Metrika (2022) 85:289–322
https://doi.org/10.1007/s00184-021-00830-5

Nonparametric estimation of cumulative distribution function from noisy data in the presence of Berkson and classical errors

Cao Xuan Phuong[1] · Le Thi Hong Thuy[2] · Vo Nguyen Tuyet Doan[3]

Received: 20 October 2020 / Accepted: 2 July 2021 / Published online: 9 July 2021

Abstract
Let X, Y, W, δ and ε be continuous univariate random variables defined on a probability space such that $Y = X + \varepsilon$ and $W = X + \delta$. Herein X, δ and ε are assumed to be mutually independent. The variables ε and δ are called classical and Berkson errors, respectively. Their distributions are known exactly. Suppose we only observe a random sample $Y_1, \ldots, Y_n$ from the distribution of Y. This paper is devoted to a nonparametric estimation of the unknown cumulative distribution function F_W of W based on the observations as well as on the distributions of ε, δ. An estimator for F_W depending on a smoothing parameter is suggested. It is shown to be consistent with respect to the mean squared error. Under certain regularity assumptions on the densities of X, δ and ε, we establish some upper and lower bounds on the convergence rate of the proposed estimator. Finally, we perform some numerical examples to illustrate our theoretical results.

Keywords Cumulative distribution function · Deconvolution · Berkson errors · Classical errors

Mathematics Subject Classification 62G05 · 62G20.

✉ Cao Xuan Phuong
caoxuanphuong@tdtu.edu.vn

Le Thi Hong Thuy
thuy.lth@vlu.edu.vn

Vo Nguyen Tuyet Doan
tuyetdoan201096@gmail.com

1 Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam

2 Faculty of Basic Sciences, Van Lang University, Ho Chi Minh City, Vietnam

3 Department of Electronic Engineering, Myongji University, Yongin 17058, Korea

1 Introduction

Let W be a continuous random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and distributed with an unknown distribution. Denote by F_W the unknown cumulative distribution distribution (cdf for short) of W, i.e.

$$F_W(x) := \mathbb{P}(W \leq x), \quad x \in \mathbb{R}.$$

Our interest in this paper is the nonparametric estimation problem of $F_W(x)$. In practice, there exist some situations in which, instead of observing W, one only observes a surrogate variable X, related to W via the model

$$W = X + \delta. \tag{1.1}$$

Here δ is a random error, called a Berkson error. It is assumed to be independent of X; in addition, its distribution is known exactly. The model (1.1) is plausible in many situations. For example, in exposure studies, an important problem is to access the effect of an air pollutant to the residents of a certain city. In that situation, W can represent as the personal exposure of the residents to the air pollutant. However, suppose that only X, the concentration of the air pollutant at certain observation stations in the city, can be measured. Mathematically, W varies randomly around X and can be expressed as X plus to a random error δ, which reflects the individual variation in the exposure from the measured exposure. More motivations and explanations of the model (1.1) can be found in Fuller (1987) and Carroll et al. (2006).

Measurements often contain errors which may be generated from many sources as measurement devices, or intrinsic difficulty in measuring variables of interest. Hence, it is more realistic to introduce an observable contaminated version Y of X linked to X through the model

$$Y = X + \varepsilon, \tag{1.2}$$

where the variable ε, assumed to be independent of both X and δ, reflects an error when measuring X with an inaccurate measurement device or procedure. We call ε a classical error. Similar to the Berkson error δ, we also assume that the distribution of δ is fully known. Now let $Y_1, \ldots, Y_n$ be a sample of i.i.d. observations taken from the distribution of Y. Based on the observations as well as the knowledge about the distributions of δ and ε, we aim to estimate $F_W(x)$ in a nonparametric strategy.

Berkson errors are in nature different from classical errors. In the model (1.1), W plays the role of an unobserved true variable and X is only a proxy for W. The Berkson error δ is independent of X, but not independent of W. Nevertheless, in the model (1.2), the classical error ε is independent of X, considered as an unobserved true variable. The model (1.1) was first considered in Berkson (1950). It has attracted less interest and mainly treated in regression problems with errors-in-variables, see, e.g. Berkson (1950), Delaigle et al. (2006), Carroll et al. (2009), Delaigle and Meister (2011), Delaigle (2014). The model (1.2) leads to so-called deconvolution problems for the density or the cdf of X that have been widely and intensively studied in the statistical literature. See, for instance, Carroll and Hall (1988), Stefanski and Carroll (1990), Fan (1991), Pensky and Vidakovic (1999), Comte et al. (2006), Hall and Meister (2007) for density estimation context, and Gaffey (1959), Fan (1991), Cordy and Thomas (1997), Meister (2009), Dattner et al. (2011), Dattner and Reiser (2013), Trong and Phuong (2019) for cdf estimation context. To the best of our knowledge, however, it seems that there is very little research referring to the two models simultaneously. In density estimation context, we refer to Delaigle (2007) and Rimal and Pensky (2019) as a handful of related works. Delaigle (2007) introduced two nonparametric methods for estimating the density f_W of W and studied some asymptotic properties of proposed estimators. Rimal and Pensky (2019) considered a nonparametric density estimation of f_W in the case of small Berkson error δ. In cdf estimation context, we have not yet found any related work so far. Therefore, our aim is to fill partially this gap in the current work.

More concretely, we study the problem of estimating the cdf F_W in some various cases of the errors δ and ε, where the characteristic function φ_ε of ε is assumed to be non-vanishing on the whole real line. Such cases of ε are referred to as “standard” cases of its distribution. In order to build a nonparametric estimator of F_W, we apply a direct inversion formula for cdf’s as well as nonparametric deconvolution tools. Our proposed estimator depends only on a regularization parameter and can be viewed as a generalized version of the estimator for the cdf of X introduced in Dattner et al. (2011), who only considered the classical error model (1.2). We discuss to the consistency of the proposed estimator with respect to the mean squared error under certain conditions on the regularization parameter. After that we establish upper and minimax lower bounds on the convergence rate of our estimator uniformly over a Sobolev class of the unknown density of X.

The rest of this paper is structured as follows. In Sect. 2, we first introduce some notations and then discuss our estimation method of the cdf F_W. In Section 3, we study mean consistency of our estimator. In Sect. 4, we establish some upper and minimax lower bounds on the convergence rate of the proposed estimator. In Sect. 5, we perform some numerical examples to illustrate our theoretical results. Finally, we present all proofs in Sect. 6.

2 Notations and the estimator

2.1 Notations

We introduce some notations which will be used later. The characteristic function of a random variable U is defined by $\varphi_U(t) := \mathbb{E}(\mathrm{e}^{\mathrm{i}tU})$ with $t \in \mathbb{R}$ and $\mathrm{i}^2 = -1$. For a complex number z, the notations $\Re\{z\}$, $\Im\{z\}$ and $\mathrm{Arg}\{z\}$ stand respectively for the real part, the imaginary part and the principal argument of z. The notation $\|\phi\|_1$ denotes the usual $L^1(\mathbb{R})$-norm of a function $\phi \in L^1(\mathbb{R})$, i.e. $\|\phi\|_1 := \int_{-\infty}^{\infty} |\phi(x)|\mathrm{d}x$. The convolution of two measurable functions u, v on $\mathbb{R}$ is the function $(u * v)(x) := \int_{-\infty}^{\infty} u(x-s)v(s)\mathrm{d}s$. For positive deterministic parameters a_n and b_n depending on the sample size n, the notation $a_n \lesssim b_n$ means that there exists a constant $c > 0$

independent of n and a positive integer n_0 such that $a_n \le c\, b_n$, for all $n \ge n_0$. Also, we write $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$.

2.2 The estimator

Throughout this paper, the characteristic function φ_ε of ε is assumed to satisfy

$$\varphi_\varepsilon(t) \neq 0, \quad \text{for all} t \in \mathbb{R}. \tag{2.1}$$

The assumption (2.1) is standard in deconvolution topics, see, e.g., Fan (1991).

Now we describe our method for estimating F_W. First, since W is a continuous type random variable, the cdf F_W is continuous on $\mathbb{R}$ and so (see, e.g. Gil-Pelaez (1951))

$$F_W(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}\Im\{\varphi_W(t)\mathrm{e}^{-\mathrm{i}tx}\}\mathrm{d}t. \tag{2.2}$$

The independence of X, δ and ε gives that $\varphi_Y(t) = \varphi_X(t)\varphi_\varepsilon(t)$ and $\varphi_W(t) = \varphi_X(t)\varphi_\delta(t)$, for all $t \in \mathbb{R}$. Combining the two last equalities with (2.2) and applying the assumption (2.1), we derive that

$$F_W(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}\Im\left\{\frac{\varphi_Y(t)\varphi_\delta(t)}{\varphi_\varepsilon(t)}\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t. \tag{2.3}$$

In the equation (2.3), only the quantity $\varphi_Y(t)$ is unknown and thus have to be estimated. Indeed, with the available data $Y_1, \ldots, Y_n$, we consider the quantity

$$\widehat{\varphi}_Y(t) := \frac{1}{n}\sum_{j=1}^{n}\mathrm{e}^{\mathrm{i}tY_j}, \quad t \in \mathbb{R},$$

known as the empirical characteristic function of Y. We see that $\mathbb{E}\widehat{\varphi}_Y(t) = \varphi_Y(t)$, $\mathrm{Var}\widehat{\varphi}_Y(t) \to 0$ as $n \to \infty$; moreover, $\widehat{\varphi}_Y(t) \to \varphi_Y(t)$ almost surely as $n \to \infty$, due to the strong law of large numbers. Thus, we propose to estimate $\varphi_Y(t)$ by $\widehat{\varphi}_Y(t)$. In doing so, we then can define a naive estimator of $F_W(x)$ in the form

$$\widetilde{F}_W(x) := \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}\Im\left\{\frac{\widehat{\varphi}_Y(t)\varphi_\delta(t)}{\varphi_\varepsilon(t)}\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t. \tag{2.4}$$

Nevertheless, since $\lim_{t\to\infty}\varphi_\varepsilon(t) = 0$, the integrand in (2.4) could not be integrable and hence $\widetilde{F}_W(x)$ could not be well-defined. To overcome this, we truncate the integration domain to obtain the following estimator for $F_W(x)$:

$$\widehat{F}_{W;T}(x) := \frac{1}{2} - \frac{1}{\pi}\int_0^T \frac{1}{t}\Im\left\{\frac{\widehat{\varphi}_Y(t)\varphi_\delta(t)}{\varphi_\varepsilon(t)}\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t. \tag{2.5}$$

Here $T > 0$ is a suitably chosen parameter according to the sample size n later.

Remark 1 If there is no the presence of the Berkson error, i.e. $\delta \equiv 0$, then the estimator $\widehat{F}_{W;T}$ at (2.5) coincides with the estimator introduced in Dattner et al. (2011). Note that $\widehat{F}_{W;T}$ has two minor disadvantages: it may be not a non-decreasing function on $\mathbb{R}$; in addition, it may take some values outside the interval [0, 1]. However, the second disadvantage can be corrected easily. Indeed, we can set the correction version

$$\widetilde{F}_{W;T}(x) := \begin{cases} 0 & \text{if } \widehat{F}_{W;T}(x) < 0, \\ \widehat{F}_{W;T}(x) & \text{if } 0 \leq \widehat{F}_{W;T}(x) \leq 1, \\ 1 & \text{if } \widehat{F}_{W;T}(x) > 1. \end{cases}$$

It is obvious that $\widetilde{F}_{W;T}$ takes all possible values in [0, 1]. Also, the mean squared error of $\widetilde{F}_{W;T}$ is smaller than or equal to the one of $\widehat{F}_{W;T}$.

3 Consistency

Our purpose in this section is to study the convergence of the mean squared error between the estimator $\widehat{F}_{W;T}(x)$ and the target $F_W(x)$, defined by $\mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2$, according to the sample size n. Herein the notation $\mathbb{E}$ stands for the mathematical expectation with respect to the joint distribution of the random variables $Y_1, \ldots, Y_n$.

First, we derive a general bound for the bias of $\widehat{F}_{W;T}(x)$ in the following lemma.

Lemma 3.1 *Let the assumption (2.1) hold. In addition, assume that*

$$\int_r^\infty \frac{|\varphi_X(t)||\varphi_\delta(t)|}{t} \mathrm{d}t < \infty, \quad \textit{for some } r > 0. \tag{3.1}$$

Then we obtain for any $x \in \mathbb{R}$ that

$$|\mathbb{E}\widehat{F}_{W;T}(x) - F_W(x)| \leq \frac{1}{\pi} \int_T^\infty \frac{|\varphi_X(t)||\varphi_\delta(t)|}{t} \mathrm{d}t.$$

In Lemma 3.1, the upper bound for the bias of $\widehat{F}_{W;T}(x)$ is uniform with respect to x. Moreover, it decreases to zero when T is increased to the plus infinity.

In order to bound the variance $\mathrm{Var}\widehat{F}_{W;T}(x)$, we need the following assumption:

(A) There exist positive constants $b_\varepsilon, \tau_\varepsilon, \omega_\varepsilon$ such that $|\varphi_\varepsilon(t)|^2 \geq 1 - b_\varepsilon t^{\tau_\varepsilon}$ for all $t \in [0, \omega_\varepsilon]$.

