

Applications of Mathematics

Le Thi Hong Thuy; Cao Xuan Phuong

Density deconvolution with associated stationary data

Applications of Mathematics, Vol. 68 (2023), No. 5, 685–708

Persistent URL: <http://dml.cz/dmlcz/151839>

Terms of use:

© Institute of Mathematics AS CR, 2023

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

DENSITY DECONVOLUTION WITH ASSOCIATED STATIONARY DATA

LE THI HONG THUY, CAO XUAN PHUONG, Ho Chi Minh City

Received June 12, 2022. Published online September 3, 2023.

Abstract. We study the density deconvolution problem when the random variables of interest are an associated strictly stationary sequence and the random noises are i.i.d. with a nonstandard density. Based on a nonparametric strategy, we introduce an estimator depending on two parameters. This estimator is shown to be consistent with respect to the mean integrated squared error. Under additional regularity assumptions on the target function as well as on the density of noises, some error estimates are derived. Several numerical simulations are also conducted to illustrate the efficiency of our method.

Keywords: associated process; density deconvolution; nonstandard noise density

MSC 2020: 62G05, 62G07, 62G20

1. INTRODUCTION

We consider the model

$$(1.1) \quad Y_j = X_j + \varepsilon_j, \quad j \in \mathbb{N}$$

with the following basic assumptions:

- (A1) $\{X_j\}_{j \in \mathbb{N}}$ is a strictly stationary sequence of random variables with a common unknown density $f_X \in L^2(\mathbb{R})$.
- (A2) $\{\varepsilon_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. random variables with a known density f_ε .
- (A3) $\{X_j\}_{j \in \mathbb{N}}$ and $\{\varepsilon_j\}_{j \in \mathbb{N}}$ are independent sequences.

Suppose we only observe n random variables $Y_1, \dots, Y_n$ of the process $\{Y_j\}_{j \in \mathbb{N}}$. Based on the observations as well as on the complete knowledge about f_ε , we aim to estimate the density f_X in a nonparametric strategy.

The process $\{\varepsilon_j\}_{j \in \mathbb{N}}$ is usually referred to as random noises. Hence, the density f_ε is called a noise density. Under the assumptions, $\{Y_j\}_{j \in \mathbb{N}}$ is also a strictly stationary

process. Let f_Y be the common density of Y_j 's. It follows from the assumptions that $f_Y = f_X * f_\varepsilon$, where the notation $*$ stands for the usual convolution on $\mathbb{R}$. Therefore, the problem of estimating f_X from the observations $Y_1, \dots, Y_n$ and the density f_ε is called a density deconvolution problem. This is a statistical inverse problem.

In the setting of independent $\{X_j\}_{j \in \mathbb{N}}$, the problem has been widely and intensively studied in a large amount of papers, such as Carroll and Hall [5], Liu and Taylor [19], Stefanski and Carroll [30], Fan [13], Pensky and Vidakovic [28], Comte et al. [8], Hall and Meister [16], Butucea and Tsybakov [3] and many others. Obviously, the independent assumption on X_j 's can be unrealistic for some applications. As an example, we mention from van Es et al. [31] a continuous-time stochastic volatility model of the form $dS_t = \sigma_t dW_t$ for $t \geq 0$ ($S_0 = 0$), where $\{S_t\}_{t \geq 0}$ denotes the log-price process of some stock on a financial market, $\{W_t\}_{t \geq 0}$ is a standard Brownian motion and $\{\sigma_t^2\}_{t \geq 0}$ is a strictly stationary, predictable process which is independent of $\{W_t\}_{t \geq 0}$. Under observed discrete-time data $S_\Delta, S_{2\Delta}, \dots, S_{n\Delta}$, one can get $Y_j^\Delta \approx X_j^\Delta + \varepsilon_j^\Delta$ for small Δ , where $Y_j^\Delta := \ln((S_{j\Delta} - S_{(j-1)\Delta})^2 / \Delta)$, $X_j^\Delta := \ln(\sigma_{(j-1)\Delta}^2)$ and $\varepsilon_j^\Delta := \ln((W_{j\Delta} - W_{(j-1)\Delta})^2 / \Delta)$, $j = \overline{1, n}$. The independence and stationarity of Brownian increments give that $\{\varepsilon_j^\Delta\}_{j=\overline{1, n}}$ is an independent strictly stationary sequence. Moreover, the sequence is independent of $\{X_j^\Delta\}_{j=\overline{1, n}}$, which is a strictly stationary sequence. Thus, the problem of estimating the density of $\ln(\sigma_{i\Delta}^2)$ reduces to a density deconvolution problem for dependent data.

In contrast to the setting of independent $\{X_j\}_{j \in \mathbb{N}}$, there is not much research in the setting of dependent $\{X_j\}_{j \in \mathbb{N}}$. We refer to Masry [20], Masry [21] and Masry [22] as the first related works for the problem with dependent $\{X_j\}_{j \in \mathbb{N}}$. These papers addressed some asymptotic properties of a kernel type estimator under various mixing assumptions on the sequence $\{X_j\}_{j \in \mathbb{N}}$. As a continuation of the three latter papers, Masry [23] considered the setting where $\{X_j\}_{j \in \mathbb{N}}$ is positively or negatively associated in the sense of Esary et al. [12]. Another notable research is the work of Comte et al. [7], where the authors studied the performance of an adaptive penalized contrast estimator when the sequence $\{X_j\}_{j \in \mathbb{N}}$ is absolutely regular in the sense of Volkonskii and Rozanov [33] or when it is τ -dependent in the sense of Dedecker and Prieur [9]. Finally, we refer to the works of van Zanten and Zareba [32] and Chesneau [6], where the performance of wavelet type estimators were investigated under strongly mixing conditions on the sequence $\{Y_j\}_{j \in \mathbb{N}}$.

Let φ_ε be the characteristic function with respect to the noise density f_ε , i.e. $\varphi_\varepsilon(t) := \int_{-\infty}^{\infty} f_\varepsilon(x) e^{itx} dx$ for $t \in \mathbb{R}$ and $i = \sqrt{-1}$. As mentioned in Stefanski and Carroll [30], the difficulty of the density deconvolution problem depends strongly on the decaying rate of φ_ε , which is closely related to the smoothness of f_ε . Roughly speaking, the problem is more difficult when f_ε is smoother. There are two major types of f_ε which have been received much attention in the literature: ordinary

smooth and supersmooth (see Fan [13]). In particular, it is usually assumed that

$$(1.2) \quad \varphi_\varepsilon^{-1}(\{0\}) = \emptyset,$$

where $\varphi_\varepsilon^{-1}(\{0\}) := \{t \in \mathbb{R} : \varphi_\varepsilon(t) = 0\}$ is the set of all possible zeros of φ_ε . Assumption (1.2) is rather standard in the field of density deconvolution. It is satisfied for many usual densities, e.g., Gaussian, Cauchy, Laplace, chi-square densities, etc. Of course, there are still some densities that do not fulfill the assumption, for instance, uniform densities, convolution of a uniform density with an arbitrary density, or compactly supported densities in general. Throughout this paper, the density f_ε will be said to be standard if it satisfies (1.2); otherwise, it is called nonstandard.

In the setting of independent $\{X_j\}_{j \in \mathbb{N}}$, the density deconvolution problem with nonstandard f_ε has been studied by some researchers, such as Devroye [11], Groeneboom and Jongbloed [15], Hall and Meister [16], Meister [24], Delaigle and Meister [10], Carrasco and Florens [4]. However, to the best of our knowledge, the problem with dependent $\{X_j\}_{j \in \mathbb{N}}$ and nonstandard f_ε has not yet been studied in any work. Thus, our aim in the present paper is to partially fill this gap. More concretely, we consider the problem of estimating f_X when the underlying process $\{X_j\}_{j \in \mathbb{N}}$ is positively associated in the sense of Esary et al. [12]; in addition, the noise density f_ε is assumed to be compactly supported in a bounded interval. It is emphasized that the just mentioned assumption on f_ε is appropriate in circumstances that the ranges of ε_j 's are restricted to a bounded domain. Despite the fact that Masry [23] also considered the setting of positively associated $\{X_j\}_{j \in \mathbb{N}}$, the results in the paper were only established in the ordinary smooth and supersmooth cases of f_ε , which are of course not the nonstandard cases of f_ε .

The rest of the present paper is organized as follows. In Section 2, we derive our estimator. In Section 3, we investigate some asymptotic properties of our estimator, including consistency and rates of convergence. In Section 4, we perform a numerical experiment to examine the convergence of the proposed estimator according to the sample size n . All proofs are presented in Section 5.

2. THE ESTIMATOR

We begin this section by introducing some notations and notions. We define the characteristic function of a random variable U by $\varphi_U(t) := \mathbb{E}(e^{itU})$, $t \in \mathbb{R}$. Here i is the imaginary unit, i.e., $i^2 = -1$. For $u \in L^p(\mathbb{R})$ ($p \in \{1, 2\}$), the function $\varphi_u(t) := \int_{-\infty}^{\infty} u(x)e^{itx} dx$ ($t \in \mathbb{R}$) is called the Fourier transform of u . As known, if $\varphi_u \in L^2(\mathbb{R})$, then one has the Fourier inversion formula $u(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} \varphi_u(t)e^{-itx} dt$ for almost all $x \in \mathbb{R}$; moreover, $\|\varphi_u\|^2 = 2\pi\|u\|^2$, which is called the Parseval identity. The

convolution of two functions u and v is defined by $(u * v)(x) := \int_{-\infty}^{\infty} u(x-s)v(s) ds$. For $a_n, b_n > 0$ depending on the sample size n , the notation $a_n \lesssim b_n$ means that there is a positive constant C independent of n such that $a_n \leq Cb_n$. The notation $\Re\{z\}$ denotes the real part of a complex number z . Finally, the notation $\text{Cov}(U, V)$ stands for the covariance of two random variables U and V .

To ensure the identifiability of our problem, i.e., the unique reconstruction of f_X from the observation density f_Y , we need the following assumption:

(A4) $\lambda(\varphi_\varepsilon^{-1}(\{0\})) = 0$, where λ is the Lebesgue measure on $\mathbb{R}$.

As indicated in Meister [25], page 24, the identifiability is a necessary condition for the existence of consistent estimators of f_X . It is obvious that (A4) is weaker than (1.2).

