h\rho k_0)}{\mathcal{F}f_\varepsilon(\rho k_0)} J(\rho r)J(\rho S_n)\rho d\rho$$
$$= \frac{1}{N}\sum_{n=1}^{N} T_{S_n}\mathcal{F}^{-1}\left[\frac{\mathcal{F}K(h\cdot)}{\mathcal{F}f_\varepsilon(\cdot)}\right](rk_0)$$
$$= \frac{1}{N}\sum_{n=1}^{N} T_{S_n}L_h(rk_0).$$

□

Lemma 15: We have

$$B(x_0) \le M_{42}h^{\alpha-\frac{1}{2}}, \ x_0 \neq 0 \quad \text{and} \quad B(0) \le M_{43}h^{\alpha-1}.$$

Proof: As can be seen in the proof of Lemma 12, the bias term does not depend on the error distribution, and thus the calculations in the proof of Lemma 4 yield the desired upper bounds (see also [16]). □

Lemma 16: Suppose $h \to 0$ as $N \to \infty$. We have, as $N \to \infty$,

$$V(x_0) \le \frac{M_{44}}{Nh^{2\beta+1}}, \ x_0 \neq 0 \quad \text{and} \quad V(0) \le \frac{M_{45}}{Nh^{2\beta+2}}$$

Proof: For $x_0 = r_0 k_{u_0} \in \mathbb{R}^2$, it follows from Lemma 1, Lemma 14, and the Plancherel identity that

$$V(x_0) \le \frac{2\pi}{N}\int_{s=0}^{\infty}[T_s L_h(r_0k_0)]^2 f_Y(sk_0)sds$$
$$\le \frac{2\pi M_{25}}{N}\int_{s=0}^{\infty}[T_s L_h(r_0k_0)]^2 sds$$
$$= \frac{M_{25}}{2\pi N}\int_{\rho=0}^{\infty}|\mathcal{F}L(h\rho k_0)|^2|J(\rho r_0)|^2\rho d\rho.$$

For $x_0 \neq 0$, the inequality $|J(a)| \le \sqrt{\frac{2}{\pi a}}$ for $a > 0$ implies

$$V(x_0) \le \frac{M_{25}}{2\pi N}\frac{2}{\pi r_0}\int_{\rho=0}^{\infty}|\mathcal{F}L(h\rho k_0)|^2 d\rho$$
$$= \frac{C_1}{Nh}\int_{\xi=0}^{\infty}|\mathcal{F}L(\xi k_0)|^2 d\xi$$
$$= \frac{C_1}{Nh}\int_{\xi=0}^{\infty}\frac{|\mathcal{F}K(\xi k_0)|^2}{\mathcal{F}f_\varepsilon((\xi/h)k_0)}d\xi \le \frac{M_{44}}{Nh^{2\beta+1}},$$

where the last inequality follows from the same line of calculations as in the proof of Lemma 13 under (D1), (D2), and (D3). When $x_0 = 0$, the variance upper bound amounts to the upper bound in Lemma 13. To observe this, note

$$V(0) \le \frac{M_{25}}{4\pi^2 N}\int_{\theta=0}^{2\pi}\int_{\rho=0}^{\infty}|\mathcal{F}L(h\rho k_\theta)|^2\rho d\rho d\theta$$
$$= \frac{M_{25}}{4\pi^2 N}\int_{\gamma\in\mathbb{R}^2}|\mathcal{F}L(h\gamma)|^2 d\gamma$$
$$= \frac{M_{25}}{4\pi^2 Nh^2}\int_{\nu\in\mathbb{R}^2}|\mathcal{F}L(\nu)|^2 d\nu$$
$$= \frac{M_{25}}{4\pi^2 Nh^2}\int_{\nu\in\mathbb{R}^2}\frac{|\mathcal{F}K(\nu)|^2}{|\mathcal{F}f_\varepsilon(\gamma/h)|^2}d\nu.$$

Thus, the variance upper bound at $x_0 = 0$ is of order $N^{-1}h^{-(2\beta+2)}$. □