Assumption **(A)** describes behavior of the characteristic function φ_ε in a neighborhood around the point $t = 0$. It holds for many usual densities, such as Gaussian, Cauchy, uniform densities and many others. For example, if ε has the Gaussian distribution with mean 0 and variance σ^2, then $\varphi_\varepsilon(t) = \mathrm{e}^{-(\sigma t)^2/2}$, $t \in \mathbb{R}$. Then $|\varphi_\varepsilon(t)|^2 = \mathrm{e}^{-(\sigma t)^2} \geq 1 - \sigma^2 t^2$, for all $t \in \mathbb{R}$, so **(A)** is satisfied with $b_\varepsilon \equiv \sigma^2$, $\tau_\varepsilon \equiv 2$ and $\omega_\varepsilon > 0$. Another example is the case where ε has the uniform distribution on the interval $(-a, a)$, $a > 0$. In that case, φ_ε is continuously differentiable up to order 2 on $\mathbb{R}$, $\varphi_\varepsilon(0) = 1$,

$\varphi_\varepsilon'(0) = \mathrm{i}\mathbb{E}(\varepsilon) = 0$ and $\varphi_\varepsilon''(0) = \mathrm{i}^2\mathbb{E}(\varepsilon^2) = -a^2/3$; hence, the applying of the Taylor expansion formula yields

$$\varphi_\varepsilon(t) = \varphi_\varepsilon(0) + \varphi_\varepsilon'(0)t + \frac{\varphi_\varepsilon''(0)}{2}(\theta t)^2 = 1 - \frac{a^2\theta^2 t^2}{6},$$

for some $\theta \in (0, 1)$. This gives $|\varphi_\varepsilon(t)|^2 \geq 1 - \frac{a^2\theta^2t^2}{3} \geq 1 - \frac{a^2t^2}{3}$, for all $t \in \mathbb{R}$. Thus, **(A)** is also satisfied for $b_\varepsilon \equiv a^2/3$, $\tau_\varepsilon \equiv 2$ and $\omega_\varepsilon > 0$.
In what follows, we set

$$\omega_* := \min\{1; \omega_\varepsilon; (2b_\varepsilon)^{-1/\tau_\varepsilon}\} \tag{3.2}$$

whenever the assumption **(A)** holds.
The following lemma provides a general upper bound for $\mathrm{Var}\widehat{F}_{W;T}(x)$.

Lemma 3.2 *Let the assumptions (2.1) and* **(A)** *hold. In addition, suppose* $\varphi_X\varphi_\varepsilon \in L^1(\mathbb{R})$. *Then we obtain for any* $x \in \mathbb{R}$ *that*

$$\mathrm{Var}\widehat{F}_{W;T}(x) \leq \frac{2}{\pi^2 n}\left(18 + \frac{2\ln^2 2}{\tau_\varepsilon^2}\right) + \frac{2\|\varphi_X\varphi_\varepsilon\|_1}{\pi^2 n}\int_{\omega_*}^{T}\frac{|\varphi_\delta(t)|^2}{t^2|\varphi_\varepsilon(t)|^2}\mathrm{d}t.$$

The upper bound in the latter lemma is also uniform with respect to x. However, it becomes bigger if T is bigger, unlike the upper bound in Lemma 3.1.
Applying the standard bias-variance decomposition

$$\mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = |\mathbb{E}\widehat{F}_{W;T}(x) - F_W(x)|^2 + \mathrm{Var}\widehat{F}_{W;T}(x)$$

and the results of Lemmas 3.1 and 3.2, we derive the following result.

Proposition 3.3 *Let the assumptions (2.1),* **(A)** *and (3.1) hold; in addition,* $\varphi_X\varphi_\varepsilon \in L^1(\mathbb{R})$. *Then we obtain*

$$\sup_{x\in\mathbb{R}}\mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 \leq \frac{1}{\pi^2}I_T^2 + \frac{2}{\pi^2 n}\left(18 + \frac{2\ln^2 2}{\tau_\varepsilon^2}\right) + \frac{2\|\varphi_X\varphi_\varepsilon\|_1}{\pi^2}J_{n,T},$$

where

$$I_T := \int_T^\infty \frac{|\varphi_X(t)||\varphi_\delta(t)|}{t}\mathrm{d}t, \quad J_{n,T} := \frac{1}{n}\int_{\omega_*}^{T}\frac{|\varphi_\delta(t)|^2}{t^2|\varphi_\varepsilon(t)|^2}\mathrm{d}t.$$

An important criterion to evaluate the quality of an estimation procedure is consistency. The following theorem establishes the (mean) consistency of the estimator $\widehat{F}_{W;T}$ under some regularity conditions on the functions φ_δ and φ_ε.

Theorem 3.4 *Let the assumptions of Proposition 3.3 hold.*

(a) *Suppose that there exist positive constants* C_δ, C_ε, $\widetilde{C}_\varepsilon$, β_δ *and* β_ε *such that*

$$|\varphi_\delta(t)| \le C_\delta(1+|t|)^{-\beta_\delta}, \quad \text{for all } t \in \mathbb{R}, \tag{3.3}$$

$$C_\varepsilon(1+|t|)^{-\beta_\varepsilon} \le |\varphi_\varepsilon(t)| \le \widetilde{C}_\varepsilon(1+|t|)^{-\beta_\varepsilon}, \quad \text{for all } t \in \mathbb{R}. \tag{3.4}$$

If the parameter T *depends on the sample size* n *such that*

$$\begin{cases} \lim_{n\to\infty} T = \infty & \text{if } \beta_\varepsilon < \beta_\delta + 1/2, \\ \lim_{n\to\infty} T = \infty,\ \lim_{n\to\infty} n^{-1}\ln T = 0 & \text{if } \beta_\varepsilon = \beta_\delta + 1/2, \\ \lim_{n\to\infty} T = \infty,\ \lim_{n\to\infty} n^{-1}T^{2(\beta_\varepsilon-\beta_\delta)-1} = 0 & \text{if } \beta_\varepsilon > \beta_\delta + 1/2, \end{cases} \tag{3.5}$$

then

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0.$$

(b) *Suppose that there exist positive constants* D_δ, D_ε, $\widetilde{D}_\varepsilon$, k_δ, k_ε, γ_δ *and* γ_ε *such that*

$$|\varphi_\delta(t)| \le D_\delta e^{-k_\delta|t|^{\gamma_\delta}}, \quad \text{for all } t \in \mathbb{R}, \tag{3.6}$$

$$D_\varepsilon e^{-k_\varepsilon|t|^{\gamma_\varepsilon}} \le |\varphi_\varepsilon(t)| \le \widetilde{D}_\varepsilon e^{-k_\varepsilon|t|^{\gamma_\varepsilon}}, \quad \text{for all } t \in \mathbb{R}. \tag{3.7}$$

If the parameter T *depends on the sample size* n *such that*

$$\begin{cases} \lim_{n\to\infty} T = \infty & \text{if } \gamma_\varepsilon < \gamma_\delta \text{ or } (\gamma_\varepsilon = \gamma_\delta, k_\varepsilon \le k_\delta), \\ \lim_{n\to\infty} T = \infty,\ \lim_{n\to\infty} \frac{e^{2(k_\varepsilon-k_\delta)T^\gamma}}{nT^{1+\gamma}} = 0 & \text{if } \gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0, k_\varepsilon > k_\delta, \\ \lim_{n\to\infty} T = \infty,\ \lim_{n\to\infty} \frac{e^{2k_\varepsilon T^{\gamma_\varepsilon}}}{nT^{1+\gamma_\varepsilon}} = 0 & \text{if } \gamma_\varepsilon > \gamma_\delta, \end{cases} \tag{3.8}$$

then

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0.$$

(c) *Suppose we have the conditions (3.3) and (3.7). If* T *depends on* n *in such a way that* $\lim_{n\to\infty} T = \infty$ *and* $\lim_{n\to\infty} n^{-1}T^{-(2\beta_\delta+\gamma_\varepsilon+1)}e^{2k_\varepsilon T^{\gamma_\varepsilon}} = 0$*, then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0.$$

(d) *Suppose we have the conditions (3.6) and (3.4). If* T *depends on* n *such that* $\lim_{n\to\infty} T = \infty$*, then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0.$$

We have some remarks on the conditions of $\varphi_\varepsilon, \varphi_\delta$ in the latter theorem. First, the conditions (3.4) and (3.7) are stronger than the condition (2.1). The first two conditions are closely related to the smoothness of the corresponding density f_ε of ε. Mathematically, the smoothness of a density corresponds to the decay rate of its characteristic function and vice verse. The decay rate becomes faster if the density is smoother. The characteristic function φ_ε from (3.4) has a polynomial decay rate at the infinity, so the density f_ε has a finite number of derivaties. While, the decay rate of the function φ_ε from (3.7) is exponential and thus the density f_ε is infinitely many times differentiable. Formally, the density f_ε is said to be ordinary smooth of order β_ε (respectively, super smooth of order γ_ε) if it holds (3.4) (respectively, (3.7)). For example, the Gamma and Laplace densities are ordinary smooth, while the Gaussian, Cauchy, t, logistic densities as well as finite mixtures of Gaussian densities are super smooth. The terminologies were first introduced by Fan (1991). Note that the conditions (3.4) and (3.7) are rather standard in nonparametric deconvolution topics, see, e.g. (Meister 2009) and references therein. Concerning the function φ_δ, we see that only the upper bounds on $|\varphi_\delta|$ are imposed in (3.3) and (3.6). This allows the function φ_δ to have some possible zeros on $\mathbb{R}$, unlike the cases of φ_ε where φ_ε must be non-vanishing on $\mathbb{R}$.
In view of Theorem 3.4, we realize that in the cases where the density f_δ of δ is smoother than the density f_ε, any selection of T with $\lim_{n\to\infty} T = \infty$ is sufficient for obtaining a mean consistency result of $\widehat{F}_{W;T}$. In any other case, the parameter T must be depended on the smoothness degrees of f_δ and f_ε.
Next, we state a mean consistency result of $\widehat{F}_{W;T}$ when the characteristic functions $\varphi_\delta, \varphi_\varepsilon$ satisfy an integrable condition.

Theorem 3.5 *Let the assumptions of Proposition 3.3 hold. In addition, assume that*

$$\int_\kappa^\infty \frac{|\varphi_\delta(t)|^2}{t^2|\varphi_\varepsilon(t)|^2}\mathrm{d}t < \infty, \quad \textit{for some } \kappa > 0. \tag{3.9}$$

If T depends on n such that $\lim_{n\to\infty} T = \infty$, then

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0.$$

A typical example for the condition (3.9) is the case where $\sup_{t>0} t^q|\varphi_\delta(t)/\varphi_\varepsilon(t)| < \infty$, for some $q > -1/2$. Of course, it is obvious that the condition also includes the cases: (i) the conditions (3.3) and (3.4) with $\beta_\varepsilon < \beta_\delta + 1/2$; (ii) the conditions (3.6) and (3.7) with $\gamma_\varepsilon < \gamma_\delta$, or $\gamma_\varepsilon = \gamma_\delta$ and $k_\varepsilon \le k_\delta$; (iii) the conditions (3.6) and (3.4).

4 Upper and minimax lower bounds

From now on, we focus on establishing convergence rates of the estimator $\widehat{F}_{W;T}$. To do that, we need some additional knowledge on the unknown density f_X of X. We introduce the following definition.

Definition 4.1 For $\alpha_X > -1/2$ and $L_X > 0$, we define $\mathcal{S}(\alpha_X, L_X)$ to be the set of all univariate densities f satisfying $\int_{-\infty}^{\infty} |\varphi_f(t)|^2(1+t^2)^{\alpha_X}\mathrm{d}t \le L_X$, where $\varphi_f(t) := \int_{-\infty}^{\infty} f(x)\mathrm{e}^{itx}\mathrm{d}x$, $t \in \mathbb{R}$.

The class $\mathcal{S}(\alpha_X, L_X)$ contains many usual univariate densities, e.g. the Gamma, Gaussian, Cauchy densities and many others. The parameter α_X represents smoothness degree for elements in this class. If $\alpha_X > 1/2$, then elements of the class are bounded continuous on $\mathbb{R}$; moreover, they have derivatives up to order ℓ, in which ℓ is the largest integer number satisfying $\ell < \alpha_X - 1/2$. The class is usually referred to as a Sobolev class of univariate densities.

Remark 2 If $f_X \in \mathcal{S}(\alpha_X, L_X)$ with $\alpha_X > -1/2$ and $L_X > 0$, then the condition (3.1) becomes evidently. This follows from the fact that, for any $r > 0$,

$$\int_r^{\infty} \frac{|\varphi_X(t)||\varphi_\delta(t)|}{t}\mathrm{d}t \le \int_r^{\infty} \frac{|\varphi_X(t)|(1+t^2)^{\alpha_X/2}(1+t^2)^{-\alpha_X/2}}{t}\mathrm{d}t$$
$$\le \sqrt{\frac{L_X}{2}\int_r^{\infty}\frac{(1+t^2)^{-\alpha_X}}{t^2}\mathrm{d}t} < \infty.$$

For the estimator $\widehat{F}_{W;T}$, its *maximal risk* over the class $\mathcal{S}(\alpha_X, L_X)$ is defined as

$$\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)] := \sup_{f_X \in \mathcal{S}(\alpha_X, L_X)} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2.$$

In the following theorem, we derive some upper bounds on the convergence rate of $\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)]$ under some different conditions on φ_δ and φ_ε.