Now we estimate f_X . Let φ_Y and φ_X be the characteristic functions of Y_j 's and X_j 's, respectively. To derive an estimator of f_X , we first estimate the unknown function φ_X and then apply a Fourier inversion procedure on proposed estimator of φ_X . For the first step, we begin from the relation $\varphi_Y = \varphi_X \varphi_\varepsilon$ obtained by model (1.1) and assumptions (A1)–(A3). The relation gives

$$(2.1) \quad \varphi_X(t) = \frac{\varphi_Y(t)}{\varphi_\varepsilon(t)} \quad \forall t \notin \varphi_\varepsilon^{-1}(\{0\}).$$

For the identically distributed observations Y_j 's, we can estimate $\varphi_Y(t)$ at every $t \in \mathbb{R}$ by the so-called empirical characteristic function $\widehat{\varphi}_{Y;n}(t) := n^{-1} \sum_{j=1}^n e^{itY_j}$. In doing so, in view of (2.1), $\varphi_X(t)$ can be estimated by

$$\widehat{\varphi}_{X;n}(t) := \frac{\widehat{\varphi}_{Y;n}(t)}{\varphi_\varepsilon(t)} \quad \forall t \notin \varphi_\varepsilon^{-1}(\{0\}).$$

However, since $\lim_{t \rightarrow s} \widehat{\varphi}_{X;n}(t) = \infty$ for any $s \in \{\pm\infty\}$ or $s \in \varphi_\varepsilon^{-1}(\{0\})$ (if $\varphi_\varepsilon^{-1}(\{0\}) \neq \emptyset$), the integrability or square integrability of $\widehat{\varphi}_{X;n}$ on the whole real line cannot be guaranteed. Hence, we must regularize $\widehat{\varphi}_{X;n}$ before the Fourier inversion procedure is applied.

To this end, we recall in the context of independent $\{X_j\}_{j \in \mathbb{N}}$ that, to avoid dividing by numbers too close to zero, Hall and Meister [16] proposed to replace $\widehat{\varphi}_{X;n}(t)$ by

$$\widehat{\varphi}_{X;n,\zeta,\varrho,r}^{\text{HM}}(t) := \frac{\varphi_\varepsilon(-t)|\varphi_\varepsilon(t)|^r \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|; n^{-\zeta}|t|^\varrho\}^{r+2}}$$

with $r, \varrho \geq 0$ and $\zeta > 0$. The function $n^{-\zeta}|t|^\varrho$ is called a ridge parameter function and the method of using $\widehat{\varphi}_{X;n,\zeta,\varrho,r}^{\text{HM}}(t)$ is known as the ridge parameter regularization.

In the context of dependent $\{X_j\}_{j \in \mathbb{N}}$, it is possible to use $\widehat{\varphi}_{X;n,\zeta,\varrho,r}^{\text{HM}}(t)$ to estimate $\varphi_X(t)$. However, instead of using the ridge parameter function, we use a more general

ridge parameter function in the form $\delta|t|^a$ with $\delta > 0$ and $a > \frac{1}{2}$; in addition, for simplicity, we take $r = 0$. In other words, we can estimate $\varphi_X(t)$ by

$$\tilde{\varphi}_{X;n,\delta,a}(t) := \frac{\varphi_\varepsilon(-t)\hat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}}.$$

It is emphasized that the estimated object φ_X is a characteristic function, so it is bounded by 1. For this reason, we finally propose to estimate $\varphi_X(t)$ for every $t \in \mathbb{R}$ by

$$\hat{\varphi}_{X;n,\delta,a}(t) := \frac{\tilde{\varphi}_{X;n,\delta,a}(t)}{\max\{|\tilde{\varphi}_{X;n,\delta,a}(t)|; 1\}}.$$

For $a > \frac{1}{2}$, $\hat{\varphi}_{X;n,\delta,a}$ belongs to $L^2(\mathbb{R})$ almost surely. From there, we apply the inverse Fourier transform for $\hat{\varphi}_{X;n,\delta,a}$ to derive the final estimator of f_X in the form

$$(2.2) \quad \hat{f}_{X;n,\delta,a}(x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\varphi}_{X;n,\delta,a}(t) e^{-itx} dt.$$

Remark 2.1. It can be seen that $\int_{-\infty}^{\infty} \hat{f}_{X;n,\delta,a}(x) dx = \hat{\varphi}_{X;n,\delta,a}(0) = 1$. The estimator $\hat{f}_{X;n,\delta,a}$ at (2.2) has a disadvantage: it may take some complex values. However, this is of minor importance, because we can always use the absolute value function $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ of the real part function $\Re\{\hat{f}_{X;n,\delta,a}\}$ whose risk, quantified by $\mathbb{E}\| |\Re\{\hat{f}_{X;n,\delta,a}\}| - f_X \|^2$, is smaller than or equal to the risk of $\hat{f}_{X;n,\delta,a}$.

3. CONVERGENCE RESULTS

In this section, our aim is to understand how well the estimator $\hat{f}_{X;n,\delta,a}$ approximates f_X . We use the mean integrated squared error (MISE) $\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2$ to evaluate the accuracy of $\hat{f}_{X;n,\delta,a}$. The error type was not mentioned in Masry [23]. First, the following proposition gives a general upper bound for the MISE.

Proposition 3.1. Consider model (1.1) under assumptions (A1)–(A4). Then for $\delta > 0$ and $a > \frac{1}{2}$ we derive

$$(3.1) \quad \mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \leq \frac{1}{2\pi} I_{1;\delta,a} + \frac{1}{2\pi n} I_{2;\delta,a} + \frac{1}{\pi n} I_{3;n,\delta,a},$$

where

$$\begin{aligned} I_{1;\delta,a} &:= \int_{-\infty}^{\infty} \left(1 - \frac{|\varphi_\varepsilon(t)|^2}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}}\right)^2 |\varphi_X(t)|^2 dt, \\ I_{2;\delta,a} &:= \int_{-\infty}^{\infty} \frac{(1 - |\varphi_X(t)|^2) |\varphi_\varepsilon(t)|^2 |\varphi_\varepsilon(t)|^2}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}^2} dt, \\ I_{3;n,\delta,a} &:= \int_{-\infty}^{\infty} \frac{|\varphi_\varepsilon(t)|^4}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}^2} \sum_{j=2}^n |\text{Cov}(e^{itX_1}, e^{itX_j})| dt. \end{aligned}$$

Estimate (3.1) is nonasymptotic. It depends on the dependence structure of the data $Y_1, \dots, Y_n$. The term $I_{1;\delta,a}$ comes from the integrated square bias of $\hat{f}_{X;n,\delta,a}$, while the terms $I_{2;\delta,a}$ and $I_{3;n,\delta,a}$ come from the integrated variance of $\hat{f}_{X;n,\delta,a}$. For every $a > \frac{1}{2}$, $I_{1;\delta_1,a} \leq I_{1;\delta_2,a}$ if $0 < \delta_1 \leq \delta_2$; however, $I_{2;\delta_1,a} \geq I_{2;\delta_2,a}$ and $I_{3;n,\delta_1,a} \geq I_{3;n,\delta_2,a}$ if $0 < \delta_1 \leq \delta_2$. In other words, when δ is smaller, $I_{1;\delta,a}$ becomes smaller, but $I_{2;\delta,a}$ and $I_{3;n,\delta,a}$ become bigger. Note that the term $I_{3;n,\delta,a}$ disappears if the observations Y_j 's are independent.

To derive a further estimate on the MISE, we will assume that the process $\{X_j\}_{j \in \mathbb{N}}$ is positively associated or associated for short. This terminology was first introduced in Esary et al. [12] as an extension of the bivariate notion of positive quadrant dependence in Lehmann [17]. For convenience, we restate this notion in the following definition.

Definition 3.2 (see Esary et al. [12]). Random variables $U_1, \dots, U_n$ are said to be *positively associated*, or just *associated*, if for every pair of coordinatewise nondecreasing functions $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ there holds $\text{Cov}(f(U_1, \dots, U_n), g(U_1, \dots, U_n)) \geq 0$ whenever the covariance is defined. A sequence $\{V_j\}_{j \in \mathbb{N}}$ of random variables is *associated* if for every $n \in \mathbb{N}$, the family of variables $V_1, \dots, V_n$ is associated.

Homogeneous Markov chains are typical examples of associated processes. Associated processes have found many applications in many fields, such as in physics (see Fortuin et al. [14]), finance (see Pan [27]), and especially in reliability theory (see Bagai and Rao [1]). For more details on the associated processes, we refer the readers to Oliveira [26].

Remark 3.3. From assumption (A2) and the fact that independent random variables are associated (see Esary et al. [12], Theorem 2.1), we conclude that the process $\{\varepsilon_j\}_{j \in \mathbb{N}}$ is associated. Also, it was shown in the latter paper that if two sets of associated random variables are independent of one another, then their union is a set of associated random variables (Property (P2)); in addition, nondecreasing functions of associated random variables are associated (Property (P4)). These assertions and assumption (A3) imply that if the process $\{X_j\}_{j \in \mathbb{N}}$ is associated, then the process $\{Y_j\}_{j \in \mathbb{N}}$ is also associated.

Next we consider the assumption:

(A5) $\{X_j\}_{j \in \mathbb{N}}$ is an associated process such that $\sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) < \infty$.

The summable condition $\sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) < \infty$ is not restrictive. For example, if $\{W_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. random variables with finite variance and $X_j = W_j + W_{j+1}$ for $j \in \mathbb{N}$, then $\{X_j\}_{j \in \mathbb{N}}$ is an associated process with $\sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) =$

$\text{Var}(W_2) < \infty$. Another example is the case where $X_j = \sum_{k=0}^{\infty} (\sqrt{3}/2^{k+1})U_{j-k}$ for $j \in \mathbb{N}$, where $\{U_p\}_{p \in \mathbb{Z}}$ is a sequence of i.i.d. Gaussian random variables with zero mean and unit variance. This linear process $\{X_j\}_{j \in \mathbb{N}}$ is associated with $\text{Cov}(X_1, X_j) = 2^{1-j}$ for all $j \geq 2$, so $\sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) = 1$. In Masry [23], the author used a similar condition of the form $\sum_{j=2}^{\infty} j^{\theta} \text{Cov}(X_1, X_j) < \infty$, for some $\theta > 0$. Obviously, our summable condition is weaker than the one of the latter paper.

Using (A5) and the previous assumptions, we derive a more explicit upper bound for the MISE.