1) PROOF OF THEOREM 4

Lemma 15 and 16 imply that

$$\mathbb{E}\left[\hat{f}_X(x_0) - f_X(x_0)\right]^2 = B^2(x_0) + V(x_0)$$
$$\le M_{42}^2 h^{2\alpha-1} + \frac{M_{44}}{Nh^{2\beta+1}}$$
$$\le C_2 N^{-\frac{2\alpha-1}{2\alpha+2\beta}} + C_3 N^{-1} N^{\frac{2\beta+1}{2\alpha+2\beta}}$$
$$\le M_{12} N^{-\frac{2\alpha-1}{2\alpha+2\beta}}, \ x_0 \neq 0,$$

and

$$\mathbb{E}\left[\hat{f}_X(0) - f_X(0)\right]^2 = B^2(0) + V(0)$$
$$\le M_{43}^2 h^{2\alpha-2} + \frac{M_{45}}{Nh^2}$$
$$\le C_4 N^{-\frac{2\alpha-2}{2\alpha+2\beta}} + C_5 N^{-1} N^{\frac{2\beta+2}{2\alpha+2\beta}}$$
$$\le M_{13} N^{-\frac{2\alpha-2}{2\alpha}}.$$

F. PROOF OF THEOREM 5

We obtain a minimax lower bound for the deconvolution problem in $\mathbb{R}^2$ when X and ε are both symmetric. When $x_0 = 0$, the evaluation functional is homogeneous, and thus the renormalization argument implies that the optimal convergence rate is of order $N^{-\frac{2\alpha-2}{2\alpha+2\beta}}$. Here, we assume that $x_0 = r_0 k_{u_0} \neq 0$ and analyze the effect of the radial symmetry on the minimax convergence rate.

Let ψ be a radial function satisfying the following (see [16]):

(DL-1) $\int_{r=0}^{\infty} \psi(rk_0) r dr = 0$
(DL-2) $\psi(rk_0) = 0$ for $r \notin [0, 1]$
(DL-3) $\|\Delta^{(\alpha-1/2)/2}\psi\|^2 \leq Q^2/4$
(DL-4) $\mathcal{F}\psi(\rho k_0) = 0$ for $\rho \notin [1, \infty)$
(DL-5) $\int_{\rho=1}^{\infty} |\mathcal{F}\psi(\rho k_0)|^2 \rho^{-2\beta} d\rho \leq M_{46}$
(DL-6) $\rho^{\vartheta} \frac{\partial^j}{\partial \rho^j} \mathcal{F}\psi$ are continuous and bounded for $j = 0, 1, 2$ and $\vartheta > 3/2$.

Since we are considering the ordinary smooth case, we assume that

(DL-7) $\left|\frac{\partial^j}{\partial \rho^j} \mathcal{F} f_\varepsilon(\rho k_0)\right| \leq M_{47} \rho^{-\beta-j}$ as $\rho \to \infty$ for $j = 0, 1, 2$

Choose f_0 to be a density in $\mathcal{S}_\alpha(Q/2)$ such that

$$g_0(rk_0) = (f_0 * f_\varepsilon)(rk_0) \geq M_{48} r^{-2\kappa} \quad \text{as} \quad r \to \infty, \tag{7}$$

where $1 < \kappa < 1.5$, and define, for $(r, u) \in [0, 1] \times [0, 2\pi)$,

$$f_1(rk_u) = f_0(rk_0) + M_{49}\delta^{\alpha-3/2} T_{r_0}\psi\left(\frac{r}{\delta}k_0\right),$$

where M_{49} will be determined later. We also define $g_1 = f_1 * f_\varepsilon$. For $M_{50} > 0$, which will be specified later, we choose

$$\delta = \delta_N = M_{50} N^{-\frac{1}{2\alpha+2\beta}}.$$

With this construction, we need the following results to obtain a minimax lower bound using the Le Cam method. The proofs are identical to those of Lemma 10 and Lemma 8, and will be omitted.