Theorem 4.2 *Let $\alpha_X > -1/2$, $L_X > 0$ and the assumption* **(A)** *hold.*

(a) Suppose we have the conditions (3.3) and (3.4); in addition, suppose $\alpha_X + \beta_\varepsilon > 1/2$. Put

$$T_1 := \begin{cases} n^{1/(2\alpha_X+2\beta_\delta+1)} & \text{if } \beta_\varepsilon \le \beta_\delta + 1/2, \\ n^{1/(2\alpha_X+2\beta_\varepsilon)} & \text{if } \beta_\varepsilon > \beta_\delta + 1/2. \end{cases} \tag{4.1}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_1}; \mathcal{S}(\alpha_X, L_X)] \lesssim \begin{cases} n^{-1} & \text{if } \beta_\varepsilon < \beta_\delta + 1/2, \\ n^{-1}\ln n & \text{if } \beta_\varepsilon = \beta_\delta + 1/2, \\ n^{-(2\alpha_X+2\beta_\delta+1)/(2\alpha_X+2\beta_\varepsilon)} & \text{if } \beta_\varepsilon > \beta_\delta + 1/2. \end{cases}$$

(b) Suppose we have the conditions (3.6) and (3.7).

– *Consider the case $\gamma_\varepsilon < \gamma_\delta$ or ($\gamma_\varepsilon = \gamma_\delta$, $k_\varepsilon \le k_\delta$). Put*

$$T_2 := \left(\frac{\ln n - ((2\alpha_X + \gamma_\delta + 1)/\gamma_\delta)\ln\left(\frac{\ln n}{2k_\delta}\right)}{2k_\delta}\right)^{1/\gamma_\delta}. \tag{4.2}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_2}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-1}.$$

– *Consider the case* $\gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0$ *and* $k_\varepsilon > k_\delta$. *Put*

$$T_3 := \left(\frac{\ln n - (2\alpha_X/\gamma)\ln\left(\frac{\ln n}{2k_\varepsilon}\right)}{2k_\varepsilon} \right)^{1/\gamma}. \tag{4.3}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_3}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-k_\delta/k_\varepsilon}(\ln n)^{-[2\alpha_X(1-k_\delta/k_\varepsilon)+\gamma+1]/\gamma}.$$

– *Consider the case* $\gamma_\varepsilon > \gamma_\delta$. *Let* $T_4 \equiv T_4(n)$ *be the unique positive solution of the equation*

$$2\alpha_X \ln T_4 + 2k_\delta T_4^{\gamma_\delta} + 2k_\varepsilon T_4^{\gamma_\varepsilon} = \ln n + \ln\left(\frac{k_\delta \gamma_\delta}{k_\varepsilon \gamma_\varepsilon}\right). \tag{4.4}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_4}; \mathcal{S}(\alpha_X, L_X)] \lesssim (\ln n)^{-(2\alpha_X+\gamma_\delta+1)/\gamma_\varepsilon} \mathrm{e}^{-2k_\delta\left(\frac{\ln n}{2k_\varepsilon}\right)^{\gamma_\delta/\gamma_\varepsilon}[1+o(1)]}.$$

(c) Suppose we have the conditions (3.3) and (3.7). Put

$$T_5 := \left(\frac{\ln n - ((2\alpha_X - \gamma_\varepsilon)/\gamma_\varepsilon)\ln\left(\frac{\ln n}{2k_\varepsilon}\right)}{2k_\varepsilon} \right)^{1/\gamma_\varepsilon}. \tag{4.5}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_5}; \mathcal{S}(\alpha_X, L_X)] \lesssim (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}.$$

(d) Suppose we have the conditions (3.6) and (3.4) (with $\alpha_X + \beta_\varepsilon > 1/2$*). Then for* T_2 *as in (4.2) we obtain that*

$$\mathcal{R}[\widehat{F}_{W;T_2}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-1}.$$

In Theorem 4.2, all the convergence rates become better when the sample size n is increased. This is compatible with the mean consistency results stated in Theorems 3.4 and 3.5. The parameter rate $\mathcal{O}(n^{-1})$ is achieved in some cases, including the part (a) with $\beta_\varepsilon < \beta_\delta + 1/2$, the part (b) with $\gamma_\varepsilon < \gamma_\delta$, or $\gamma_\varepsilon = \gamma_\delta$ and $k_\varepsilon \le k_\delta$, and the part (d). In any other case, all corresponding rates depend on the smoothness degree α_X of the density f_X. The rates become faster when α_X is larger. Finally, in the parts

(a) and (c), if there is no presence of the Berkson error δ, i.e. $\beta_\delta = 0$ in (3.3), then we obtain the same rates as in the setting in which only the classical error model (1.2) is considered, see, e.g. (Dattner et al. 2011) and (Dattner and Reiser 2013).

Remark 3 In Theorem 4.2, the proposed parameters T_i $(i = 1, \ldots, 5)$ are all asymptotically optimal. However, since they are all depended on the class parameter α_X which is usually unknown in practice, the resulting estimators are non-adaptive and cannot be implemented. In fact, this does not pose a serious problem in the parts (b), (c) and (d) of the theorem. Indeed, all the convergence rates derived in the parts (b), (c) and (d) are still preserved when we replace T_i $(i = 2, \ldots, 5)$ by $\widetilde{T}_i$, where

$$\widetilde{T}_2 := \left(\frac{\ln n - (\ln(\ln n))^2}{2k_\delta}\right)^{1/\gamma_\delta}, \quad \widetilde{T}_3 := \left(\frac{\ln n - (\ln(\ln n))^2}{2k_\varepsilon}\right)^{1/\gamma},$$

$\widetilde{T}_4$ is the unique positive solution of the equation: $2k_\delta \widetilde{T}_4^{\gamma_\delta} + 2k_\varepsilon \widetilde{T}_4^{\gamma_\varepsilon} = \ln n - (\ln(\ln n))^2$,

$$\widetilde{T}_5 := \left(\frac{\ln n - (\ln(\ln n))^2}{2k_\varepsilon}\right)^{1/\gamma_\varepsilon}.$$

In principle, the parameters k_δ, k_ε, γ_δ and γ_ε are all known because the functions φ_δ and φ_ε are known exactly. So the parameters $\widetilde{T}_i$ for $i = 2, \ldots, 5$ are fully data-driven. In other words, we can implement the estimator $\widehat{F}_{W;T}$ with the following selections of the regularization parameter T:

$$T = \begin{cases} \widetilde{T}_2 & \text{under (3.6), (3.7) with } \gamma_\varepsilon < \gamma_\delta \text{ or } (\gamma_\varepsilon = \gamma_\delta, k_\varepsilon \le k_\delta), \\ \widetilde{T}_3 & \text{under (3.6), (3.7) with } \gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0 \text{ and } k_\varepsilon > k_\delta, \\ \widetilde{T}_4 & \text{under (3.6), (3.7) with } \gamma_\varepsilon > \gamma_\delta, \\ \widetilde{T}_5 & \text{under (3.3), (3.7)}, \\ \widetilde{T}_2 & \text{under (3.6), (3.4)}. \end{cases} \tag{4.6}$$

Concerning the case where φ_δ satisfies (3.3) and φ_ε satisfies (3.4), we set

$$\widetilde{T}_{1,1} := n^{1/(2\beta_\delta)} \quad \text{if } \beta_\varepsilon \le \beta_\delta + 1/2.$$

Then

$$\mathcal{R}[\widehat{F}_{W;\widetilde{T}_{1,1}}; \mathcal{S}(\alpha_X, L_X)] \lesssim \begin{cases} n^{-1} & \text{if } \beta_\varepsilon < \beta_\delta + 1/2, \\ n^{-1}\ln n & \text{if } \beta_\varepsilon = \beta_\delta + 1/2. \end{cases}$$

Thus, the convergence rates in Theorem 4.2(a) are also preserved with $\beta_\varepsilon \le \beta_\delta + 1/2$. So, in the implementation of $\widehat{F}_{W;T}$, we propose to choose

$$T = \widetilde{T}_{1,1} \quad \text{under the conditions (3.3), (3.4) with } \beta_\varepsilon \le \beta_\delta + 1/2. \tag{4.7}$$

Unlike as in the case $\beta_\varepsilon \le \beta_\delta + 1/2$, the problem of constructing an adaptive estimate of T_1 in the case $\beta_\varepsilon > \beta_\delta + 1/2$, say $\widetilde{T}_{1,2}$, such that the resulting estimator $\widehat{F}_{W;\widetilde{T}_{1,2}}$

and the estimator $\widehat{F}_{W;T_1}$ have the same convergence rate is in fact very intricate in our case. This problem is thus left for our future research.

When $\varphi_\delta, \varphi_\varepsilon$ satisfy **(A)** and (3.9), the estimator $\widehat{F}_{W;T}$ can achieve the parameter rate $\mathcal{O}(n^{-1})$, as shown in the following theorem.

Theorem 4.3 *Let $\alpha_X > -1/2$, $L_X > 0$ and the assumptions* **(A)**, *(3.9) hold. Put*

$$T_6 := n^{1/(2\alpha_X+1)}. \tag{4.8}$$

Then

$$\mathcal{R}[\widehat{F}_{W;T_6}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-1}.$$

In the rest of this section, we discuss optimality of our estimation procedure. First, the convergence rates derived in Theorem 4.3 as well as in the part (a) with $\beta_\varepsilon < \beta_\delta + 1/2$, the part (b) with $\gamma_\varepsilon < \gamma_\delta$, or $\gamma_\varepsilon = \gamma_\delta$ and $k_\varepsilon \leq k_\delta$, and the part (d) of Theorem 4.2 are of the form $\mathcal{O}(n^{-1})$. As known, the rate $\mathcal{O}(n^{-1})$ is the best rate for nonparametric estimation procedures. Hence, we conclude that the resulting estimators $\widehat{F}_{W;T}$ are minimax optimal in order in these cases. In the sequel, the convergence rates in the part (a) with $\beta_\varepsilon > \beta_\delta + 1/2$ and in the part (c) of Theorem 4.2 will also be demonstrated to be optimal. In order to do that, we formulate some lower bounds on the convergence rate of $\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)]$, for any possible estimator $\widehat{F}_W$ of F_W based on the data $Y_1, \ldots, Y_n$.

Theorem 4.4 *Let $\widehat{F}_W(x; Y_1, \ldots, Y_n)$ be an arbitrary estimator of $F_W(x)$ based on the data $Y_1, \ldots, Y_n$. Let $L_X > 0$ be large enough and $\alpha_X > 1/2$. Suppose the function $a(t) := \mathrm{Arg}\{\varphi_\delta(t)\}$ is twice continuously differentiable on $\mathbb{R}$ with $|a^{(j)}(t)| \leq a_* < \infty$, $j = 0, 1, 2$.*

(a) Suppose that we have (3.4) and that there exists a constant $\overline{C}_\varepsilon > 0$ such that

$$|\varphi_\varepsilon'(t)| \leq \overline{C}_\varepsilon (1 + |t|)^{-\beta_\varepsilon}, \quad \textit{for all } t \in \mathbb{R}. \tag{4.9}$$

In addition, suppose that there is a constant $\widetilde{C}_\delta > 0$ such that

$$|\varphi_\delta(t)| \geq \widetilde{C}_\delta (1 + |t|)^{-\beta_\delta}, \quad \textit{for all } t \in \mathbb{R}. \tag{4.10}$$

Then we obtain

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \gtrsim n^{-(2\alpha_X+2\beta_\delta+1)/(2\alpha_X+2\beta_\varepsilon)}.$$

(b) Suppose that we have (3.7) and that there exists a constant $\overline{D}_\varepsilon > 0$ such that

$$|\varphi_\varepsilon'(t)| \leq \overline{D}_\varepsilon \mathrm{e}^{-k_\varepsilon |t|^{\gamma_\varepsilon}}, \quad \textit{for all } t \in \mathbb{R}. \tag{4.11}$$

In addition, suppose that there is a constant $\widetilde{D}_\delta > 0$ such that

$$|\varphi_\delta(t)| \geq \widetilde{D}_\delta \mathrm{e}^{-k_\delta |t|^{\gamma_\delta}}, \quad \textit{for all } t \in \mathbb{R}. \tag{4.12}$$

Then we obtain

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \gtrsim (\ln n)^{-(2\alpha_X+1)/\gamma_\varepsilon} \mathrm{e}^{-2^{1+\gamma_\delta} k_\delta \left(\frac{\ln n}{2k_\varepsilon}\right)^{\gamma_\delta/\gamma_\varepsilon}}.$$

(c) *Let the conditions (4.10), (3.7) and (4.11) hold. Then we obtain*

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \gtrsim (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}.$$

Remark 4 We have the following conclusions from Theorems 4.2 and 4.4:

(a) Suppose φ_δ satisfies (3.3) and (4.10), and φ_ε satisfies (3.4) and (4.9). If $\beta_\varepsilon > \beta_\delta + 1/2$, then

$$\inf_{\widehat{F}_W} \mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \asymp n^{-(2\alpha_X+2\beta_\delta+1)/(2\alpha_X+2\beta_\varepsilon)}.$$

So $n^{-(2\alpha_X+2\beta_\delta+1)/(2\alpha_X+2\beta_\varepsilon)}$ is the minimax optimal rate of estimators over the class $\mathcal{S}(\alpha_X, L_X)$, and $\widehat{F}_{W;T_1}$ is a rate optimal estimator. However, if $\beta_\varepsilon = \beta_\delta + 1/2$, then the upper bound $\mathcal{O}(n^{-1}\ln n)$ of $\mathcal{R}[\widehat{F}_{W;T_1}; \mathcal{S}(\alpha_X, L_X)]$ does not match to the lower bound $\mathcal{O}(n^{-1})$ of $\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)]$, and thus the estimator $\widehat{F}_{W;T_1}$ is not optimal.