Proposition 3.4. *Consider model (1.1) under assumptions (A1)–(A5). Then for $\delta > 0$ and $a > \frac{3}{2}$,*

$$(3.2) \quad \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \leq \frac{1}{2\pi} I_{1;\delta,a} + \frac{1}{2\pi n} I_{2;\delta,a} + \frac{4 \sum_{j=2}^{\infty} \text{Cov}(X_1, X_j)}{\pi n} \tilde{I}_{3;\delta,a},$$

where $I_{1;\delta,a}$, $I_{2;\delta,a}$ are as in Proposition 3.1, and

$$\tilde{I}_{3;\delta,a} := \int_{-\infty}^{\infty} \frac{t^2 |\varphi_{\varepsilon}(t)|^4}{\max\{|\varphi_{\varepsilon}(t)|^2; \delta |t|^a\}^2} dt.$$

Note that the constraint $a > \frac{3}{2}$ in the latter proposition ensures the finiteness of $I_{1;\delta,a}$, $I_{2;\delta,a}$ and $\tilde{I}_{3;\delta,a}$.

We are now ready to state a consistency result of the estimator $\hat{f}_{X;n,\delta,a}$.

Theorem 3.5. *Let the assumptions of Proposition 3.4 hold. If δ depends on n in such a way that $\lim_{n \rightarrow \infty} \delta = 0$ and $\lim_{n \rightarrow \infty} n\delta^2 = \infty$, then*

$$\lim_{n \rightarrow \infty} \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 = 0.$$

In the sequel, we will establish some error estimates under further assumptions on f_X and f_{ε} . First, concerning f_X , we consider the assumption:

(A6) There exist $\alpha > \frac{1}{2}$ and $L > 0$ such that $\int_{-\infty}^{\infty} |\varphi_X(t)|^2 (1+t^2)^{\alpha} dt \leq L$.

Assumption (A6) is known as the Sobolev condition on f_X . It implies that f_X has derivatives up to order l , where l is the largest integer number satisfying $l < \alpha - \frac{1}{2}$. That explains why the number α is usually called smoothness degree of f_X . This assumption is satisfied for many common univariate densities, such as Gaussian, Cauchy, Laplace, Gamma, chi-square densities and many others. Note that assumption (A6) has been mentioned in many papers in nonparametric deconvolution topics. See, for instance, Pensky and Vidakovic [28], Hall and Meister [16], Comte et al. [7] and many others.

Next, regarding the density f_ε , we introduce the following assumption:

(A7) There exists a constant $M > 0$ such that $f_\varepsilon(x) = 0$ for all $x \notin [-M, M]$.

Assumption (A7) means that the density f_ε is compactly supported in $[-M, M]$. Typical examples for the assumption are the uniform and beta densities. Note also that the function φ_ε under (A7) can be extended into an entire function (i.e., an analytic function on the whole complex plane) of exponential type. This implies that possible zeros of φ_ε are all isolated, so φ_ε automatically satisfies (A4).

Remark 3.6. If $\varphi_\varepsilon \in L^2(\mathbb{R})$, then by the Paley-Wiener theorem (see, e.g., Rudin [29]), assumption (A7) is equivalent to the assumption that $|\varphi_\varepsilon(z)| \leq e^{M|z|}$, for all $z \in \mathbb{C}$.

The following theorem establishes an upper bound on the convergence rate of the MISE.

Theorem 3.7. Consider model (1.1) under assumptions (A1)–(A3), (A5)–(A7). Let $a > \frac{3}{2}$. Choosing $\delta = n^{-l}$ with $0 < l < \frac{1}{2}$, we obtain

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \left(\frac{\ln n}{\ln(\ln n)}\right)^{-2\alpha}.$$

In Theorem 3.7, the convergence rate depends on the sample size n and is improved with increasing n . Also, the rate becomes faster when the smoothness degree α is larger, i.e., the target density f_X is smoother. It is emphasized that the selection of δ does not depend on the parameters α, L related to the unknown density f_X , so the estimator $\hat{f}_{X;n,\delta,a}$ is fully data-driven, meaning that it can be computed without any knowledge concerning the parameters.

Although our main attention is focused on the case of compactly supported f_ε , to convince that our estimator is well chosen, we need to show that it also achieves some classical optimal rates of convergence under some typical standard cases of f_ε . For that purpose, we introduce the following assumptions:

(A8) There exist constants $c_2 \geq c_1 > 0$ and $\beta > 0$ such that

$$c_1(1 + |t|)^{-\beta} \leq |\varphi_\varepsilon(t)| \leq c_2(1 + |t|)^{-\beta} \quad \forall t \in \mathbb{R}.$$

(A9) There exist constants $k_2 \geq k_1 > 0$ and $d, \gamma > 0$ such that

$$k_1 e^{-d|t|^\gamma} \leq |\varphi_\varepsilon(t)| \leq k_2 e^{-d|t|^\gamma} \quad \forall t \in \mathbb{R}.$$

Assumptions (A8) and (A9), first introduced by Fan [13], have become very common in the field of nonparametric deconvolution. Densities satisfying (A8) or (A9) are

called ordinary smooth densities of order β or supersmooth of order γ , respectively. For example, the Laplace and Gamma densities are ordinary smooth, while the Gaussian, Cauchy, t and logistic densities are supersmooth. Under the assumptions, $\varphi_\varepsilon^{-1}(\{0\}) = \emptyset$ and hence φ_ε still satisfies assumption (A4).

Theorem 3.8. *Consider model (1.1) under assumptions (A1)–(A6). Let $a > \frac{3}{2}$.*

(a) *Let (A8) hold. Choosing*

$$\delta = \begin{cases} n^{-(a+2\beta)/(2\alpha+2\beta+1)} & \text{if } \beta \geq 1, \\ n^{-(a+2\beta)/(2\alpha+3)} & \text{if } 0 < \beta < 1, \end{cases}$$

we obtain

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \begin{cases} n^{-2\alpha/(2\alpha+2\beta+1)} & \text{if } \beta \geq 1, \\ n^{-2\alpha/(2\alpha+3)} & \text{if } 0 < \beta < 1. \end{cases}$$

(b) *Let (A9) hold. Choosing*

$$\delta = k_1^2 \left(\frac{\ln n - (\ln(\ln n))^2}{2d} \right)^{-a/\gamma} e^{-\ln n + (\ln(\ln n))^2},$$

we obtain

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim (\ln n)^{-2\alpha/\gamma}.$$

We have some comments on the results of Theorem 3.8 as follows:

- ▷ In part (a), where f_ε is ordinary smooth of order β , we obtain the rate $n^{-2\alpha/(2\alpha+2\beta+1)}$ with $\beta \geq 1$. This rate coincides with the classical optimal rate of deconvolution estimators in the setting of independent $\{X_j\}_{j \in \mathbb{N}}$ (see, e.g., Pensky and Vidakovic [28], Comte et al. [8]). Also, in comparison to Comte et al. [7], where the process $\{X_j\}_{j \in \mathbb{N}}$ is assumed to be absolutely regular or τ -dependent, this rate is the same as the rate of their penalized contrast estimator. For $0 < \beta < 1$, our rate $n^{-2\alpha/(2\alpha+3)}$ is slower than the classical optimal rate. This can be viewed as a payment for the case of associated $\{X_j\}_{j \in \mathbb{N}}$. It is noted that, in the latter paper, the integrability of φ_ε is required. That implies the authors only derived the rate $n^{-2\alpha/(2\alpha+2\beta+1)}$ for the case $\beta > 1$.
- ▷ In part (b), where f_ε is supersmooth of order γ , we obtain the logarithm rate $(\ln n)^{-2\alpha/\gamma}$, which gives that $\hat{f}_{X;n,\delta,a}$ converges extremely to f_X . However, this rate is the same as the optimal rate derived in the settings of independent $\{X_j\}_{j \in \mathbb{N}}$ (see, e.g., Pensky and Vidakovic [28], Comte et al. [8]) as well as of absolutely regular or τ -dependent $\{X_j\}_{j \in \mathbb{N}}$ (see Comte et al. [7]). Finally, it

is worth mentioning that the selection of δ in this part depends only on the parameters k_1 , d , γ and a . In our estimator, $a > \frac{3}{2}$ is given. Also, in principle, the parameters k_1 , d , γ are all known since the density f_ε is assumed to be known exactly. Hence, the estimator $\hat{f}_{X;n,\delta,a}$ is fully data-driven.

4. NUMERICAL EXPERIMENT

We now present some numerical simulations to illustrate the convergence of our estimator to the sample size n . All computations are carried by using the R software. We consider three following examples of the process $\{X_j\}_{j \in \mathbb{N}}$:

- E1: $X_j = \frac{1}{\sqrt{2}}(W_j + W_{j+1})$ for $j \in \mathbb{N}$, where $\{W_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. Gaussian random variables with zero mean and unit variance.
- E2: $X_j = \frac{1}{2\sqrt{2}}(W_j + W_{j+1})$ for $j \in \mathbb{N}$, where $\{W_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. chi-square random variables with 2 degrees of freedom.
- E3: $X_j = \sum_{k=0}^{\infty} (\sqrt{3}/2^{k+1})W_{j-k}$ for $j \in \mathbb{N}$, where $\{W_j\}_{j \in \mathbb{Z}}$ is a sequence of i.i.d. Gaussian random variables with zero mean and unit variance.

The processes $\{X_j\}_{j \in \mathbb{N}}$ in the three examples satisfy assumptions (A1) and (A5). In E1 and E3, X_j 's have the standard Gaussian distribution, so $f_X(x) = (1/\sqrt{2\pi}) \times e^{-x^2/2}$. In E2, $\text{Var}(X_j) = 1$ and $f_X(x) = 2\sqrt{2}f_{\chi^2(4)}(2\sqrt{2}x)$, where $f_{\chi^2(4)}$ is the density of the chi-square distribution with 4 degrees of freedom.

Corresponding to each example, we consider the following cases of the process $\{\varepsilon_j\}_{j \in \mathbb{N}}$:

- C1: ε_j 's have the uniform distribution on the interval $(-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$. In that case, φ_ε satisfies (A7).
- C2: ε_j 's have the triangular distribution with lower limit $-\frac{\sqrt{6}}{4}$, upper limit $\frac{\sqrt{6}}{4}$ and mode 0. In that case, φ_ε also satisfies (A7).
- C3: ε_j 's have the Laplace distribution with location parameter 0 and scale parameter $\frac{1}{2\sqrt{2}}$. In that case, φ_ε satisfies (A8).
- C4: ε_j 's have the Gaussian distribution with mean 0 and variance $\frac{1}{4}$. In that case, φ_ε satisfies (A9).

In all the cases of ε_j 's, $\text{Var}(\varepsilon_j) = \frac{1}{4}$ and hence the signal ratios, i.e., $\sqrt{\text{Var}(\varepsilon_j)/\text{Var}(X_j)}$, are equal to 0.5 corresponding to 50% noise contamination.