Lemma 17: We have

$$|f_0(r_0k_0) - f_1(r_0k_0)| \geq M_{51} N^{-\frac{\alpha-1/2}{2\alpha+2\beta}}$$

Lemma 18: We have $f_0, f_1 \in \mathcal{S}_\alpha(Q)$.

We now obtain an upper bound for the χ^2 divergence between $f_0 * f_\varepsilon$ and $f_1 * f_\varepsilon$.

Lemma 19: Let $\varrho > 0$ be given. Then, for sufficiently large N, we have

$$\chi^2(g_1, g_0) \leq \frac{\varrho}{N}.$$

Proof: Let $\varrho > 0$ be given. By the distributive property of the convolution, we have

$$\begin{aligned}
&\chi^2(g_1, g_0)\\
&= 2\pi \int_{r=0}^{\infty} \frac{(g_1(rk_0) - g_0(rk_0))^2}{g_0(rk_0)} r dr\\
&= 2\pi M_{49}^2 \delta^{2\alpha-3} \int_{r=0}^{\infty} \left[\left(T_{r_0}\psi(\cdot/\delta) * f_\varepsilon\right)(rk_0)\right]^2 g_0^{-1}(rk_0) r dr.
\end{aligned}$$

Note, by the Plancherel identity, Lemma 1, (DL-5), (DL-7), and the inequality $|J(a)| \leq \sqrt{\frac{2}{\pi a}}$ for $0 > 0$,

$$\begin{aligned}
R_1 &\triangleq \int_{r=0}^{1} \left[\left(T_{r_0}\psi(\cdot/\delta) * f_\varepsilon\right)(rk_0)\right]^2 g_0^{-1}(rk_0) r dr\\
&\leq M_{52} \int_{r=0}^{\infty} \left[\left(T_{r_0}\psi(\cdot/\delta) * f_\varepsilon\right)(rk_0)\right]^2 r dr,\\
&\quad \times \left(M_{52} = \max_{0 \leq r \leq 1} g_0^{-1}(rk_0)\right)\\
&= \frac{M_{52}}{4\pi^2} \int_{\rho=0}^{\infty} \left|\mathcal{F}[T_{r_0}\psi(\cdot/\delta)](\rho k_0) \mathcal{F} f_\varepsilon(\rho k_0)\right|^2 \rho d\rho\\
&= \frac{M_{52}}{4\pi^2} \int_{\rho=0}^{\infty} \left|J(r_0\rho)\delta^2 \mathcal{F}\psi(\delta\rho k_0) \mathcal{F} f_\varepsilon(\rho k_0)\right|^2 \rho d\rho\\
&\leq \frac{2M_{52}}{4\pi^3 r_0} \int_{\xi=1}^{\infty} \left|\delta^2 \mathcal{F}\psi(\xi k_0) \mathcal{F} f_\varepsilon((\xi/\delta)k_0)\right|^2 \frac{d\xi}{\delta}\\
&\leq \frac{2M_{52}M_{47}\delta^{2\beta+3}}{4\pi^3 r_0} \int_{\xi=1}^{\infty} |\mathcal{F}\psi(\xi k_0)|^2 \xi^{-2\beta} d\xi\\
&\leq M_{53}\delta^{2\beta+3}.
\end{aligned} \tag{8}$$

Observe that a change of variables $r \mapsto \delta r$ and $s \mapsto \delta s$ implies

$$\begin{aligned}
R_2 &\triangleq \int_{r=1}^{\infty} \left[\left(T_{r_0}\psi(\cdot/\delta) * f_\varepsilon\right)(rk_0)\right]^2 g_0^{-1}(rk_0) r dr\\
&= \int_{r=1}^{\infty} \left[\int_{s=0}^{\infty} \int_{v=0}^{2\pi} T_{r_0}\psi\left(\frac{rk_u - sk_v}{\delta}\right) f_\varepsilon(sk_v) s ds dv\right]^2\\
&\quad \times g_0^{-1}(rk_0) r dr\\
&= \delta^2 \int_{r=1/\delta}^{\infty} \left[\left(T_{r_0}\psi * \delta^2 f_\varepsilon(\delta\cdot)\right)(rk_0)\right]^2 g_0^{-1}(\delta rk_0) r dr.
\end{aligned}$$