(b) Consider the case where φ_δ satisfies (3.3) and (4.10), and φ_ε satisfies (3.7) and (4.11). Then from Theorem 4.2(c) and Theorem 4.4(c) we infer that

$$\inf_{\widehat{F}_W} \mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \asymp (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}.$$

Thus, $(\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}$ is the minimax optimal rate of estimators over the class $\mathcal{S}(\alpha_X, L_X)$, and $\widehat{F}_{W;T_5}$ is also a rate optimal estimator.

5 Numerical examples

In this section, we present several numerical examples to illustrate finite sample performance of the estimator $\widehat{F}_{W;T}$ at (2.5). We also have some numerical comparisons between the estimator $\widehat{F}_{W;T}$ and another possible estimator of F_W. All calculations are conducted by using the R software.

For numerical tests, we have to compute $\widehat{F}_{W;T}(x)$ in many times. Hence, we need a robust method to perform this task fastly. Let us describe the method as follows. Suppose that we wish to compute numerically $\widehat{F}_{W;T}(x)$ on an interest interval $[B_1, B_2]$ which covers the support of W. Define $A_1 := 0$, $A_2 := T$,

$$v(t) := \frac{\varphi_\delta(t)\widehat{\varphi}_Y(t)}{t\varphi_\varepsilon(t)}, \quad t \in [A_1, A_2],$$

and

$$I_v(x) := \int_{A_1}^{A_2} v(t)\mathrm{e}^{-\mathrm{i}xt}\mathrm{d}t, \quad x \in [B_1, B_2].$$

Then we rewrite $\widehat{F}_{W;T}(x)$ from (2.5) in the form

$$\widehat{F}_{W;T}(x) = \frac{1}{2} - \frac{1}{\pi}\Im\{I_v(x)\}, \quad x \in [B_1, B_2]. \tag{5.1}$$

Denote $\delta_A := A_2 - A_1$, $\delta_B := B_2 - B_1$. For a fixed positive integer $M > 0$ large enough, we set $\delta_t := \delta_A/M$, $t_j := (j - 1/2)\delta_t + A_1$ for $j = \overline{1, M}$, $\delta_x := \delta_B/M$, $x_k := (k - 1/2)\delta_x + B_1$ for $k = \overline{1, M}$. We then approximate the integral I_v over the grid $\{x_k\}_{k=\overline{1,M}}$ as follows:

$$I_v(x_k) \approx \delta_t \sum_{j=1}^{M} v(t_j)\mathrm{e}^{-\mathrm{i}t_j x_k} = \delta_t \sum_{j=1}^{M} v(t_j)\mathrm{e}^{-\mathrm{i}(j\delta_t + A_1 - \delta_t/2)(k\delta_x + B_1 - \delta_x/2)}.$$

For brevity, put $\alpha := \delta_t\delta_x/2$, $\widetilde{A}_1 := A_1 - \delta_t/2$ and $\widetilde{B}_1 := B_1 - \delta_x/2$. We then derive

$$\begin{aligned} I_v(x_k) &\approx \delta_t \sum_{j=1}^{M} v(t_j)\mathrm{e}^{-\mathrm{i}(j\delta_t + \widetilde{A}_1)(k\delta_x + \widetilde{B}_1)} \\ &= \frac{\delta_A}{M}\sum_{j=1}^{M} v(t_j)\exp\left\{-2\alpha\mathrm{i}jk - \mathrm{i}j\frac{\widetilde{B}_1\delta_A}{M} - \mathrm{i}k\frac{\widetilde{A}_1\delta_B}{M} - \mathrm{i}\widetilde{A}_1\widetilde{B}_1\right\} \\ &= \frac{\delta_A}{M}\mathrm{e}^{-\mathrm{i}\widetilde{A}_1(\widetilde{B}_1 + k\delta_B/M)}\sum_{j=1}^{M}\left(v(t_j)\mathrm{e}^{-\mathrm{i}(\widetilde{B}_1\delta_A/M)j}\right)\mathrm{e}^{-2\alpha\mathrm{i}jk}. \end{aligned} \tag{5.2}$$

From (5.1) and (5.2), we obtain for $k = \overline{1, M}$ that

$$\widehat{F}_{W;T}(x_k) \approx \frac{1}{2} - \frac{1}{\pi}\Im\left\{\frac{\delta_A}{M}\mathrm{e}^{-\mathrm{i}\widetilde{A}_1(\widetilde{B}_1 + k\delta_B/M)}\sum_{j=1}^{M}\left(v_j\mathrm{e}^{-\mathrm{i}(\widetilde{B}_1\delta_A/M)j}\right)\mathrm{e}^{-2\alpha\mathrm{i}jk}\right\}. \tag{5.3}$$

The approximation (5.3) will be used to calculate $\widehat{F}_{W;T}$ over the point grid $\{x_k\}_{k=\overline{1,M}}$. Next, we present the results of our numerical experiment. We consider the two following examples for the distribution of W:

- (E1) $W = X + \delta$, where X has the Gamma distribution with shape parameter 4 and scale parameter $1/2$ and δ is chosen as below.
- (E2) $W = X + \delta$, where $X = U/\sqrt{8}$ with $U \sim 0.6\mathcal{L}(0, 1) + 0.4\mathcal{L}(5, 1)]$, and δ is chosen as below. Here $\mathcal{L}(m, s)$ denotes the Laplace distribution with location parameter m and scale parameter s.

The variable X has unit variance in both examples. When X is as in (E1), we have $\varphi_{f_X}(t) = (1 - \mathrm{i}t/2)^{-4}$ with $t \in \mathbb{R}$, and so $f_X \in \mathcal{S}(\alpha_X, L_X)$ with $\alpha_X < 7/2$ and $L_X > 0$ large enough. For X as in (E2), we have $f_X \in \mathcal{S}(\alpha_X, L_X)$ with $\alpha_X < 3/2$ and $L_X > 0$ large enough.
For each the example, we choose the error variables δ, ε such that the signal ratios σ_δ/σ_X and $\sigma_\varepsilon/\sigma_X$ are equal to 0.5, corresponding to 50% noise levels. Herein σ_X, σ_δ and σ_ε are the standard deviations of X, δ and ε, respectively. More concretely, we consider the four following cases of δ, ε:

(C1) $\delta \sim \mathcal{L}(0, 1/\sqrt{8})$ and $\varepsilon \sim \mathcal{L}(0, 1/\sqrt{8})$. In that case, $\varphi_\delta(t) = \varphi_\varepsilon(t) = 8/(t^2 + 8)$, and hence φ_δ and φ_ε satisfy (3.3) and (3.4), respectively, with $\beta_\delta = \beta_\varepsilon = 2$. In view of (4.7), we choose $T = n^{1/4}$.
(C2) $\delta \sim \mathcal{L}(0, 1/\sqrt{8})$ and $\varepsilon \sim \mathcal{N}(0, (1/2)^2)$, i.e. ε has the Gaussian distribution with mean 0 and variance $(1/2)^2$. Then $\varphi_\delta(t) = 8/(t^2 + 8)$ and $\varphi_\varepsilon(t) = \mathrm{e}^{-t^2/8}$. So φ_δ satisfy (3.3) with $\beta_\delta = 2$, and φ_ε satisfy (3.7) with $\gamma_\varepsilon = 2$, $k_\varepsilon = 1/8$. By (4.6), we choose $T = 2\sqrt{\ln n - (\ln(\ln n))^2}$.
(C3) $\delta \sim \mathcal{N}(0, (1/2)^2)$ and $\varepsilon \sim \mathcal{L}(0, 1/\sqrt{8})$. Then $\varphi_\delta(t) = \mathrm{e}^{-t^2/8}$ and $\varphi_\varepsilon(t) = 8/(t^2 + 8)$. Hence, φ_δ satisfy (3.6) with $\gamma_\delta = 2$, $k_\delta = 1/8$, and φ_ε satisfy (3.4) with $\beta_\varepsilon = 2$. By (4.6), we choose $T = 2\sqrt{\ln n - (\ln(\ln n))^2}$.
(C4) $\varepsilon \sim \mathcal{N}(0, (1/2)^2)$ and $\delta \sim \mathcal{N}(0, (1/2)^2)$. In that case, $\varphi_\delta(t) = \varphi_\varepsilon(t) = \mathrm{e}^{-t^2/8}$. Hence, φ_δ satisfy (3.6) with $\gamma_\delta = 2$, $k_\delta = 1/8$, and φ_ε satisfy (3.7) with $\gamma_\varepsilon = 2$, $k_\varepsilon = 1/8$. In this case, by (4.6), we select $T = 2\sqrt{\ln n - (\ln(\ln n))^2}$.

As mentioned in Sect. 1, Delaigle (2007) considered the problem of estimating the density f_W of W. The author estimated $f_W(x)$ by

$$\widehat{f}_W^{\mathrm{De}}(x) := \begin{cases} \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\varphi_\delta(t)\widehat{\varphi}_Y(t)}{\varphi_\varepsilon(t)} \mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t & \text{if } \varphi_\delta/\varphi_\varepsilon \in L^1(\mathbb{R}), \\ \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_K(ht)\frac{\varphi_\delta(t)\widehat{\varphi}_Y(t)}{\varphi_\varepsilon(t)} \mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t & \text{if } \varphi_\delta/\varphi_\varepsilon \notin L^1(\mathbb{R}), \end{cases}$$

where K is a kernel with compactly supported and $h > 0$ is a bandwidth parameter. In view of the relation $F_W(x) = \int_{-\infty}^{x} f_W(u)\mathrm{d}u$, it is possible to estimate the cdf $F_W(x)$ at every x by the plug-in type of estimator

$$\widehat{F}_W^{\mathrm{De}}(x) := \int_{-\infty}^{x} \widehat{f}_W^{\mathrm{De}}(u)\mathrm{d}u,$$

called Delaigle type of estimator. We will compare numerically the two estimators $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$. To calculate the estimator $\widehat{F}_W^{\mathrm{De}}$, the kernel K is chosen of the form

$$K(x) := \frac{48\cos(x)}{\pi x^4}\left(1 - \frac{15}{x^2}\right) - \frac{144\sin(x)}{\pi x^5}\left(2 - \frac{5}{x^2}\right).$$

This kernel has order 2 and been commonly used in deconvolution problems (see, e.g. Wand (1998), Delaigle and Hall (2006)). It was also used in the numerical simulations of Delaigle (2007). Its Fourier transform is of the form $\varphi_K(t) = (1 - t^2)^3\mathbb{I}_{[-1,1]}(t)$, which is supported in $[-1, 1]$. Herein $\mathbb{I}_{[-1,1]}$ denotes the indicator function of $[-1, 1]$.

Regarding h, in the cases (C1) and (C4), $\varphi_\delta/\varphi_\varepsilon = 1 \notin L^1(\mathbb{R})$, so we choose $h = n^{-1/5}$, as suggested in Delaigle (2007), Theorem 3. In the case (C2), $\varphi_\delta(t)/\varphi_\varepsilon(t) = 8(t^2 + 8)^{-1}e^{t^2/8} \notin L^1(\mathbb{R})$; hence, by Delaigle (2007), Theorem 4, we select $h = 1/(2\sqrt{\ln n})$. For the case (C3), we see that $\varphi_\delta/\varphi_\varepsilon \in L^1(\mathbb{R})$ and thus we do not need to choose any bandwidth in the computation of $\widehat{F}_W^{\mathrm{De}}$.