Because $\hat{f}_{X;n,\delta,a}$ can take some complex values, we set up $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ instead of $\hat{f}_{X;n,\delta,a}$ in our experiment. Note that the risk of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ is not bigger than the risk of $\hat{f}_{X;n,\delta,a}$ (see Remark 2.1). For that purpose, the observations $Y_1, \dots, Y_n$ of the process $\{Y_j\}_{j \in \mathbb{N}}$ must be available. Indeed, these observations are derived in the following way:

1. For E1 and E2, generate i.i.d. data $W_1, \dots, W_{n+1}$. For E3, generate i.i.d. data $W_{1-N}, W_{2-N}, \dots, W_0, W_1, \dots, W_n$, where N is chosen large enough. In our setup, we take $N := 5000$.
2. For $j = \overline{1, n}$, set $X_j = \frac{1}{\sqrt{2}}(W_j + W_{j+1})$ in E1, $X_j = \frac{1}{2\sqrt{2}}(W_j + W_{j+1})$ in E2, and $X_j \approx \sum_{k=0}^N (\sqrt{3}/2^{k+1})W_{j-k}$ in E3.
3. Generate i.i.d. data $\varepsilon_1, \dots, \varepsilon_n$ from the density f_ε .
4. For $j = \overline{1, n}$, set $Y_j = X_j + \varepsilon_j$.

Concerning δ and a , we take $a = 2$ in all cases of $\{\varepsilon_j\}_{j \in \mathbb{N}}$ (i.e., C1, C2, C3, C4), $\delta = n^{-1/4}$ in C1, C2, C3 (in view of Theorems 3.7 and 3.5), and $\delta = 4(\ln n - (\ln(\ln n))^2)^{-1}e^{-\ln n + (\ln(\ln n))^2}$ in C4 (in view of Theorem 3.8 (b)).

Let $I := [A, B]$ be a sufficiently large interval on which we plot the graphs of f_X and $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ and compute some error estimates. Concretely, we take $A = -5$, $B = 5$ in E1 and E3, and $A = -4$, $B = 6$ in E2. Let $M := 1000$ and $\{x_k\}_{k=\overline{1, M+1}}$ be the equidistant point grid in I . To approximate the MISE of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$, we replicate our computations R times independently, where R is chosen large enough. In our setup, we let $R := 100$. At the q th computation (for $q = \overline{1, R}$), a data sample $\mathbf{Y}_q := (Y_{1,q}, \dots, Y_{n,q})$ is generated and then the integrated squared error (ISE) of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$, i.e., the quantity $\text{ISE}_q := \|\Re\{\hat{f}_{X;n,\delta,a}\}(\cdot, \mathbf{Y}_q) - f_X\|^2$, is approximated by

$$\text{EISE}_q := \frac{B-A}{M} \sum_{k=1}^M |\Re\{\hat{f}_{X;n,\delta,a}\}(x_k, \mathbf{Y}_q) - f_X(x_k)|^2,$$

where

$$\begin{aligned} |\Re\{\hat{f}_{X;n,\delta,a}\}|(x_k, \mathbf{Y}_q) &:= \left| \frac{1}{2\pi} \Re \left\{ \int_{-\infty}^{\infty} \frac{\varphi_\varepsilon(-t) n^{-1} \sum_{j=1}^n e^{itY_{j,q}}}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right. \right. \\ &\quad \times \left. \left(\max \left\{ \left| \frac{\varphi_\varepsilon(-t) n^{-1} \sum_{j=1}^n e^{itY_{j,q}}}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right|; 1 \right\} \right)^{-1} e^{-itx_k} dt \right\} \Big|. \end{aligned}$$

After that we estimate the MISE of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ by

$$\text{EMISE} := \frac{1}{R} \sum_{q=1}^R \text{EISE}_q.$$

In Tables 1, 2 and 3, we report the values of EMISE for some different sample sizes n (i.e., $n = 100, 300, 500, 700, 900$) and for some different cases of ε_j 's (i.e., the cases C1, C2, C3 and C4). As expected, the values of EMISE decrease with

increasing n . This result is compatible to the consistency result of $\hat{f}_{X;n,\delta,a}$, as shown in Theorem 3.5. To visualize more specifically the convergence trend of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$, in Figures 1, 2 and 3, we show some sub-figures plotted with different sizes n and different cases of ε_j 's. In each sub-figure, we plot six curves, where the black curve is the graph of f_X , and the curves of red, blue, green, orange and purple colors are the graphs of $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ corresponding, respectively, to the 25th, 50th, 75th, 90th and 100th percentiles of the R calculated EISE $_q$ s. We realize that the curves become closer when the sample sizes n are bigger, and this indicates that $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ converges quite well to f_X . Eventually, we see that the convergence tendency of the numerical estimators in Figures 1 and 3 is more clear than the one in Figure 2. In particular, in each case of ε_j 's, the values of EMISE in Tables 1 and 3 are smaller than those in Table 2. These observations can be explained by the fact that the densities f_X in E1 and E3 are smoother than the one in E2.

n	EMISE			
	C1	C2	C3	C4
100	0.0074	0.0077	0.0074	0.0075
300	0.0035	0.0032	0.0031	0.0032
500	0.0026	0.0023	0.0019	0.0024
700	0.0018	0.0019	0.0016	0.0020
900	0.0018	0.0015	0.0013	0.0017

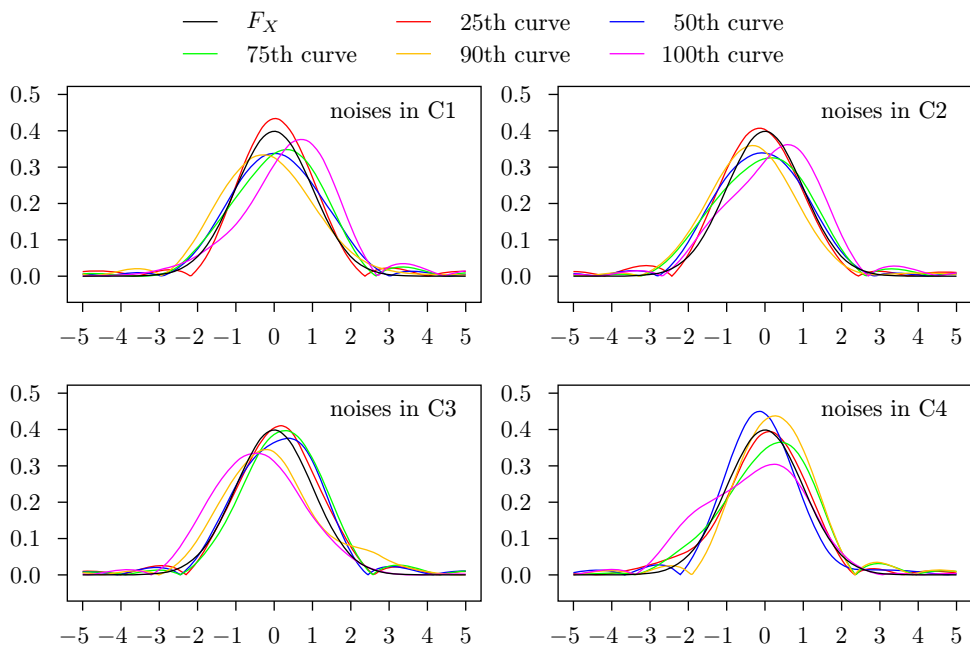
Table 1. The values of EMISE when the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E1.

n	EMISE			
	C1	C2	C3	C4
100	0.0387	0.0378	0.0332	0.0315
300	0.0323	0.0314	0.0262	0.0215
500	0.0299	0.0288	0.0239	0.0187
700	0.0283	0.0272	0.0223	0.0166
900	0.0275	0.0265	0.0213	0.0154

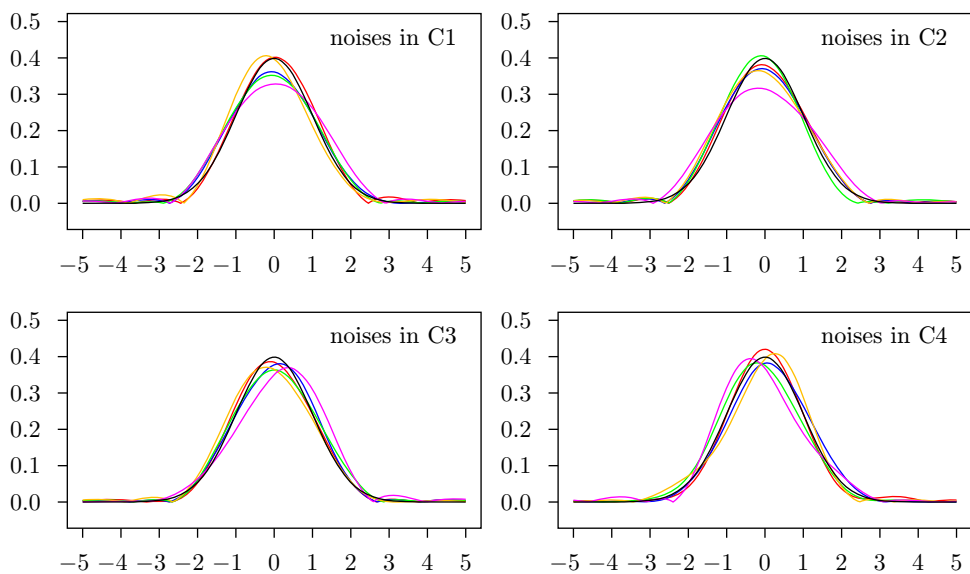
Table 2. The values of EMISE when the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E2.

n	EMISE			
	C1	C2	C3	C4
100	0.0070	0.0067	0.0077	0.0123
300	0.0053	0.0061	0.0066	0.0123
500	0.0034	0.0035	0.0040	0.0086
700	0.0031	0.0032	0.0038	0.0070
900	0.0029	0.0029	0.0035	0.0062

Table 3. The values of EMISE when the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E3.