Under (DL-6) and (DL-7), we have, for any $\rho \in [1, \infty)$,

$$|\Delta\varphi_\delta(\rho k_0)| \leq M_{54}\delta^{\beta}\rho^{-\vartheta} \quad \text{as} \quad \delta \to 0,$$

where $\varphi_\delta(\rho k_\theta) = J(\rho r_0)\mathcal{F}\psi(\rho k_0)\mathcal{F} f_\varepsilon((\rho/\delta)k_0))$ for $(\rho, \theta) \in [1, \infty) \times [0, 2\pi)$. It follows from (DL-6) and (7) that

$$\begin{aligned}
R_2 &= \delta^2 \int_{r=1/\delta}^{\infty} \left[\left(T_{r_0}\psi * \delta^2 f_\varepsilon(\delta\cdot)\right)(rk_0)\right]^2 g_0^{-1}(\delta rk_0) r dr\\
&= \delta^2 \int_{r=1/\delta}^{\infty} \left[-\frac{1}{r^2}\mathcal{F}^{-1}\Delta\varphi_\delta(rk_0)\right]^2 g_0^{-1}(\delta rk_0) r dr\\
&\leq \frac{M_{54}^2\delta^{2\beta+5}}{2\pi^3} \int_{s=1}^{\infty} \frac{1}{s^4}\left[\int_{\rho=1}^{\infty} \rho^{1/2-\vartheta} d\rho\right] g_0^{-1}(sk_0) ds\\
&\leq M_{55}\delta^{2\beta+5}.
\end{aligned} \tag{9}$$

By (8) and (9), we have

$$\chi^2(g_1, g_0) \leq 2\pi M_{49}^2 \delta^{2\alpha-3}(R_1 + R_2) \leq M_{56}\delta^{2\alpha+2\beta} \leq \frac{\varrho}{N},$$

provided that we choose M_{49} small enough. □

1) PROOF OF THEOREM 5

The Le Cam method with Lemma 17, Lemma 18, Lemma 19 provides a lower bound for $x_0 \neq 0$:

$$\inf_{T} \sup_{f_X \in \mathcal{S}_\alpha(Q)} \mathbb{E}\left[N^{\frac{2\alpha-1}{2\alpha+2\beta}} |T(x_0) - f_X(x_0)|^2\right] \geq M_{57},$$

as $N \to \infty$. As discussed above, a lower bound for $x_0 = 0$ can be determined by the renormalization argument.

ACKNOWLEDGMENT

This article is based in part on the Kwan-Young Bak's Ph.D. dissertation.