In each example of X and each case of the errors δ and ε, we obtain the data sample $Y_1, \ldots, Y_n$ based on the formula $Y_j = X_j + \varepsilon_j$ for $j = \overline{1, n}$, where $X_1, \ldots, X_n$ are generated randomly from the distribution of X and $\varepsilon_1, \ldots, \varepsilon_n$ are generated randomly from the distribution of ε. To have data samples for computing approximately mean squared errors, the process of generating data sample $Y_1, \ldots, Y_n$ is replicated in R times independently. In this experiment, we take $R := 500$. The data sample in the q-th repetition is denoted by $(Y_1^{(q)}, \ldots, Y_n^{(q)})$, $q = \overline{1, R}$. Afterwards, the mean squared errors of $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$ at x are respectively approximated by so-called empirical mean squared errors of $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$ at x, defined respectively as

$$\mathrm{EMSE}(\widehat{F}_{W;T}; x) := \frac{1}{R}\sum_{q=1}^{R} |\widehat{F}_{W;T}(x; Y_1^{(q)}, \ldots, Y_n^{(q)}) - F_W(x)|^2,$$

$$\mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x) := \frac{1}{R}\sum_{q=1}^{R} |\widehat{F}_W^{\mathrm{De}}(x; Y_1^{(q)}, \ldots, Y_n^{(q)}) - F_W(x)|^2.$$

Let $M := 300$, $B_1 = -8$, $B_2 = 12$ and $\{x_k\}_{k=\overline{1,M}}$ be the equidistant point grid in $[B_1, B_2]$, as mentioned in the first part of this section. In Tables 1 and 2, we report the values of the maximum empirical mean squared errors $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_{W;T}; x_k)$ and $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x_k)$ for $n = 100$ (the left column) and $n = 500$ (the right column) and for different scenarios of δ, ε. Note that the last two quantities are respectively the estimates of the worst mean squared errors $\sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2$ and $\sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_W^{\mathrm{De}}(x) - F_W(x)|^2$. Observe from the tables that, as expected, the values of $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_{W;T}; x_k)$ and $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x_k)$ decrease when the sample sizes n increase. This confirms the convergence of $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$.

In order to visualize both numerical estimators and their convergence tendency, in Figs. 1 and 2, we plot three curves with black, red and blue colors, where the black curve is the graph of the target cdf F_W, the red curve is the graph of $\widehat{F}_{W;T}$, and the blue curve is the graph of $\widehat{F}_W^{\mathrm{De}}$. From the top to the bottom of each the figure, the left and right sub-figures are respectively plotted with $n = 100$ and $n = 500$. Obviously, the larger the sample size n is, the closer the curves become. This reflects the convergence of the estimators $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$. In comparison of the two estimators $\widehat{F}_{W;T}$ and $\widehat{F}_W^{\mathrm{De}}$, we see in Figs. 1 and 2 that the red curves are always better than the blue curves in the cases (C1), (C2) and (C4); however, in the case (C3) where the estimator $\widehat{F}_W^{\mathrm{De}}$ does not contain any regularization parameter, the blue curves are better than the red curves. These observations are consistent with the numerical results in Tables 1 and 2.

Finally, it seems that the convergence tendency of the numerical estimators in Fig. 1 is shown more clearly than the one in Fig. 2. Also, in each the case of the errors δ and ε, the values of $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_{W;T}; x_k)$ and $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x_k)$ in Table

Table 1 The values of $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_{W;T}; x_k)$ and $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x_k)$ when W is as in (E1)

δ, ε	$n = 100$ $\widehat{F}_{W;T}$	$\widehat{F}_W^{\mathrm{De}}$	$n = 500$ $\widehat{F}_{W;T}$	$\widehat{F}_W^{\mathrm{De}}$
(C1)	0.00281	0.02047	0.00192	0.01115
(C2)	0.00285	0.00822	0.00205	0.00598
(C3)	0.00321	0.00309	0.00250	0.00242
(C4)	0.00321	0.02092	0.00248	0.01140

Table 2 The values of $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_{W;T}; x_k)$ and $\max_{k=\overline{1,M}} \mathrm{EMSE}(\widehat{F}_W^{\mathrm{De}}; x_k)$ when W is as in (E2)

δ, ε	$n = 100$ $\widehat{F}_{W;T}$	$\widehat{F}_W^{\mathrm{De}}$	$n = 500$ $\widehat{F}_{W;T}$	$\widehat{F}_W^{\mathrm{De}}$
(C1)	0.00628	0.02187	0.00365	0.01378
(C2)	0.00692	0.01172	0.00478	0.00887
(C3)	0.00696	0.00580	0.00554	0.00488
(C4)	0.00729	0.02225	0.00542	0.01384

1 are respectively smaller than those in Table 2. These observations can be explained by the fact that the distribution of W in (E1) is smoother than the one in (E2).

6 Proofs

Proof of Lemma 3.1 Using the Fubini theorem and the equality $\mathbb{E}\widehat{\varphi}_Y(t) = \varphi_X(t)\varphi_\varepsilon(t)$ (for $t \in \mathbb{R}$) gives

$$\begin{aligned}\mathbb{E}\widehat{F}_{W;T}(x) &= \frac{1}{2} - \frac{1}{\pi}\int_0^T \frac{1}{t}\Im\left\{\frac{\varphi_\delta(t)\mathbb{E}\widehat{\varphi}_Y(t)}{\varphi_\varepsilon(t)}\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t \\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^T \frac{1}{t}\Im\left\{\varphi_X(t)\varphi_\delta(t)\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t. \end{aligned} \tag{6.1}$$

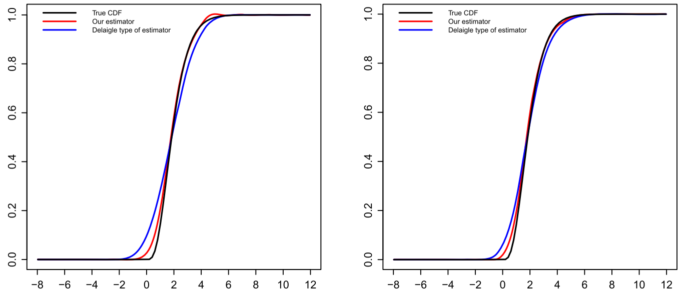
It follows from (6.1), (2.2) and the inequality $|\Im\{z\}| \le |z|$ (for $z \in \mathbb{C}$) that

$$|\mathbb{E}\widehat{F}_{W;T}(x) - F_W(x)| = \left|\frac{1}{\pi}\int_T^\infty \frac{1}{t}\Im\left\{\varphi_X(t)\varphi_\delta(t)\mathrm{e}^{-\mathrm{i}tx}\right\}\mathrm{d}t\right| \le \frac{1}{\pi}\int_T^\infty \frac{|\varphi_X(t)||\varphi_\delta(t)|}{t}\mathrm{d}t,$$

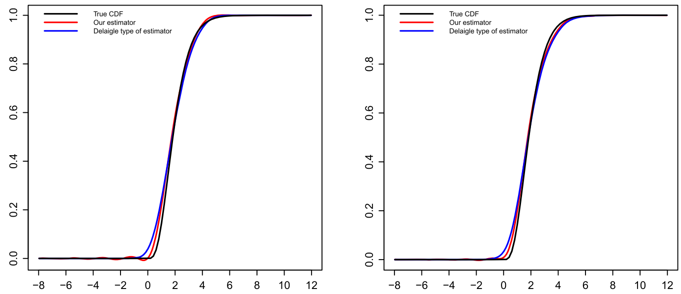
and this completes the proof. □

Proof of Lemma 3.2 First, we rewrite the estimator $\widehat{F}_{W;T}(x)$ in the form

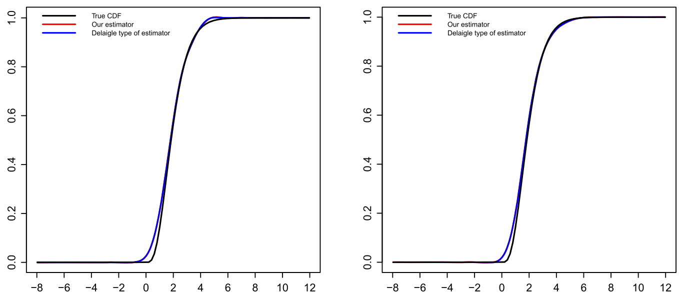
$$\widehat{F}_{W;T}(x) = \frac{1}{2} - \frac{1}{n\pi}\sum_{j=1}^n U_{j,T}(x),$$



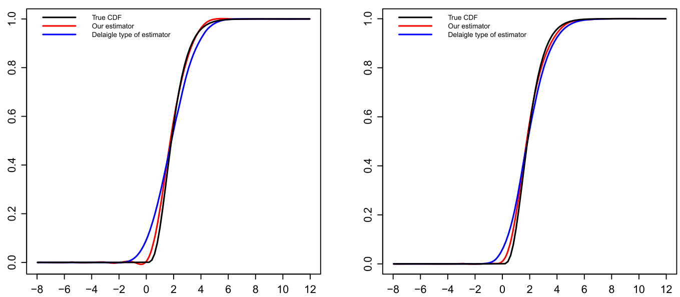
(a) δ, ε are in (C1).



(b) δ, ε are in (C2).

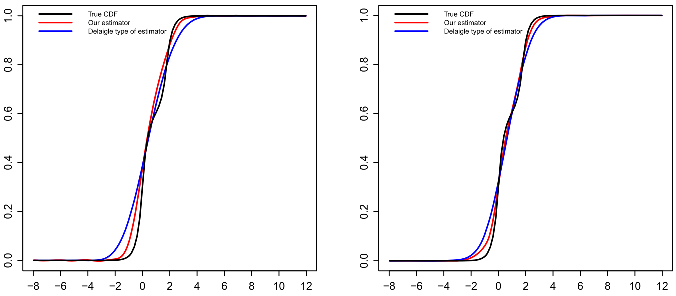


(c) δ, ε are in (C3).

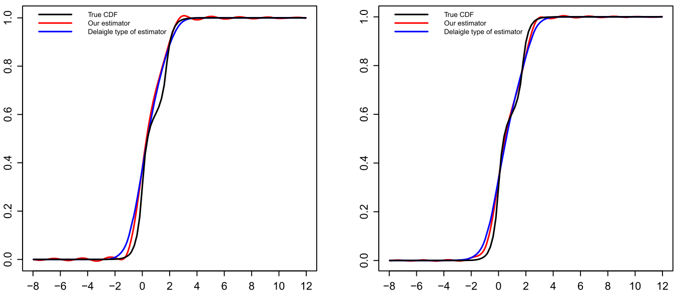


(d) δ, ε are in (C4).

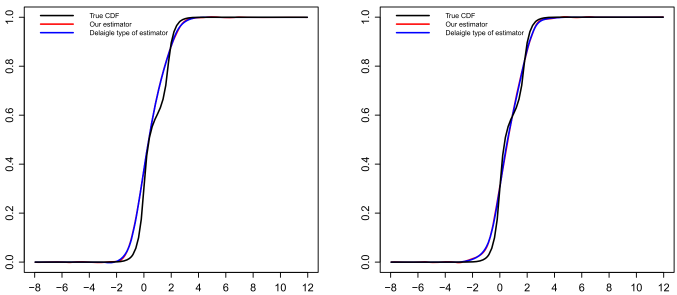
Fig. 1 The variable W is as in (E1). The left and right sub-figures are plotted with $n = 100$ and $n = 500$, respectively



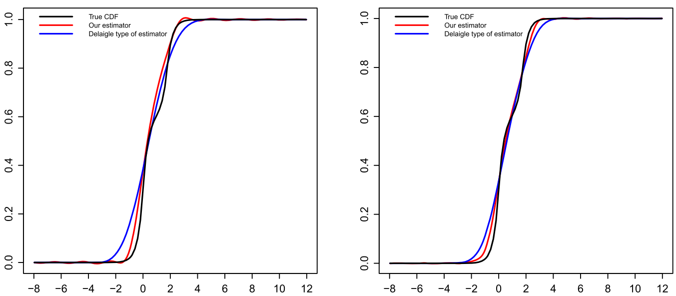
(a) δ, ε are in (C1).



(b) δ, ε are in (C2).



(c) δ, ε are in (C3).



(d) δ, ε are in (C4).

Fig. 2 The variable W is as in (E2). The left and right sub-figures are plotted with $n = 100$ and $n = 500$, respectively where

$$U_{j,T}(x) := \int_0^T \frac{1}{t|\varphi_\varepsilon(t)|^2} \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_j - x)} \right\} dt.$$

Since $U_{1,T}(x), \ldots, U_{n,T}(x)$ are i.i.d. random variables, we obtain

$$\operatorname{Var} \widehat{F}_{W;T}(x) = \frac{1}{n\pi^2} \operatorname{Var} U_{1,T}(x) \le \frac{1}{n\pi^2} \mathbb{E}|U_{1,T}(x)|^2.$$

After that using the inequality $|u+v|^2 \le 2|u|^2 + 2|v|^2$ $(u, v \in \mathbb{R})$ gives

$$\operatorname{Var} \widehat{F}_{W;T}(x) \le \frac{2}{\pi^2 n} V_1 + \frac{2}{\pi^2 n} V_2 \tag{6.2}$$

with

$$V_1 := \mathbb{E} \left| \int_0^{\omega_*} \frac{1}{t|\varphi_\varepsilon(t)|^2} \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_1 - x)} \right\} dt \right|^2,$$

$$V_2 := \mathbb{E} \left| \int_{\omega_*}^T \frac{1}{t|\varphi_\varepsilon(t)|^2} \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_1 - x)} \right\} dt \right|^2.$$

Here ω_* is defined by (3.2).