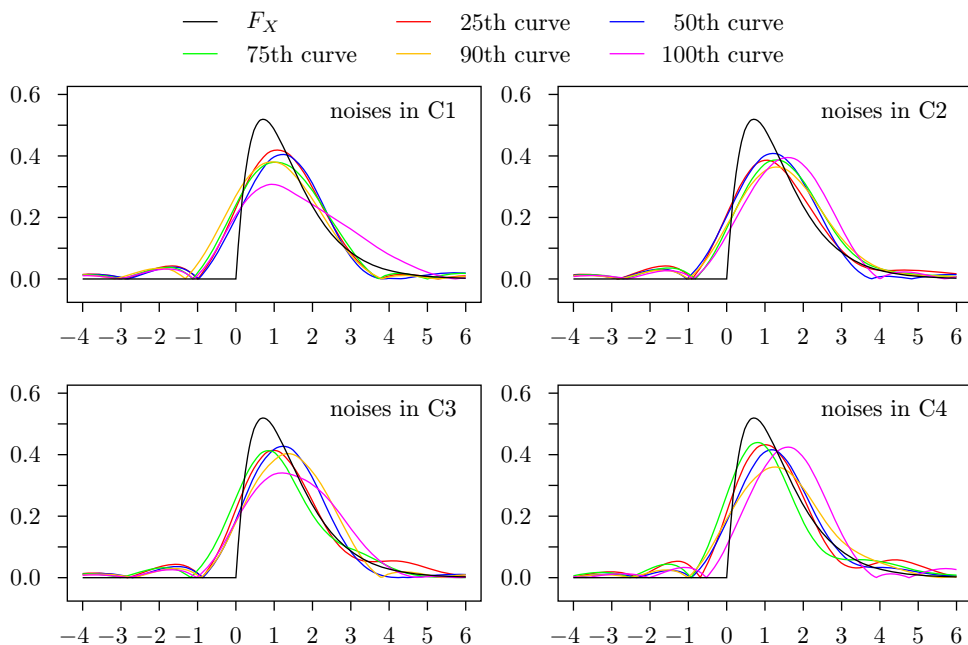


(a) $n = 100$

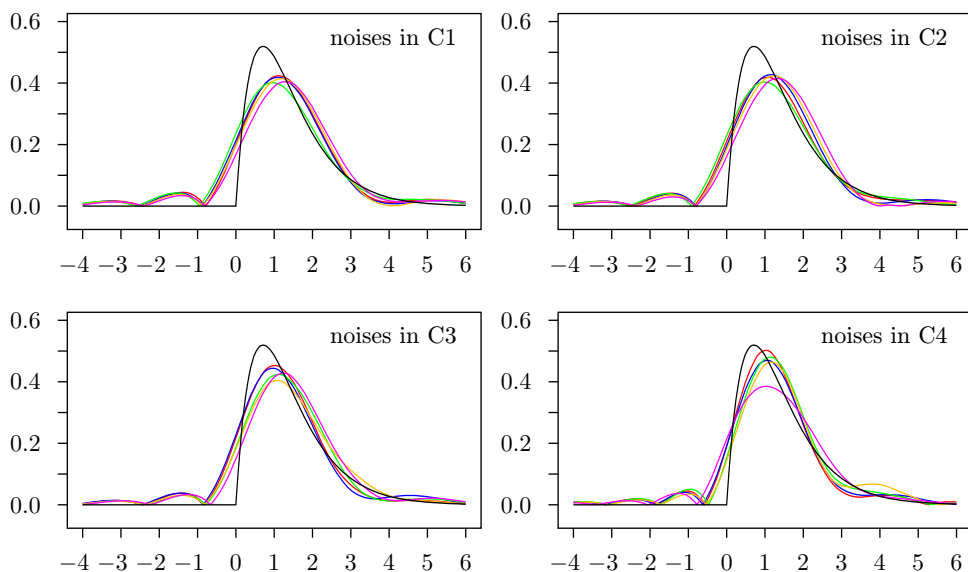


(b) $n = 500$

Figure 1. Graphs of f_X and $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ corresponding to the 25th, 50th, 75th, 90th and 100th percentiles of 100 calculated ISE_q s. Here the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E1.

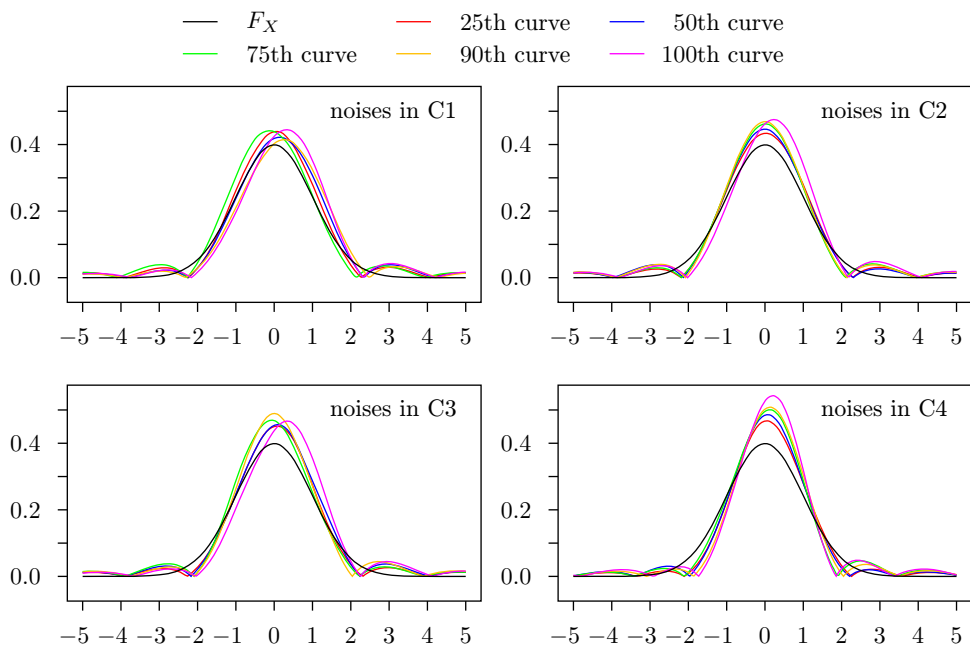


(a) $n=100$

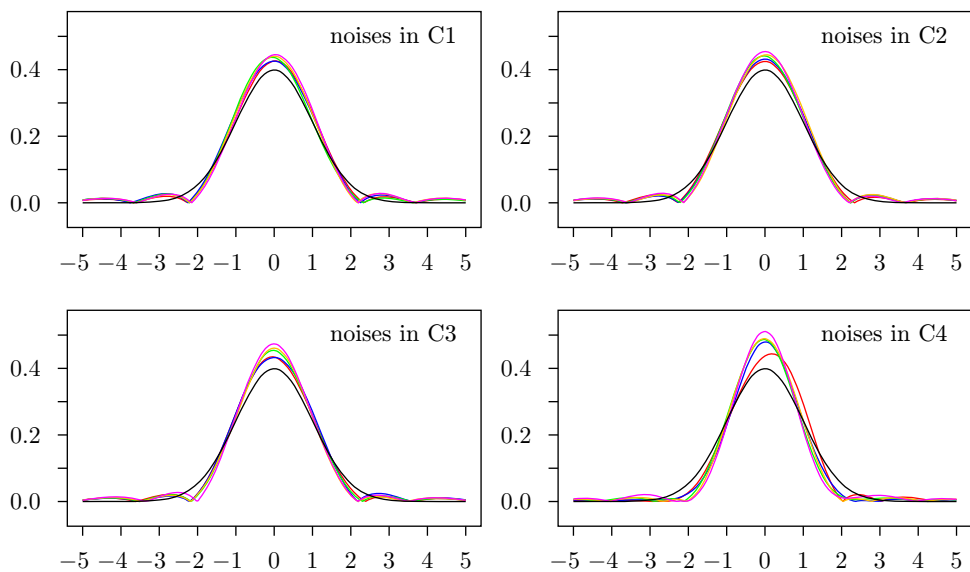


(b) $n=500$

Figure 2. Graphs of f_X and $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ corresponding to the 25th, 50th, 75th, 90th and 100th percentiles of 100 calculated ISE_q 's. Here the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E2.



(a) $n = 100$



(b) $n = 500$

Figure 3. Graphs of f_X and $|\Re\{\hat{f}_{X;n,\delta,a}\}|$ corresponding to the 25th, 50th, 75th, 90th and 100th percentiles of 100 calculated ISE_q s. Here the process $\{X_j\}_{j \in \mathbb{N}}$ is as in E3.

5. PROOFS

P r o o f of Proposition 3.1. By the Parseval identity and the Fubini theorem, we obtain that

$$\begin{aligned} & \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbb{E} \left| \frac{\varphi_\varepsilon(-t) \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \left(\max\left\{ \left| \frac{\varphi_\varepsilon(-t) \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right|; 1 \right\} \right)^{-1} - \varphi_X(t) \right|^2 dt \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbb{E} \left| \frac{\varphi_\varepsilon(-t) \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} - \varphi_X(t) \right|^2 dt. \end{aligned}$$

After that, using the standard bias-variance decomposition yields

$$\begin{aligned} (5.1) \quad \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \mathbb{E} \left(\frac{\varphi_\varepsilon(-t) \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right) - \varphi_X(t) \right|^2 dt \\ &\quad + \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{Var} \left(\frac{\varphi_\varepsilon(-t) \widehat{\varphi}_{Y;n}(t)}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{\varphi_\varepsilon(-t) \mathbb{E}(\widehat{\varphi}_{Y;n}(t))}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} - \varphi_X(t) \right|^2 dt \\ &\quad + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\varphi_\varepsilon(t)|^2 \text{Var}(\widehat{\varphi}_{Y;n}(t))}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}^2} dt. \end{aligned}$$

It is obvious that

$$(5.2) \quad \mathbb{E}(\widehat{\varphi}_{Y;n}(t)) = \varphi_Y(t) = \varphi_X(t) \varphi_\varepsilon(t).$$

For the quantity $\text{Var}(\widehat{\varphi}_{Y;n}(t))$ we have

$$\begin{aligned} \text{Var}(\widehat{\varphi}_{Y;n}(t)) &= \frac{1}{n^2} \left[\text{Var} \left(\sum_{j=1}^n e^{itY_j} \right) + \sum_{1 \leq j \neq k \leq n} \text{Cov}(e^{itY_j}, e^{itY_k}) \right] \\ &= \frac{1}{n} \text{Var}(e^{itY_1}) + \frac{1}{n^2} \sum_{1 \leq j \neq k \leq n} \text{Cov}(e^{itY_j}, e^{itY_k}). \end{aligned}$$

Note that $\text{Var}(e^{itY_1}) = 1 - |\mathbb{E}(e^{itY_1})|^2 = 1 - |\varphi_X(t)|^2 |\varphi_\varepsilon(t)|^2$ and that for all $1 \leq j \neq k \leq n$,

$$\begin{aligned} \text{Cov}(e^{itY_j}, e^{itY_k}) &= \mathbb{E}(e^{it(Y_j - Y_k)}) - \mathbb{E}(e^{itY_j}) \mathbb{E}(e^{-itY_k}) \\ &= |\varphi_\varepsilon(t)|^2 [\mathbb{E}(e^{it(X_j - X_k)}) - \mathbb{E}(e^{itX_j}) \mathbb{E}(e^{-itX_k})] \\ &= |\varphi_\varepsilon(t)|^2 \text{Cov}(e^{itX_j}, e^{itX_k}). \end{aligned}$$

Hence,

$$\text{Var}(\widehat{\varphi}_{Y;n}(t)) = \frac{1 - |\varphi_X(t)|^2 |\varphi_\varepsilon(t)|^2}{n} + \frac{|\varphi_\varepsilon(t)|^2}{n^2} \sum_{1 \leq j \neq k \leq n} \text{Cov}(e^{itX_j}, e^{itX_k}).$$

By the latter estimate and the stationarity of the sequence $\{X_j\}_{j \in \mathbb{N}}$ in (A1), we derive

$$(5.3) \quad \text{Var}(\widehat{\varphi}_{Y;n}(t)) \leq \frac{1 - |\varphi_X(t)|^2 |\varphi_\varepsilon(t)|^2}{n} + \frac{2|\varphi_\varepsilon(t)|^2}{n} \sum_{j=2}^n |\text{Cov}(e^{itX_1}, e^{itX_j})|.$$

From (5.1), (5.2) and (5.3), estimate (3.1) follows. $\square$

To prove Proposition 3.4, we need the following lemma.