REFERENCES

[1] K. Fang and Y. T. Zhang, *Generalized Multivariate Analysis*. Beijing, China: Science Press, 1990.
[2] R. F. W. Bader, *An Introduction to the Electronic Structure of Atoms and Molecules*. Toronto, ON, Canada: Clarke & Irwin, 1970.
[3] J. P. Nolan, "Fitting data and assessing goodness-of-fit with stable distributions," in *Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*. Washington, DC, USA: American Univ., 1999.
[4] J. S. Jaffe, "Three-dimensional probability density functions via tomographic inversion," *SIAM J. Appl. Math.*, vol. 65, no. 5, pp. 1506–1525, Jan. 2005.
[5] S. F. Huckemann, P. T. Kim, J.-Y. Koo, and A. Munk, "Möbius deconvolution on the hyperbolic plane with application to impedance density estimation," *Ann. Statist.*, vol. 38, no. 4, pp. 2465–2498, 2010.
[6] L. R. Haff, P. T. Kim, J.-Y. Koo, and D. S. P. Richards, "Minimax estimation for mixtures of Wishart distributions," *Ann. Statist.*, vol. 39, no. 6, pp. 3417–3440, Dec. 2011.
[7] L. LeCam, "Convergence of estimates under dimensionality restrictions," *Ann. Statist.*, vol. 1, no. 1, pp. 38–53, Jan. 1973.
[8] G. Geenens, "Mellin–Meijer kernel density estimation on $\mathbb{R}^+$," *Ann. Inst. Stat. Math.*, vol. 73, no. 5, pp. 953–977, 2021.
[9] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, vol. 27, no. 3, pp. 832–837, Sep. 1956.
[10] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.
[11] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Boca Raton, FL, USA: CRC Press, 1986.
[12] M. P. Wand and M. C. Jones, *Kernel Smoothing*. Boca Raton, FL, USA: CRC Press, 1994.
[13] G. S. Watson and M. R. Leadbetter, "On the estimation of the probability density, I," *Ann. Math. Statist.*, vol. 34, no. 2, pp. 480–491, Jun. 1963.
[14] P. J. Diggle and P. Hall, "A Fourier approach to nonparametric deconvolution of a density estimate," *J. Roy. Stat. Soc. B, Methodol.*, vol. 55, no. 2, pp. 523–531, Jan. 1993.
[15] J. Fan, "Global behavior of deconvolution kernel estimates," *Statistica Sinica*, vol. 1, no. 2, pp. 541–551, Jul. 1991.
[16] J. Fan, "On the optimal rates of convergence for nonparametric deconvolution problems," *Ann. Statist.*, vol. 19, no. 3, pp. 1257–1272, Sep. 1991.
[17] J. Fan and J.-Y. Koo, "Wavelet deconvolution," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 734–747, Mar. 2002.
[18] J.-Y. Koo, "Optimal rates of convergence for nonparametric statistical inverse problems," *Ann. Statist.*, vol. 21, no. 2, pp. 590–599, Jun. 1993.
[19] B. A. Mair and F. H. Ruymgaart, "Statistical inverse estimation in Hilbert scales," *SIAM J. Appl. Math.*, vol. 56, no. 5, pp. 1424–1444, Oct. 1996.
[20] C.-H. Zhang, "Fourier methods for estimating mixing densities and distributions," *Ann. Statist.*, vol. 18, no. 2, pp. 806–831, Jun. 1990.
[21] A. Meister, *Deconvolution Problems in Nonparametric Statistics*. Berlin, Germany: Springer, 2009.
[22] I. N. Sneddon, *Fourier Transforms*. North Chelmsford, MA, USA: Courier Corporation, 1995.
[23] A. Terras, *Harmonic Analysis on Symmetric Spaces—Euclidean Space, the Sphere, and the Poincaré Upper Half-Plane*. New York, NY, USA: Springer, 2013.
[24] G. N. Watson, *A Treatise on the Theory of Bessel Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1995.
[25] D. L. Donoho and M. G. Low, "Renormalization exponents and optimal pointwise rates of convergence," *Ann. Statist.*, vol. 20, no. 2, pp. 944–970, Jun. 1992.
[26] B. Yu, "Assouad, Fano, and Le Cam," in *Festschrift for Lucien Le Cam*. New York, NY, USA: Springer, 1997, pp. 423–435.
[27] A. B. Tsybakov, *Introduction to Nonparametric Estimation*. New York, NY, USA: Springer, 2009.
[28] A. Terras, *Harmonic Analysis on Symmetric Spaces—Higher Rank Spaces, Positive Definite Matrix Space and Generalizations*. New York, NY, USA: Springer, 2016.
[29] C. Butucea, "Exact adaptive pointwise estimation on Sobolev classes of densities," *ESAIM, Probab. Statist.*, vol. 5, pp. 1–31, Aug. 2001.

KWAN-YOUNG BAK received the B.S. and Ph.D. degrees in statistics from Korea University, in 2017 and 2021, respectively. He was a Research Professor with the Department of Statistics, Korea University. He is currently an Assistant Professor with the School of Mathematics, Statistics and Data Science, Sungshin Women's University. His research interests include function estimation and minimax theory.

JA-YONG KOO received the B.S. degree in statistics from Seoul National University, in 1984, and the Ph.D. degree in statistics from the University of California at Berkeley, Berkeley, in 1988. He is currently a Professor with the Department of Statistics, Korea University. His research interests include function estimation and data mining.