Estimate V_1: For all $t \in (0, \omega_*)$, we have $1 - |\varphi_\varepsilon(t)|^2 \le b_\varepsilon t^{\tau_\varepsilon} \le b_\varepsilon \omega_*^{\tau_\varepsilon} \le 1/2$, so we apply the Taylor expansion formula to obtain

$$\frac{1}{|\varphi_\varepsilon(t)|^2} = \frac{1}{1 - (1 - |\varphi_\varepsilon(t)|^2)} = \sum_{k=0}^{\infty} (1 - |\varphi_\varepsilon(t)|^2)^k = 1 + \sum_{k=1}^{\infty} (1 - |\varphi_\varepsilon(t)|^2)^k.$$

Thus,

$$V_1 = \mathbb{E} \left| \int_0^{\omega_*} \frac{1}{t} \left(1 + \sum_{k=1}^{\infty} (1 - |\varphi_\varepsilon(t)|^2)^k \right) \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_1 - x)} \right\} dt \right|^2.$$

Once again, applying the inequality $|u+v|^2 \le 2|u|^2 + 2|v|^2$ $(u, v \in \mathbb{R})$ gives

$$V_1 \le 2V_{1,1} + 2V_{1,2}$$

with

$$V_{1,1} := \mathbb{E} \left| \int_0^{\omega_*} \frac{1}{t} \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_1 - x)} \right\} dt \right|^2,$$

$$V_{1,2} := \mathbb{E} \left| \int_0^{\omega_*} \frac{1}{t} \Im \left\{ \varphi_\varepsilon(-t)\varphi_\delta(t) e^{it(Y_1 - x)} \right\} \sum_{k=1}^{\infty} (1 - |\varphi_\varepsilon(t)|^2)^k dt \right|^2.$$

Using the Fubini theorem, we get

$$\begin{aligned} V_{1,1} &= \mathbb{E}\left|\int_0^{\omega_*} \frac{1}{t}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v)\sin(t(Y_1 - x - u + v))\mathrm{d}u\mathrm{d}v\right)\mathrm{d}t\right|^2 \\ &= \mathbb{E}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v)\left(\int_0^{\omega_*}\frac{\sin(t(Y_1 - x - u + v))}{t}\mathrm{d}t\right)\mathrm{d}u\mathrm{d}v\right|^2 \\ &\le \mathbb{E}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v)\left|\int_0^{\omega_*}\frac{\sin(t(Y_1 - x - u + v))}{t}\mathrm{d}t\right|\mathrm{d}u\mathrm{d}v\right)^2 \\ &= \mathbb{E}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v)\left|\int_0^{(Y_1 - x - u + v)\omega_*}\frac{\sin(\tau)}{\tau}\mathrm{d}\tau\right|\mathrm{d}u\mathrm{d}v\right)^2. \end{aligned}$$

By the fact that $\sup_{a>0}\left|\int_0^a \tau^{-1}\sin(\tau)\mathrm{d}\tau\right| \le 3$ (see, e.g. Kawata (1972)), we deduce

$$V_{1,1} \le \left(3\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v)\mathrm{d}u\mathrm{d}v\right)^2 = 9.$$

Next, using **(A)** and the equality $\sum_{k=1}^{\infty} x^k/k = \ln[1/(1-x)]$ (for $0 < x < 1$), we have

$$\begin{aligned} V_{1,2} &\le \left(\int_0^{\omega_*}\frac{1}{t}\sum_{k=1}^{\infty}(1 - |\varphi_\varepsilon(t)|^2)^k\mathrm{d}t\right)^2 \le \left(\int_0^{\omega_*}\frac{1}{t}\sum_{k=1}^{\infty}(b_\varepsilon t^{\tau_\varepsilon})^k\mathrm{d}t\right)^2 \\ &= \left(\frac{1}{\tau_\varepsilon}\sum_{k=1}^{\infty}\frac{(b_\varepsilon\omega_*^{\tau_\varepsilon})^k}{k}\right)^2 = \left(\frac{1}{\tau_\varepsilon}\ln\left(\frac{1}{1 - b_\varepsilon\omega_*^{\tau_\varepsilon}}\right)\right)^2 \le \frac{\ln^2 2}{\tau_\varepsilon^2}. \end{aligned}$$

From the bounds of $V_{1,1}$ and $V_{1,2}$, we obtain

$$V_1 \le 18 + \frac{2\ln^2 2}{\tau_\varepsilon^2}. \tag{6.3}$$

Estimate V_2: By the Fubini theorem,

$$\begin{aligned} V_2 &= \left|\mathbb{E}\int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{\Im\left\{\varphi_\varepsilon(-t)\varphi_\delta(t)\mathrm{e}^{\mathrm{i}t(Y_1 - x)}\right\}\Im\left\{\varphi_\varepsilon(-s)\varphi_\delta(s)\mathrm{e}^{\mathrm{i}s(Y_1 - x)}\right\}}{ts|\varphi_\varepsilon(t)|^2|\varphi_\varepsilon(s)|^2}\mathrm{d}t\mathrm{d}s\right| \\ &= \left|\int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{1}{ts|\varphi_\varepsilon(t)|^2|\varphi_\varepsilon(s)|^2}\left(\int_{-\infty}^{\infty} A(t,s,y) f_Y(y)\mathrm{d}y\right)\mathrm{d}t\mathrm{d}s\right| \\ &\le \int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{1}{ts|\varphi_\varepsilon(t)|^2|\varphi_\varepsilon(s)|^2}\left|\int_{-\infty}^{\infty} A(t,s,y) f_Y(y)\mathrm{d}y\right|\mathrm{d}t\mathrm{d}s, \end{aligned}$$

where

$$A(t,s,y) := \Im\left\{\varphi_\varepsilon(-t)\varphi_\delta(t)\mathrm{e}^{\mathrm{i}t(y-x)}\right\}\Im\left\{\varphi_\varepsilon(-s)\varphi_\delta(s)\mathrm{e}^{\mathrm{i}s(y-x)}\right\}.$$

Using the Fubini theorem and the equality $\sin(a)\sin(b) = 2^{-1}[\cos(a-b)-\cos(a+b)]$ with $a, b \in \mathbb{R}$, we obtain

$$\begin{aligned}A(t,s,y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v) f_\varepsilon(p) f_\delta(q)\sin[t(y-x-u+v)]\\ &\quad\times \sin[s(y-x-p+q)]\mathrm{d}u\mathrm{d}v\mathrm{d}p\mathrm{d}q\\ &= \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v) f_\varepsilon(p) f_\delta(q)\\ &\quad\times \cos[y(t-s)+x(s-t)-ut+vt+ps-qs]\mathrm{d}u\mathrm{d}v\mathrm{d}p\mathrm{d}q\\ &\quad-\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\varepsilon(u) f_\delta(v) f_\varepsilon(p) f_\delta(q)\\ &\quad\times \cos[y(t+s)-x(t+s)-ut+vt-ps+qs]\mathrm{d}u\mathrm{d}v\mathrm{d}p\mathrm{d}q,\end{aligned}$$

so

$$\begin{aligned}\int_{-\infty}^{\infty} A(t,s,y) f_Y(y)\mathrm{d}y = \frac{1}{2}\Re\{\varphi_Y(t-s)\varphi_\varepsilon(-t)\varphi_\delta(t)\varphi_\varepsilon(s)\varphi_\delta(-s)\mathrm{e}^{\mathrm{i}(s-t)x}\}\\ -\frac{1}{2}\Re\{\varphi_Y(t+s)\varphi_\varepsilon(-t)\varphi_\delta(t)\varphi_\varepsilon(-s)\varphi_\delta(s)\mathrm{e}^{-\mathrm{i}(t+s)x}\},\end{aligned}$$

which yields

$$\left|\int_{-\infty}^{\infty} A(t,s,y) f_Y(y)\mathrm{d}y\right| \le \frac{(|\varphi_Y(t+s)|+|\varphi_Y(t-s)|)|\varphi_\varepsilon(t)||\varphi_\delta(t)||\varphi_\varepsilon(s)||\varphi_\delta(s)|}{2}.$$

Therefore,

$$V_2 \le \frac{1}{2}(V_{2,1}+V_{2,2}),$$

with

$$\begin{aligned}V_{2,1} &:= \int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{|\varphi_Y(t+s)||\varphi_\delta(t)||\varphi_\delta(s)|}{ts|\varphi_\varepsilon(t)||\varphi_\varepsilon(s)|}\mathrm{d}t\mathrm{d}s,\\ V_{2,2} &:= \int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{|\varphi_Y(t-s)||\varphi_\delta(t)||\varphi_\delta(s)|}{ts|\varphi_\varepsilon(t)||\varphi_\varepsilon(s)|}\mathrm{d}t\mathrm{d}s.\end{aligned}$$

Using the Cauchy-Schwarz inequality and the Fubini theorem yields

$$\begin{aligned}V_{2,1} &\le \left(\int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{|\varphi_Y(t+s)||\varphi_\delta(t)|^2}{t^2|\varphi_\varepsilon(t)|^2}\mathrm{d}t\mathrm{d}s\cdot\int_{\omega_*}^{T}\int_{\omega_*}^{T}\frac{|\varphi_Y(t+s)||\varphi_\delta(s)|^2}{s^2|\varphi_\varepsilon(s)|^2}\mathrm{d}t\mathrm{d}s\right)^{1/2}\\ &= \left(\int_{\omega_*}^{T}\left(\int_{\omega_*}^{T}|\varphi_Y(t+s)|\mathrm{d}s\right)\frac{|\varphi_\delta(t)|^2}{t^2|\varphi_\varepsilon(t)|^2}\mathrm{d}t\right.\end{aligned}$$

$$\cdot \int_{\omega_*}^{T} \left(\int_{\omega_*}^{T} |\varphi_Y(t+s)| \mathrm{d}t \right) \frac{|\varphi_\delta(s)|^2}{s^2 |\varphi_\varepsilon(s)|^2} \mathrm{d}s \Bigg)^{1/2}$$
$$\leq \|\varphi_X \varphi_\varepsilon\|_1 \int_{\omega_*}^{T} \frac{|\varphi_\delta(t)|^2}{t^2 |\varphi_\varepsilon(t)|^2} \mathrm{d}t.$$

The same bound holds for $V_{2,2}$. Hence, we derive

$$V_2 \leq \|\varphi_X \varphi_\varepsilon\|_1 \int_{\omega_*}^{T} \frac{|\varphi_\delta(t)|^2}{t^2 |\varphi_\varepsilon(t)|^2} \mathrm{d}t. \tag{6.4}$$

Finally, combining (6.2) with (6.3) and (6.4) gives the desired result. □

Lemma 6.1 *For* $0 < s < 1 < T$, *we have*

$$\int_s^T \frac{(1+t)^{2\beta}}{t^2} \mathrm{d}t \leq \begin{cases} \frac{1}{s} & \text{if } \beta \leq 0, \\ \frac{2^{2\beta}}{s} + \frac{2^{2\beta}}{1-2\beta} & \text{if } 0 < \beta < 1/2, \\ \frac{1}{s} + \ln\left(\frac{T}{s}\right) & \text{if } \beta = 1/2, \\ \frac{2^{2\beta}}{s} + \frac{2^{2\beta}}{2\beta-1} T^{2\beta-1} & \text{if } \beta > 1/2. \end{cases}$$

Proof For brevity, put $I := \int_s^T (1+t)^{2\beta}/t^2 \mathrm{d}t$. For $\beta \leq 0$, we have

$$I \leq \int_s^T \frac{1}{t^2} \mathrm{d}t = \frac{1}{s} - \frac{1}{T} \leq \frac{1}{s}.$$

For $\beta = 1/2$,

$$I = \int_s^T \left(\frac{1}{t^2} + \frac{1}{t} \right) \mathrm{d}t = \frac{1}{s} - \frac{1}{T} + \ln\left(\frac{T}{s}\right) \leq \frac{1}{s} + \ln\left(\frac{T}{s}\right).$$

For $0 < \beta \neq 1/2$, we have

$$I = \int_s^1 \frac{(1+t)^{2\beta}}{t^2} \mathrm{d}t + \int_1^T \frac{(1+t)^{2\beta}}{t^2} \mathrm{d}t \leq 2^{2\beta} \int_s^1 \frac{1}{t^2} \mathrm{d}t + 2^{2\beta} \int_1^T t^{2\beta-2} \mathrm{d}t$$
$$\leq \frac{2^{2\beta}}{s} + \frac{2^{2\beta}}{2\beta - 1} (T^{2\beta-1} - 1) \leq \begin{cases} \frac{2^{2\beta}}{s} + \frac{2^{2\beta}}{2\beta-1} T^{2\beta-1} & \text{for } \beta > 1/2, \\ \frac{2^{2\beta}}{s} + \frac{2^{2\beta}}{1-2\beta} & \text{for } 0 < \beta < 1/2. \end{cases}$$

The lemma is proved. □

Lemma 6.2 *Let* $A, r > 0$. *There exists a constant* $C_1 > 0$ *depending on* A, r *such that*

$$\int_M^\infty \mathrm{e}^{-At^r} \mathrm{d}t \leq C_1 M^{1-r} \mathrm{e}^{-AM^r}, \quad \textit{for any } M > 1.$$