Lemma 5.1 (see Birkel [2]). *Let A, B be finite sets and $\{U_j\}_{j \in A \cup B}$ be associated random variables. If $f: \mathbb{R}^{\#A} \rightarrow \mathbb{R}$ and $g: \mathbb{R}^{\#B} \rightarrow \mathbb{R}$ are partially differentiable with bounded partial derivatives, then*

$$|\text{Cov}(f((U_j)_{j \in A}), g((U_k)_{k \in B}))| \leq \sum_{j \in A} \sum_{k \in B} \left\| \frac{\partial f}{\partial t_j} \right\|_\infty \left\| \frac{\partial g}{\partial t_k} \right\|_\infty \text{Cov}(U_j, U_k),$$

where $\|\cdot\|_\infty$ denotes the usual sup-norm.

Proof of Proposition 3.4. We first have for $j = \overline{2, n}$ that

$$\begin{aligned} \text{Cov}(e^{itX_1}, e^{itX_j}) &= \text{Cov}(\cos(tX_1) + i \sin(tX_1), \cos(tX_j) + i \sin(tX_j)) \\ &= \text{Cov}(\cos(tX_1), \cos(tX_j)) - i \text{Cov}(\cos(tX_1), \sin(tX_j)) \\ &\quad + i \text{Cov}(\sin(tX_1), \cos(tX_j)) + \text{Cov}(\sin(tX_1), \sin(tX_j)), \end{aligned}$$

so

$$\begin{aligned} |\text{Cov}(e^{itX_1}, e^{itX_j})| &\leq |\text{Cov}(\cos(tX_1), \cos(tX_j))| + |\text{Cov}(\cos(tX_1), \sin(tX_j))| \\ &\quad + |\text{Cov}(\sin(tX_1), \cos(tX_j))| + |\text{Cov}(\sin(tX_1), \sin(tX_j))|. \end{aligned}$$

Under assumption (A5), applying Lemma 5.1 gives

$$|\text{Cov}(e^{itX_1}, e^{itX_j})| \leq 4t^2 \text{Cov}(X_1, X_j),$$

which yields

$$I_{3;n,\delta,a} \leq 4 \int_{-\infty}^{\infty} \frac{t^2 |\varphi_\varepsilon(t)|^4}{\max\{|\varphi_\varepsilon(t)|^2; \delta |t|^a\}^2} \sum_{j=2}^n \text{Cov}(X_1, X_j) dt \leq 4 \sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) \tilde{I}_{3;\delta,a}.$$

From the latter estimate and estimate (3.1) we obtain (3.2). $\square$

P r o o f of Theorem 3.5. From assumption (A4) we have

$$I_{1;\delta,a} = \int_{\mathbb{R} \setminus \varphi_\varepsilon^{-1}(\{0\})} \left(1 - \frac{|\varphi_\varepsilon(t)|^2}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}}\right)^2 |\varphi_X(t)|^2 dt,$$

from which we apply the Lebesgue dominated convergence theorem to get

$$(5.4) \quad \lim_{n \rightarrow \infty} I_{1;\delta,a} = 0.$$

Since $|\varphi_\varepsilon|$ is continuous on $\mathbb{R}$ and $|\varphi_\varepsilon(0)| = 1$, there exists a constant $t_0 \in (0, 1)$ depending only on φ_ε such that $|\varphi_\varepsilon(t)| \geq \frac{1}{2}$, for all $t \in [-t_0, t_0]$. Thus,

$$\begin{aligned} I_{2;\delta,a} &\leq \int_{|t| \leq t_0} \frac{1}{|\varphi_\varepsilon(t)|^2} dt + \frac{1}{\delta^2} \int_{|t| > t_0} \frac{1}{|t|^{2a}} dt \leq 8t_0 + \frac{1}{\delta^2} \int_{|t| > t_0} \frac{1}{|t|^{2a}} dt, \\ \tilde{I}_{3;\delta,a} &\leq \int_{|t| \leq t_0} \frac{t^2}{|\varphi_\varepsilon(t)|^2} dt + \frac{1}{\delta^2} \int_{|t| > t_0} \frac{1}{|t|^{2a-2}} dt \leq \frac{8t_0^3}{3} + \frac{1}{\delta^2} \int_{|t| > t_0} \frac{1}{|t|^{2a-2}} dt. \end{aligned}$$

It follows from the two latter estimates and the assumption $\lim_{n \rightarrow \infty} n\delta^2 = \infty$ that

$$(5.5) \quad \lim_{n \rightarrow \infty} \frac{I_{2;\delta,a}}{n} = 0,$$

$$(5.6) \quad \lim_{n \rightarrow \infty} \frac{\tilde{I}_{3;\delta,a}}{n} = 0.$$

From (3.2), (5.4), (5.5) and (5.6), we deduce $\lim_{n \rightarrow \infty} \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 = 0$. $\square$

To prove Theorem 3.7, we need the following lemma.

Lemma 5.2. *Let assumption (A7) hold and $q > 1$. Define $L_{\varphi_\varepsilon}(R, \varrho) := \{|t| \leq R: |\varphi_\varepsilon(t)| \leq \varrho\}$ for $R, \varrho > 0$. For $\varrho > 0$ small enough, let R_ϱ be a unique positive solution of the equation*

$$(5.7) \quad 2eMR_\varrho(q \ln R_\varrho + \ln(15e^3)) = \ln(1/\varrho).$$

Then there exist positive constants C_1, C_2 and C_3 depending only on M, q such that

$$(5.8) \quad C_1 \frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))} \leq R_\varrho \leq C_2 \frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))}$$

and

$$(5.9) \quad \lambda(L_{\varphi_\varepsilon}(R_\varrho, \varrho)) \leq C_3 \left(\frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))} \right)^{1-q}$$

for all $\varrho > 0$ small enough.

Proof. Put $\psi(R) := 2eMR(q \ln R + \ln(15e^3))$, $R > 1$. Obviously, the function ψ is increasing on $(1, \infty)$, $\psi(R) > 0$ for all $R > 1$, $\lim_{R \rightarrow \infty} \psi(R) = \infty$ and $\lim_{R \rightarrow 1^+} \psi(R) = 2eM \ln(15e^3)$. Hence, for every $0 < \varrho < (15e^3)^{-2eM}$, there exists a unique number $R_\varrho > 1$ such that $\psi(R_\varrho) = \ln(1/\varrho)$, i.e., R_ϱ satisfies (5.7). Also, we have $R_\varrho \uparrow \infty$ as $\varrho \downarrow 0$.

We next verify estimate (5.8). For $\varrho > 0$ small enough, we derive from (5.7) that

$$(5.10) \quad \frac{\ln(1/\varrho)}{2eM(q+1)} \leq R_\varrho \ln R_\varrho \leq \frac{\ln(1/\varrho)}{2eMq},$$

so

$$(5.11) \quad R_\varrho \leq R_\varrho \ln R_\varrho \leq \frac{1}{2eMq} \ln(1/\varrho).$$

From (5.10) and (5.11), we have

$$(5.12) \quad R_\varrho \geq \frac{\ln(1/\varrho)}{2eM(q+1) \ln R_\varrho} \geq \ln(1/\varrho) \left(2eM(q+1) \ln \left(\frac{1}{2eMq} \ln(1/\varrho) \right) \right)^{-1} \\ \geq \frac{\ln(1/\varrho)}{4eM(q+1) \ln(\ln(1/\varrho))}.$$

From (5.10) and (5.12), we have

$$R_\varrho \leq \frac{\ln(1/\varrho)}{2eMq \ln R_\varrho} \leq \ln(1/\varrho) \left(2eMq \ln \left(\frac{\ln(1/\varrho)}{4eM(q+1) \ln(\ln(1/\varrho))} \right) \right)^{-1} \\ = \ln(1/\varrho) \left(2eMq \left[\ln \left(\frac{1}{4eM(q+1)} \right) + \ln(\ln(1/\varrho)) - \ln(\ln(\ln(1/\varrho))) \right] \right)^{-1} \\ \leq \frac{\ln(1/\varrho)}{eMq \ln(\ln(1/\varrho))}.$$

Hence, we have shown (5.8) with $C_1 := 1/[4eM(q+1)]$, $C_2 := 1/(eMq)$.

Now we prove (5.9). Define $\Phi(z) := \int_{-M}^M f_\varepsilon(x) e^{izx} dx$, $z \in \mathbb{C}$. Then Φ is a nontrivial entire function, $|\Phi(0)| = 1$ and $\ln \left(\max_{|z|=2eR_\varrho} |\Phi(z)| \right) \leq 2eMR_\varrho$. Hence, by applying Theorem 4, Section 11.3 in Levin [18] with $\eta := R_\varrho^{-q}$, there exists a set of disks $\{D(z_j, r_j)\}_{j \in J} \subset \mathbb{C}$ with $\sum_{j \in J} r_j \leq \eta R_\varrho$ such that

$$|\Phi(z)| > \exp \left\{ -\ln \left(\frac{15e^3}{\eta} \right) \ln \left(\max_{|z|=2eR_\varrho} |\Phi(z)| \right) \right\} \\ \geq \exp \{ -2eMR_\varrho(q \ln R_\varrho + \ln(15e^3)) \} = \varrho$$

for all $z \in D'(0, R_\varrho) \setminus \bigcup_{j \in J} D(z_j, r_j)$. Here $D'(0, R_\varrho) := \{z \in \mathbb{C} : |z| \leq R_\varrho\}$ and $D(z_j, r_j) := \{z \in \mathbb{C} : |z - z_j| < r_j\}$. Consequently,

$$L_{\varphi_\varepsilon}(R_\varrho, \varrho) = \{z \in \mathbb{R} : |z| \leq R_\varrho, |\Phi(z)| \leq \varrho\} \subset \bigcup_{j \in J} [D(z_j, r_j) \cap \mathbb{R}].$$

Hence,

$$\lambda(L_{\varphi_\varepsilon}(R_\varrho, \varrho)) \leq 2 \sum_{j \in J} r_j \leq 2\eta R_\varrho \leq 2 \left(C_1 \frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))} \right)^{1-q} = C_3 \left(\frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))} \right)^{1-q}$$

with $C_3 := 2C_1^{1-q}$. □

Proof of Theorem 3.7. Put $A_{\delta,a} := \{t \in \mathbb{R} : |\varphi_\varepsilon(t)|^2 < \delta|t|^a\}$. Then

$$(5.13) \quad I_{1;\delta,a} = \int_{A_{\delta,a}} \left(1 - \frac{|\varphi_\varepsilon(t)|^2}{\max\{|\varphi_\varepsilon(t)|^2; \delta|t|^a\}} \right)^2 |\varphi_X(t)|^2 dt \leq \int_{A_{\delta,a}} |\varphi_X(t)|^2 dt.$$

Let $R, \varrho > 0$ and $L_{\varphi_\varepsilon}(R, \varrho)$ be defined as in Lemma 5.2. For $\delta \in (0, R^{-a}\varrho^2]$ we have

$$\begin{aligned} A_{\delta,a} &= L_{\varphi_\varepsilon}(R, \varrho) \cup \{|t| \leq R : \varrho < |\varphi_\varepsilon(t)| < \sqrt{\delta|t|^a}\} \cup \{|t| > R : |\varphi_\varepsilon(t)| < \sqrt{\delta|t|^a}\} \\ &\subset L_{\varphi_\varepsilon}(R, \varrho) \cup \{|t| > R\}, \end{aligned}$$

which together with (5.13) yields

$$I_{1;\delta,a} \leq \int_{L_{\varphi_\varepsilon}(R, \varrho)} |\varphi_X(t)|^2 dt + 2 \int_R^\infty |\varphi_X(t)|^2 dt.$$

By the fact that $|\varphi_X(t)| \leq 1$ for any $t \in \mathbb{R}$ and by assumption (A6), we get

$$I_{1;\delta,a} \leq \lambda(L_{\varphi_\varepsilon}(R, \varrho)) + 2(1 + R^2)^{-\alpha} \int_R^\infty |\varphi_X(t)|^2 (1 + t^2)^\alpha dt \lesssim \lambda(L_{\varphi_\varepsilon}(R, \varrho)) + R^{-2\alpha}.$$

Also, according to the proof of Theorem 3.5, we have $I_{2;\delta,a}, \tilde{I}_{3;\delta,a} \lesssim 1/\delta^2$. Therefore, from the estimates of $I_{1;\delta,a}, I_{2;\delta,a}, \tilde{I}_{3;\delta,a}$, assumption (A5) and estimate (3.2) in Proposition 3.4, we derive, for $\delta \in (0, R^{-a}\varrho^2]$,

$$(5.14) \quad \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \lambda(L_{\varphi_\varepsilon}(R, \varrho)) + R^{-2\alpha} + \frac{1}{n\delta^2}.$$

For $\varrho > 0$ small enough, let R_ϱ be the unique solution of equation (5.7), where we take $q := 2\alpha + 1$. Applying (5.14) with $R := R_\varrho$ and Lemma 5.2, we have, for $\delta \in (0, R_\varrho^{-a}\varrho^2]$,

$$\mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \left(\frac{\ln(1/\varrho)}{\ln(\ln(1/\varrho))} \right)^{-2\alpha} + \frac{1}{n\delta^2}.$$

Letting $\delta = n^{-l}$, $\varrho = n^{-l_1}$ with $0 < l < \frac{1}{2}$, $0 < l_1 < l/2$, we obtain

$$\mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \left(\frac{\ln n}{\ln(\ln n)} \right)^{-2\alpha} + n^{2l-1} \lesssim \left(\frac{\ln n}{\ln(\ln n)} \right)^{-2\alpha},$$

and this completes the proof. □

Proof of Theorem 3.8. As in the proof of Theorem 3.7, we have

$$(5.15) \quad I_{1;\delta,a} \leq \int_{A_{\delta,a}} |\varphi_X(t)|^2 dt,$$

where $A_{\delta,a} := \{t \in \mathbb{R} : |\varphi_\varepsilon(t)|^2 < \delta|t|^a\}$.

(a) Let $0 < \delta < c_1^2/4^\beta$. For an arbitrary $t \in A_{\delta,a}$ we have $c_1^2(1+|t|)^{-2\beta} \leq |\varphi_\varepsilon(t)|^2 < \delta|t|^a$ by the left inequality in (A8), so $c_1^2/\delta < |t|^a(1+|t|)^{2\beta}$. This and the constraint $0 < \delta < c_1^2/4^\beta$ imply $|t| > 1$. Hence, $c_1^2/\delta < |t|^a(2|t|)^{2\beta} = 4^\beta|t|^{a+2\beta}$, which yields $|t| > T_*$ with $T_* := (c_1^2/(4^\beta\delta))^{1/(a+2\beta)}$. Thus, we have shown that $A_{\delta,a} \subset \{t \in \mathbb{R} : |t| > T_*\}$ for $0 < \delta < c_1^2/4^\beta$. Combining this with (5.15) and (A6), we obtain

$$I_{1;\delta,a} \leq \int_{T_*}^{\infty} |\varphi_X(t)|^2 dt \leq (1+T_*^2)^{-\alpha} \int_{T_*}^{\infty} |\varphi_X(t)|^2 (1+t^2)^\alpha dt \lesssim T_*^{-2\alpha}.$$

Next, by (A8), we have

$$\begin{aligned} I_{2;\delta,a} &\leq \int_{|t| \leq T_*} \frac{1}{|\varphi_\varepsilon(t)|^2} dt + \int_{|t| > T_*} \frac{|\varphi_\varepsilon(t)|^2}{\delta^2|t|^{2a}} dt \\ &\lesssim \int_{|t| \leq T_*} (1+|t|)^{2\beta} dt + \int_{|t| > T_*} \frac{(1+|t|)^{-2\beta}}{\delta^2|t|^{2a}} dt \lesssim T_*^{2\beta+1} + \frac{T_*^{1-2a-2\beta}}{\delta^2}, \\ \tilde{I}_{3;\delta,a} &\leq \int_{|t| \leq T_*} t^2 dt + \int_{|t| > T_*} \frac{t^2|\varphi_\varepsilon(t)|^4}{\delta^2|t|^{2a}} dt \\ &\lesssim T_*^3 + \int_{|t| > T_*} \frac{t^2(1+|t|)^{-4\beta}}{\delta^2|t|^{2a}} dt \lesssim T_*^3 + \frac{T_*^{3-2a-4\beta}}{\delta^2}. \end{aligned}$$

It follows from estimate (3.2) and the estimates of $I_{1;\delta,a}$, $I_{2;\delta,a}$ and $\tilde{I}_{3;\delta,a}$ that

$$(5.16) \quad \mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim T_*^{-2\alpha} + \frac{T_*^{2\beta+1}}{n} + \frac{T_*^{1-2a-2\beta}}{n\delta^2} + \frac{T_*^3}{n} + \frac{T_*^{3-2a-4\beta}}{n\delta^2}.$$

Now we distinguish two cases of β as follows:

Case 1: $\beta \geq 1$. In this case, it follows from (5.16) that

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim T_*^{-2\alpha} + \frac{T_*^{2\beta+1}}{n} + \frac{T_*^{1-2a-2\beta}}{n\delta^2}.$$

Inserting $T_* := (c_1^2/(4^\beta\delta))^{1/(a+2\beta)}$ into the right-hand side of the latter estimate yields

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \delta^{2\alpha/(a+2\beta)} + \frac{1}{n\delta^{(2\beta+1)/(a+2\beta)}}.$$

Letting $\delta = n^{-(a+2\beta)/(2\alpha+2\beta+1)}$ gives

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim n^{-2\alpha/(2\alpha+2\beta+1)}.$$

Case 2: $0 < \beta < 1$. In this case, it follows from (5.16) that

$$\mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim T_*^{-2\alpha} + \frac{T_*^3}{n} + \frac{T_*^{3-2a-4\beta}}{n\delta^2}.$$

Inserting $T_* := (c_1^2/(4^\beta\delta))^{1/(a+2\beta)}$ into the right-hand side of the latter estimate yields

$$\mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim \delta^{2\alpha/(a+2\beta)} + \frac{1}{n\delta^{3/(a+2\beta)}}.$$

Letting $\delta = n^{-(a+2\beta)/(2\alpha+3)}$ gives

$$\mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim n^{-2\alpha/(2\alpha+3)}.$$

(b) Put

$$T_{**} := \left(\frac{\ln n - (\ln(\ln n))^2}{2d} \right)^{1/\gamma}, \quad \delta := k_1^2 T_{**}^{-a} e^{-2dT_{**}^\gamma}.$$

For an arbitrary $t \in A_{\delta,a}$ we have $k_1^2 e^{-2d|t|^\gamma} \leq |\varphi_\varepsilon(t)|^2 < \delta |t|^a$ by the left inequality in (A9), so $k_1^2/\delta < |t|^a e^{2d|t|^\gamma}$. Since $k_1^2/\delta = T_{**}^a e^{2dT_{**}^\gamma}$, we infer $|t| > T_{**}$. So we have shown that $A_{\delta,a} \subset \{t \in \mathbb{R} : |t| > T_{**}\}$. Combining this with (5.15) and (A6), we obtain

$$I_{1;\delta,a} \leq \int_{T_{**}}^\infty |\varphi_X(t)|^2 dt \leq (1 + T_{**}^2)^{-\alpha} \int_{T_{**}}^\infty |\varphi_X(t)|^2 (1 + t^2)^\alpha dt \lesssim T_{**}^{-2\alpha}.$$

Next, by (A9), we have

$$\begin{aligned} I_{2;\delta,a} &\leq \int_{|t| \leq T_{**}} \frac{1}{|\varphi_\varepsilon(t)|^2} dt + \int_{|t| > T_{**}} \frac{|\varphi_\varepsilon(t)|^2}{\delta^2 |t|^{2a}} dt \\ &\lesssim \int_{|t| \leq T_{**}} e^{2d|t|^\gamma} dt + \int_{|t| > T_{**}} \frac{e^{-2d|t|^\gamma}}{\delta^2 |t|^{2a}} dt \lesssim T_{**}^{1-\gamma} e^{2dT_{**}^\gamma} + \frac{T_{**}^{-2a+1-\gamma} e^{-2dT_{**}^\gamma}}{\delta^2}, \\ \tilde{I}_{3;\delta,a} &\leq \int_{|t| \leq T_{**}} t^2 dt + \int_{|t| > T_{**}} \frac{t^2 |\varphi_\varepsilon(t)|^4}{\delta^2 |t|^{2a}} dt \\ &\lesssim T_{**}^3 + \int_{|t| > T_{**}} \frac{t^2 e^{-4d|t|^\gamma}}{\delta^2 |t|^{2a}} dt \lesssim T_{**}^3 + \frac{T_{**}^{3-2a-\gamma} e^{-4dT_{**}^\gamma}}{\delta^2}. \end{aligned}$$

It follows from estimate (3.2) and the estimates of $I_{1;\delta,a}$, $I_{2;\delta,a}$ and $\tilde{I}_{3;\delta,a}$ that

$$\begin{aligned} (5.17) \quad \mathbb{E} \|\hat{f}_{X;n,\delta,a} - f_X\|^2 &\lesssim T_{**}^{-2\alpha} + \frac{T_{**}^{1-\gamma} e^{2dT_{**}^\gamma}}{n} + \frac{T_{**}^{-2a+1-\gamma} e^{-2dT_{**}^\gamma}}{n\delta^2} + \frac{T_{**}^3}{n} + \frac{T_{**}^{3-2a-\gamma} e^{-4dT_{**}^\gamma}}{n\delta^2} \\ &\lesssim T_{**}^{-2\alpha} + \frac{T_{**}^{1-\gamma} e^{2dT_{**}^\gamma}}{n} + \frac{T_{**}^{-2a+1-\gamma} e^{-2dT_{**}^\gamma}}{n\delta^2} \\ &= T_{**}^{-2\alpha} + \frac{T_{**}^{1-\gamma} e^{2dT_{**}^\gamma}}{n} \left(1 + \frac{T_{**}^{-2a} e^{-4dT_{**}^\gamma}}{\delta^2} \right) \lesssim T_{**}^{-2\alpha} + \frac{T_{**}^{1-\gamma} e^{2dT_{**}^\gamma}}{n}. \end{aligned}$$

From the definition of T_{**} we derive for large n that

$$(5.18) \quad \left(\frac{\ln n}{4d}\right)^{1/\gamma} \leq T_{**} \leq \left(\frac{\ln n}{2d}\right)^{1/\gamma}.$$

For n large enough, $\ln(\ln n) > (2\alpha + 1 - \gamma)/\gamma$ and so

$$T_{**} \leq \left(\frac{\ln n - (2\alpha + 1 - \gamma)\gamma^{-1} \ln(\ln n)}{2d}\right)^{1/\gamma},$$

which gives

$$(5.19) \quad e^{2dT_{**}^\gamma} \lesssim n(\ln n)^{-(2\alpha+1-\gamma)/\gamma}.$$

From (5.17), (5.18) and (5.19), we derive

$$\mathbb{E}\|\hat{f}_{X;n,\delta,a} - f_X\|^2 \lesssim (\ln n)^{-2\alpha/\gamma},$$

and this completes the proof. $\square$

Acknowledgments. We would like to thank the anonymous reviewer and the associate editor for fruitful comments and suggestions which have helped to significantly improve the paper.

References

- [1] *I. Bagai, B. L. S. Prakasa Rao*: Kernel-type density and failure rate estimation for associated sequences. *Ann. Inst. Stat. Math.* **47** (1995), 253–266. [zbl](#) [MR](#) [doi](#)
- [2] *T. Birkel*: On the convergence rate in the central limit theorem for associated processes. *Ann. Probab.* **16** (1988), 1685–1698. [zbl](#) [MR](#) [doi](#)
- [3] *C. Butucea, A. B. Tsybakov*: Sharp optimality in density deconvolution with dominating bias. I. *Theory Probab. Appl.* **52** (2008), 24–39. [zbl](#) [MR](#) [doi](#)
- [4] *M. Carrasco, J.-P. Florens*: A spectral method for deconvolving a density. *Econom. Theory* **27** (2011), 546–581. [zbl](#) [MR](#) [doi](#)
- [5] *R. J. Carroll, P. Hall*: Optimal rates of convergence for deconvolving a density. *J. Am. Stat. Assoc.* **83** (1988), 1184–1186. [zbl](#) [MR](#) [doi](#)
- [6] *C. Chesneau*: On the adaptive wavelet deconvolution of a density for strong mixing sequences. *J. Korean Stat. Soc.* **41** (2012), 423–436. [zbl](#) [MR](#) [doi](#)
- [7] *F. Comte, J. Dedecker, M.-L. Taupin*: Adaptive density deconvolution with dependent inputs. *Math. Methods Stat.* **17** (2008), 87–112. [zbl](#) [MR](#) [doi](#)
- [8] *F. Comte, Y. Rozenholc, M.-L. Taupin*: Penalized contrast estimator for adaptive density deconvolution. *Can. J. Stat.* **34** (2006), 431–452. [zbl](#) [MR](#) [doi](#)
- [9] *J. Dedecker, C. Prieur*: New dependence coefficients. Examples and applications to statistics. *Probab. Theory Relat. Fields* **132** (2005), 203–236. [zbl](#) [MR](#) [doi](#)
- [10] *A. Delaigle, A. Meister*: Nonparametric function estimation under Fourier-oscillating noise. *Stat. Sin.* **21** (2011), 1065–1092. [zbl](#) [MR](#) [doi](#)

- [11] *L. Devroye*: Consistent deconvolution in density estimation. *Can. J. Stat.* *17* (1989), 235–239. [zbl](#) [MR](#) [doi](#)
- [12] *J. D. Esary, F. Proschan, D. W. Walkup*: Association of random variables, with applications. *Ann. Math. Stat.* *38* (1967), 1466–1474. [zbl](#) [MR](#) [doi](#)
- [13] *J. Fan*: On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Stat.* *19* (1991), 1257–1272. [zbl](#) [MR](#) [doi](#)
- [14] *C. M. Fortuin, P. W. Kasteleyn, J. Ginibre*: Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* *22* (1971), 89–103. [zbl](#) [MR](#) [doi](#)
- [15] *P. Groeneboom, G. Jongbloed*: Density estimation in the uniform deconvolution model. *Stat. Neerl.* *57* (2003), 136–157. [zbl](#) [MR](#) [doi](#)
- [16] *P. Hall, A. Meister*: A ridge-parameter approach to deconvolution. *Ann. Stat.* *35* (2007), 1535–1558. [zbl](#) [MR](#) [doi](#)
- [17] *E. L. Lehmann*: Some concepts of dependence. *Ann. Math. Stat.* *37* (1966), 1137–1153. [zbl](#) [MR](#) [doi](#)
- [18] *B. Y. Levin*: Lectures on Entire Functions. Translations of Mathematical Monographs 150. AMS, Providence, 1996. [zbl](#) [MR](#) [doi](#)
- [19] *M. C. Liu, R. L. Taylor*: A consistent nonparametric density estimator for the deconvolution problem. *Can. J. Stat.* *17* (1989), 427–438. [zbl](#) [MR](#) [doi](#)
- [20] *E. Masry*: Multivariate probability density deconvolution for stationary random processes. *IEEE Trans. Inf. Theory* *37* (1991), 1105–1115. [zbl](#) [MR](#) [doi](#)
- [21] *E. Masry*: Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes. *J. Multivariate Anal.* *44* (1993), 47–68. [zbl](#) [MR](#) [doi](#)
- [22] *E. Masry*: Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic Processes Appl.* *47* (1993), 53–74. [zbl](#) [MR](#) [doi](#)
- [23] *E. Masry*: Deconvolving multivariate kernel density estimates from contaminated associated observations. *IEEE Trans. Inf. Theory* *49* (2003), 2941–2952. [zbl](#) [MR](#) [doi](#)
- [24] *A. Meister*: Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions. *Inverse Probl.* *24* (2008), Article ID 015003, 14 pages. [zbl](#) [MR](#) [doi](#)
- [25] *A. Meister*: Deconvolution Problems in Nonparametric Statistics. Lecture Notes in Statistics 193. Springer, Berlin, 2009. [zbl](#) [MR](#) [doi](#)
- [26] *P. E. Oliveira*: Asymptotics for Associated Random Variables. Springer, Berlin, 2012. [zbl](#) [MR](#) [doi](#)
- [27] *J. Pan*: Tail dependence of random variables from ARCH and heavy-tailed bilinear models. *Sci. China, Ser. A* *45* (2002), 749–760. [zbl](#) [MR](#) [doi](#)
- [28] *M. Pensky, B. Vidakovic*: Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Stat.* *27* (1999), 2033–2053. [zbl](#) [MR](#) [doi](#)
- [29] *W. Rudin*: Real and Complex Analysis. McGraw-Hill, New York, 1987. [zbl](#) [MR](#)
- [30] *L. A. Stefanski, R. J. Carroll*: Deconvoluting kernel density estimators. *Statistics* *21* (1990), 169–184. [zbl](#) [MR](#) [doi](#)
- [31] *B. van Es, P. Spreij, H. van Zanten*: Nonparametric volatility density estimation. *Bernoulli* *9* (2003), 451–465. [zbl](#) [MR](#) [doi](#)
- [32] *H. van Zanten, P. Zareba*: A note on wavelet density deconvolution for weakly dependent data. *Stat. Inference Stoch. Process.* *11* (2008), 207–219. [zbl](#) [MR](#) [doi](#)
- [33] *V. A. Volkonskij, Y. A. Rozanov*: Some limit theorems for random functions. I. *Theor. Probab. Appl.* *4* (1959), 178–197. [zbl](#) [MR](#) [doi](#)

Authors' addresses: *Le Thi Hong Thuy*, Faculty of Fundamental Sciences, Van Lang University, 69/68 Dang Thuy Tram Street, Ward 13, Binh Thanh District, Ho Chi Minh City, Vietnam, e-mail: thuy.lth@vlu.edu.vn; *Cao Xuan Phuong* (corresponding author), Faculty of Mathematics and Statistics, Ton Duc Thang University, 19 Nguyen Huu Tho Street, Tan Phong Ward, District 7, Ho Chi Minh City, Vietnam, e-mail: caoxuanphuong@tdtu.edu.vn.