Proof Consider the case $r \geq 1$. Then

$$\int_M^\infty \mathrm{e}^{-At^r}\mathrm{d}t \leq \int_M^\infty \left(\frac{t}{M}\right)^{r-1} \mathrm{e}^{-At^r}\mathrm{d}t = -(Ar)^{-1}M^{1-r}\int_M^\infty (-Ar)t^{r-1}\mathrm{e}^{-At^r}\mathrm{d}t$$
$$= (Ar)^{-1}M^{1-r}\mathrm{e}^{-AM^r}.$$

Next, we consider the case $0 < r < 1$. Put $J_p := \int_M^\infty t^{-pr}\mathrm{e}^{-At^r}\mathrm{d}t$, $p \geq 0$. Using the integration by parts gives

$$J_p = \frac{1}{Ar}M^{1-(p+1)r}\mathrm{e}^{-AM^r} + \frac{1-(p+1)r}{Ar}J_{p+1}.$$

Let k be the smallest positive integer number greater than or equal to $1/r$. Using the latter recurrence relation, we can find constants $C_0, C_0' > 0$ depending only on A, r such that

$$J_0 \leq C_0\mathrm{e}^{-AM^r}(M^{1-r} + M^{1-2r} + \cdots + M^{1-(k-1)r}) + C_0' J_{k-1}.$$

Note that since $kr \geq 1$ we have $(t/M)^{kr-1} \geq 1$ for all $t \geq M$, and so

$$J_{k-1} = \int_M^\infty t^{r-kr}\mathrm{e}^{-At^r}\mathrm{d}t \leq \int_M^\infty \left(\frac{t}{M}\right)^{kr-1} t^{r-kr}\mathrm{e}^{-At^r}\mathrm{d}t$$
$$= M^{1-kr}\int_M^\infty t^{r-1}\mathrm{e}^{-At^r}\mathrm{d}t = (Ar)^{-1}M^{1-kr}\mathrm{e}^{-AM^r}.$$

Hence,

$$J_0 \leq \max\{C_0; C_0'(Ar)^{-1}\}\mathrm{e}^{-AM^r}(M^{1-r} + M^{1-2r} + \cdots + M^{1-(k-1)r} + M^{1-kr})$$
$$\leq C_1 M^{1-r}\mathrm{e}^{-AM^r},$$

for some constant $C_1 \equiv C_1(A, r) > 0$. The lemma is proved. □

Lemma 6.3 *(Comte and Lacour (2013), Lemma 1, page 586) Let $A, r > 0$ and $a \in \mathbb{R}$. There exists a constant $C_2 > 0$ depending on A, a, r such that*

$$\int_0^M (1+t^2)^a \mathrm{e}^{At^r}\mathrm{d}t \leq C_2 M^{2a+1-r}\mathrm{e}^{AM^r}, \quad \textit{for any } M > 0.$$

Proof of Theorem 3.4 First, under the assumption (3.1), we apply the Lebesgue dominated convergence theorem to derive $\lim_{T\to\infty} I_T = 0$.
(a) For the conditions (3.3) and (3.4), we have

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^T \frac{(1+t)^{2(\beta_\varepsilon-\beta_\delta)}}{t^2}\mathrm{d}t.$$

Applying Lemma 6.1 yields

$$J_{n,T} \lesssim \begin{cases} n^{-1} & \text{if } \beta_\varepsilon < \beta_\delta + 1/2, \\ n^{-1}\ln T & \text{if } \beta_\varepsilon = \beta_\delta + 1/2, \\ n^{-1}T^{2(\beta_\varepsilon - \beta_\delta)-1} & \text{if } \beta_\varepsilon > \beta_\delta + 1/2. \end{cases}$$

Hence, if T satisfies (3.5), then $\lim_{n\to\infty} J_{n,T} = 0$ and $\lim_{n\to\infty} I_T = 0$. Combining these results with Proposition 3.3 gives desired result.
(b) For the conditions (3.6) and (3.7), we have

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^{T} \frac{\mathrm{e}^{2k_\varepsilon t^{\gamma_\varepsilon} - 2k_\delta t^{\gamma_\delta}}}{t^2}\mathrm{d}t.$$

- Consider the case $0 < \gamma_\varepsilon < \gamma_\delta$. Put $\psi(t) := k_\varepsilon t^{\gamma_\varepsilon} - k_\delta t^{\gamma_\delta}$, $t \in (0,\infty)$. It is easy to verify that the function ψ attains its maximum on $(0,\infty)$ at $t_* := \left(\frac{k_\varepsilon\gamma_\varepsilon}{k_\delta\gamma_\delta}\right)^{1/(\gamma_\delta-\gamma_\varepsilon)}$. Thus, $J_{n,T} \lesssim n^{-1}\mathrm{e}^{2\psi(t_*)}\int_{\omega_*}^{T} t^{-2}\mathrm{d}t \lesssim n^{-1}$.
- Consider the case $\gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0$. Then $J_{n,T} \lesssim n^{-1}\int_{\omega_*}^{T} t^{-2}\mathrm{e}^{2(k_\varepsilon - k_\delta)t^\gamma}\mathrm{d}t$. If $k_\varepsilon \le k_\delta$, then $J_{n,T} \lesssim n^{-1}\int_{\omega_*}^{T} t^{-2}\mathrm{d}t \lesssim n^{-1}$. If $k_\varepsilon > k_\delta$, then

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^{T}(1+t^2)^{-1}\mathrm{e}^{2(k_\varepsilon - k_\delta)t^\gamma}\mathrm{d}t \lesssim n^{-1}T^{-(1+\gamma)}\mathrm{e}^{2(k_\varepsilon - k_\delta)T^\gamma},$$

  where we have used Lemma 6.3 in the last estimate.
- Consider the case $\gamma_\varepsilon > \gamma_\delta$. Then

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^{T}(1+t^2)^{-1}\mathrm{e}^{2k_\varepsilon t^{\gamma_\varepsilon}}\mathrm{d}t \lesssim \frac{1}{n}T^{-(1+\gamma_\varepsilon)}\mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}},$$

where we have also applied Lemma 6.3. In summarize, we obtain that

$$J_{n,T} \lesssim \begin{cases} n^{-1} & \text{if } \gamma_\varepsilon < \gamma_\delta \text{ or } (\gamma_\varepsilon = \gamma_\delta, k_\varepsilon \le k_\delta), \\ n^{-1}T^{-(1+\gamma)}\mathrm{e}^{2(k_\varepsilon - k_\delta)T^\gamma} & \text{if } \gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0, k_\varepsilon > k_\delta, \\ n^{-1}T^{-(1+\gamma_\varepsilon)}\mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}} & \text{if } \gamma_\varepsilon > \gamma_\delta. \end{cases}$$

Hence, if T satisfies (3.8), then $\lim_{n\to\infty} J_{n,T} = 0$ and $\lim_{n\to\infty} I_T = 0$. Combining these results with Proposition 3.3 gives desired result.
(c) For the conditions (3.3) and (3.7), we have

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^{T}\frac{(1+t)^{-2\beta_\delta}\mathrm{e}^{2k_\varepsilon t^{\gamma_\varepsilon}}}{t^2}\mathrm{d}t \lesssim \frac{1}{n}\int_{\omega_*}^{T}(1+t^2)^{-(\beta_\delta+1)}\mathrm{e}^{2k_\varepsilon t^{\gamma_\varepsilon}}\mathrm{d}t.$$

Using Lemma 6.3 yields $J_{n,T} \lesssim n^{-1}T^{-(2\beta_\delta+\gamma_\varepsilon+1)}\mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}}$. Hence, if $\lim_{n\to\infty} T = \infty$ and $\lim_{n\to\infty} n^{-1}T^{-(2\beta_\delta+\gamma_\varepsilon+1)}\mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}} = 0$, then $\lim_{n\to\infty} J_{n,T} = 0$ and $\lim_{n\to\infty} I_T = 0$. Combining these results with Proposition 3.3 gives desired result.
(d) For the conditions (3.6) and (3.4), we have

$$J_{n,T} \lesssim \frac{1}{n}\int_{\omega_*}^{T} \frac{\mathrm{e}^{-2k_\delta t^{\gamma_\delta}}(1+t)^{2\beta_\varepsilon}}{t^2}\mathrm{d}t \lesssim \frac{1}{n},$$

which gives $\lim_{n\to\infty} J_{n,T} = 0$. Moreover, if $\lim_{n\to\infty} T = \infty$, then $\lim_{n\to\infty} I_T = 0$. Combining these results with Proposition 3.3 gives desired result. □

Proof of Theorem 3.5 Under the assumption (3.9), we obtain $J_{n,T} \lesssim n^{-1}$, which together with Proposition 3.3 to give

$$\sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 \lesssim I_T^2 + n^{-1}. \tag{6.5}$$

Moreover, as shown in the proof of Theorem 3.4, we have $\lim_{T\to\infty} I_T = 0$ under the assumption (3.1). Therefore, if T depends on n such that $\lim_{n\to\infty} T = \infty$, then the estimate (6.5) yields $\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \mathbb{E}|\widehat{F}_{W;T}(x) - F_W(x)|^2 = 0$. □

Proof of Theorem 4.2 (a) Suppose $f_X \in \mathcal{S}(\alpha_X, L_X)$. Under the condition (3.3), we have

$$I_T \lesssim \int_T^\infty \frac{|\varphi_X(t)|(1+t^2)^{\alpha_X/2}(1+t^2)^{-\alpha_X/2}(1+t)^{-\beta_\delta}}{t}\mathrm{d}t.$$

By the Cauchy-Schwarz inequality and the assumption $f_X \in \mathcal{S}(\alpha_X, L_X)$,

$$I_T^2 \lesssim \int_T^\infty |\varphi_X(t)|^2(1+t^2)^{\alpha_X}\mathrm{d}t \int_T^\infty \frac{(1+t^2)^{-\alpha_X}(1+t)^{-2\beta_\delta}}{t^2}\mathrm{d}t \lesssim T^{-(2\alpha_X+2\beta_\delta+1)}. \tag{6.6}$$

Next, by the condition (3.4), the Cauchy-Schwarz inequality and the assumptions $f_X \in \mathcal{S}(\alpha_X, L_X)$ and $\alpha_X + \beta_\varepsilon > 1/2$, there exists a positive constant M depending on $L_X, \alpha_X, \beta_\varepsilon$ such that $\|\varphi_X\varphi_\varepsilon\|_1 \le M$. From this, the estimate (6.6), the estimate of $J_{n,T}$ in the proof of Theorem 3.4(a) and Proposition 3.3, we deduce that

$$\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)] \lesssim \begin{cases} T^{-(2\alpha_X+2\beta_\delta+1)} + n^{-1} & \text{if } \beta_\varepsilon < \beta_\delta + 1/2, \\ T^{-(2\alpha_X+2\beta_\delta+1)} + n^{-1}\ln T & \text{if } \beta_\varepsilon = \beta_\delta + 1/2, \\ T^{-(2\alpha_X+2\beta_\delta+1)} + n^{-1}T^{2(\beta_\varepsilon-\beta_\delta)-1} & \text{if } \beta_\varepsilon > \beta_\delta + 1/2. \end{cases}$$

By choosing $T = T_1$, where T_1 is as in (4.1), we obtain the desired result of this part.
(b) Suppose $f_X \in \mathcal{S}(\alpha_X, L_X)$. It follows from this assumption, the Cauchy-Schwarz inequality, the condition (3.6) and Lemma 6.2 that

$$I_T^2 \leq \int_T^\infty \frac{|\varphi_X(t)|^2(1+t^2)^{\alpha_X}}{t^2(1+t^2)^{\alpha_X}}\mathrm{d}t \cdot \int_T^\infty |\varphi_\delta(t)|^2\mathrm{d}t$$
$$\lesssim T^{-(2\alpha_X+2)}\int_T^\infty \mathrm{e}^{-2k_\delta t^{\gamma_\delta}}\mathrm{d}t \lesssim T^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T^{\gamma_\delta}}. \tag{6.7}$$

From (6.7), the estimate of $J_{n,T}$ in the proof of Theorem 3.4(b) and Proposition 3.3, we deduce that

$$\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)]$$
$$\lesssim \begin{cases} T^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T^{\gamma_\delta}} + n^{-1} & \text{if } \gamma_\varepsilon < \gamma_\delta \text{ or } (\gamma_\varepsilon = \gamma_\delta, k_\varepsilon \leq k_\delta), \\ T^{-(2\alpha_X+\gamma+1)}\mathrm{e}^{-2k_\delta T^{\gamma}} + n^{-1}T^{-(1+\gamma)}\mathrm{e}^{2(k_\varepsilon-k_\delta)T^\gamma} & \text{if } \gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0, k_\varepsilon > k_\delta, \\ T^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T^{\gamma_\delta}} + n^{-1}T^{-(1+\gamma_\varepsilon)}\mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}} & \text{if } \gamma_\varepsilon > \gamma_\delta. \end{cases}$$

Now we distinguish the three following cases:

- Consider the case $\gamma_\varepsilon < \gamma_\delta$ or $(\gamma_\varepsilon = \gamma_\delta, k_\varepsilon \leq k_\delta)$. Then for T_2 as in (4.2) we have

$$T_2^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T_2^{\gamma_\delta}}$$
$$\lesssim (\ln n)^{-(2\alpha_X+\gamma_\delta+1)/\gamma_\delta}\exp\left\{-\left[\ln n - ((2\alpha_X+\gamma_\delta+1)/\gamma_\delta)\ln\left(\frac{\ln n}{2k_\delta}\right)\right]\right\} \lesssim n^{-1},$$

  so $\mathcal{R}[\widehat{F}_{W;T_2}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-1}$.
- Consider the case $\gamma_\varepsilon = \gamma_\delta \equiv \gamma > 0$ and $k_\varepsilon > k_\delta$. Then for T_3 as in (4.3) we have

$$T_3^{-(2\alpha_X+\gamma+1)}\mathrm{e}^{-2k_\delta T_3^{\gamma}} \lesssim (\ln n)^{-(2\alpha_X+\gamma+1)/\gamma}\exp\left\{-2k_\delta\left(\frac{\ln n - (2\alpha_X/\gamma)\ln\left(\frac{\ln n}{2k_\varepsilon}\right)}{2k_\varepsilon}\right)\right\}$$
$$\lesssim n^{-k_\delta/k_\varepsilon}(\ln n)^{-U/\gamma},$$

  with $U := 2\alpha_X(1-k_\delta/k_\varepsilon)+\gamma+1 > 0$. Moreover,

$$n^{-1}T_3^{-(1+\gamma)}\mathrm{e}^{2(k_\varepsilon-k_\delta)T_3^\gamma}$$
$$\lesssim n^{-1}(\ln n)^{-(1+\gamma)/\gamma}\exp\left\{2(k_\varepsilon-k_\delta)\left(\frac{\ln n - (2\alpha_X/\gamma)\ln\left(\frac{\ln n}{2k_\varepsilon}\right)}{2k_\varepsilon}\right)\right\}$$
$$\lesssim n^{-k_\delta/k_\varepsilon}(\ln n)^{-U/\gamma}.$$

  Therefore, we obtain that $\mathcal{R}[\widehat{F}_{W;T_3}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-k_\delta/k_\varepsilon}(\ln n)^{-U/\gamma}$.
- Consider the case $\gamma_\varepsilon > \gamma_\delta$. First, the existence and uniqueness of the positive solution T_4 in n for the equation (4.4) are verified easily by considering the function

$$\Phi(t) := 2\alpha_X \ln t + 2k_\delta t^{\gamma_\delta} + 2k_\varepsilon t^{\gamma_\varepsilon}, \quad t > 0.$$

Also we have $\lim_{n\to\infty} T_4 = \infty$. Since $\gamma_\varepsilon > \gamma_\delta$, we deduce $\lim_{t\to\infty} \Phi(t)/(2k_\varepsilon t^{\gamma_\varepsilon}) = 1$, so $\lim_{n\to\infty} \Phi(T_4)/(2k_\varepsilon T_4^{\gamma_\varepsilon}) = 1$. On the other hand, $\Phi(T_4) = \ln n + \ln(k_\delta\gamma_\delta/(k_\varepsilon\gamma_\varepsilon))$. Thus,

$$\lim_{n\to\infty} \frac{\ln n + \ln\left(\frac{k_\delta\gamma_\delta}{k_\varepsilon\gamma_\varepsilon}\right)}{2k_\varepsilon T_4^{\gamma_\varepsilon}} = 1,$$

which gives

$$T_4 = \left(\frac{\ln n}{2k_\varepsilon}\right)^{1/\gamma_\varepsilon} [1 + o(1)].$$

Note that using (4.4) we have

$$\lim_{n\to\infty} \frac{n^{-1} T_4^{-(1+\gamma_\varepsilon)} \mathrm{e}^{2k_\varepsilon T_4^{\gamma_\varepsilon}}}{T_4^{-(2\alpha_X+\gamma_\delta+1)} \mathrm{e}^{-2k_\delta T_4^{\gamma_\delta}}} = \frac{k_\delta\gamma_\delta}{k_\varepsilon\gamma_\varepsilon} \lim_{n\to\infty} T_4^{\gamma_\delta-\gamma_\varepsilon} = 0,$$

so

$$n^{-1} T_4^{-(1+\gamma_\varepsilon)} \mathrm{e}^{2k_\varepsilon T_4^{\gamma_\varepsilon}} \lesssim T_4^{-(2\alpha_X+\gamma_\delta+1)} \mathrm{e}^{-2k_\delta T_4^{\gamma_\delta}}.$$

Thus we obtain

$$\begin{aligned}\mathcal{R}[\widehat{F}_{W;T_4}; \mathcal{S}(\alpha_X, L_X)] &\lesssim T_4^{-(2\alpha_X+\gamma_\delta+1)} \mathrm{e}^{-2k_\delta T_4^{\gamma_\delta}} \\ &\lesssim (\ln n)^{-(2\alpha_X+\gamma_\delta+1)/\gamma_\varepsilon} \mathrm{e}^{-2k_\delta\left(\frac{\ln n}{2k_\varepsilon}\right)^{\gamma_\delta/\gamma_\varepsilon}[1+o(1)]}.\end{aligned}$$

(c) Suppose $f_X \in \mathcal{S}(\alpha_X, L_X)$. Under this assumption and the condition (3.3), we have as in the proof of the part (a) that $I_T^2 \lesssim T^{-(2\alpha_X+2\beta_\delta+1)}$. Moreover, the conditions (3.3), (3.7) and Lemma 6.3 lead to $J_{n,T} \lesssim n^{-1} T^{-(2\beta_\delta+1+\gamma_\varepsilon)} \mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}}$. Combining Proposition 3.3 with the estimates of I_T^2 and $J_{n,T}$, we get that

$$\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)] \lesssim T^{-(2\alpha_X+2\beta_\delta+1)} + n^{-1} T^{-(2\beta_\delta+1+\gamma_\varepsilon)} \mathrm{e}^{2k_\varepsilon T^{\gamma_\varepsilon}}.$$

For T_5 as in (4.5), we have

$$\begin{aligned}&T_5^{-(2\alpha_X+2\beta_\delta+1)} \lesssim (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}, \\ &n^{-1} T_5^{-(2\beta_\delta+1+\gamma_\varepsilon)} \mathrm{e}^{2k_\varepsilon T_5^{\gamma_\varepsilon}} \\ &\lesssim n^{-1} (\ln n)^{-(2\beta_\delta+1+\gamma_\varepsilon)/\gamma_\varepsilon} \exp\left\{\ln n - ((2\alpha_X - \gamma_\varepsilon)/\gamma_\varepsilon) \ln\left(\frac{\ln n}{2k_\varepsilon}\right)\right\} \\ &\lesssim (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}.\end{aligned}$$

Thus, $\mathcal{R}[\widehat{F}_{W;T_5}; \mathcal{S}(\alpha_X, L_X)] \lesssim (\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon}$.

(d) Suppose $f_X \in \mathcal{S}(\alpha_X, L_X)$. Under this assumption and the condition (3.6), we have as in the proof of the part (b) that $I_T^2 \lesssim T^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T^{\gamma_\delta}}$. The conditions (3.4), (3.6) lead to $J_{n,T} \lesssim n^{-1}$. Besides, for the assumption $\alpha_X + \beta_\varepsilon > 1/2$, there exists a constant $M > 0$ such that $\|\varphi_X \varphi_\varepsilon\|_1 \le M$, for any $f_X \in \mathcal{S}(\alpha_X, L_X)$. Combining the above estimates with Proposition 3.3, we get that

$$\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)] \lesssim T^{-(2\alpha_X+\gamma_\delta+1)}\mathrm{e}^{-2k_\delta T^{\gamma_\delta}} + n^{-1},$$

By the same arguments as in the proof of the part (b) with $\gamma_\varepsilon < \gamma_\delta$ or ($\gamma_\varepsilon = \gamma_\delta$, $k_\varepsilon \le k_\delta$), we obtain the desired result. □

Proof of Theorem 4.3 Suppose $f_X \in \mathcal{S}(\alpha_X, L_X)$. Then it is not hard to verify that $I_T^2 \lesssim T^{-(2\alpha_X+1)}$. Moreover, under the condition (3.9), we have $J_{n,T} \lesssim n^{-1}$. Combining these with Proposition 3.3, we get that $\mathcal{R}[\widehat{F}_{W;T}; \mathcal{S}(\alpha_X, L_X)] \lesssim T^{-(2\alpha_X+1)} + n^{-1}$. Therefore, $\mathcal{R}[\widehat{F}_{W;T_6}; \mathcal{S}(\alpha_X, L_X)] \lesssim n^{-1}$, where T_6 is as in (4.8). □

Proof of Theorem 4.4 Let $\rho \equiv \rho_n > 0$, $N \equiv N_n > 0$ be parameters depending on n such that $\lim_{n\to\infty} \rho = 0$, $\lim_{n\to\infty} N = \infty$ and $\rho^2 N^{2\alpha_X-1} \lesssim 1$. These parameters will be selected later. Let $H : \mathbb{R} \to \mathbb{R}$ be a function satisfying the properties as follows: H is infinite continuously differentiable on $\mathbb{R}$, $H(-t) = -H(t)$ for any $t \in \mathbb{R}$, $H(t) \ge 0$ for all $t \ge 0$, $\mathrm{supp}(H) = [-2, -1] \cup [1, 2]$, and $H^{(k)}(1) = H^{(k)}(2) = 0$ for all $k = 0, 1, 2, \ldots$ Define J as the inverse Fourier transform of $\mathrm{i}H\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\{\varphi_\delta(N\cdot)\}}$, i.e.

$$J(x) := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{i}H(t)\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\{\varphi_\delta(Nt)\}}\mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t = \frac{1}{\pi}\int_1^2 H(t)\sin(\mathrm{Arg}\{\varphi_\delta(Nt)\} + tx)\mathrm{d}t.$$

Obviously, the function J is real-valued.

Next we set $f_{X,1}(x) := 1/[\pi(1+x^2)]$, $f_{X,2}(x) := f_{X,1}(x) + \rho J(Nx)$, $x \in \mathbb{R}$. We now show that $f_{X,1}, f_{X,2} \in \mathcal{S}(\alpha_X, L_X)$ for large $L_X > 0$ and for large n. First, observe that $f_{X,1}$ is a density with $\varphi_{f_{X,1}}(t) = \mathrm{e}^{-|t|}$, $t \in \mathbb{R}$. For any $|x| \le 1$, we have $f_{X,1}(x) \ge 1/(2\pi)$ and $|J(Nx)| \le \frac{1}{\pi}\int_1^2 |H(t)|\mathrm{d}t$, so $f_{X,2}(x) \ge 1/(2\pi) - \frac{\rho}{\pi}\int_1^2 |H(t)|\mathrm{d}t > 0$, for n large enough. For $|x| > 1$, using the Fourier inversion formula, we have

$$|J(Nx)| = \frac{1}{2\pi N}\left|\int_{-\infty}^{\infty} \varphi_J(t/N)\mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t\right| = \frac{1}{2\pi N}\left|\int_{-\infty}^{\infty} \widetilde{H}_N(t)\mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t\right|,$$

with $\widetilde{H}_N(t) := H(t/N)\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\{\varphi_\delta(t)\}} = H(t/N)\mathrm{e}^{-\mathrm{i}a(t)}$. Note that

$$\begin{aligned}\widetilde{H}_N''(t) = &\frac{1}{N^2}H''(t/N)\mathrm{e}^{-\mathrm{i}a(t)} - \frac{2\mathrm{i}}{N}H'(t/N)\mathrm{e}^{-\mathrm{i}a(t)}a'(t)\\ &- H(t/N)\mathrm{e}^{-\mathrm{i}a(t)}[a'(t)]^2 - \mathrm{i}H(t/N)\mathrm{e}^{-\mathrm{i}a(t)}a''(t),\end{aligned}$$

which together with the assumption on $a(t)$ gives

$$|\widetilde{H}_N''(t)| \lesssim \frac{1}{N^2}|H''(t/N)| + \frac{1}{N}|H'(t/N)| + |H(t/N)|.$$

By integration by parts,

$$|J(Nx)| = \frac{1}{2\pi Nx^2}\left|\int_{-\infty}^{\infty} \widetilde{H}_N''(t)\mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t\right|$$
$$\lesssim \frac{1}{\pi N(1+x^2)}\int_{-\infty}^{\infty}\left(\frac{1}{N^2}|H''(t/N)| + \frac{1}{N}|H'(t/N)| + |H(t/N)|\right)\mathrm{d}t$$
$$\lesssim \frac{1}{\pi(1+x^2)}.$$

Therefore, $f_{X,2}(x) = f_{X,1}(x) + \rho J(Nx) \gtrsim \frac{1}{\pi(1+x^2)}(1-\rho) > 0$ for large n. Hence, we have verified the non-negativity of $f_{X,2}$. Furthermore, we have $\int_{-\infty}^{\infty} f_{X,2}(x)\mathrm{d}x = \varphi_{f_{X,2}}(0) = \varphi_{f_{X,1}}(0) = 1$. In summary, $f_{X,2}$ is also a density for large n.
Using the equality $\varphi_{J(N\cdot)}(t) = N^{-1}\varphi_J(t/N) = \mathrm{i}N^{-1}H(t/N)\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\varphi_\delta(t)}$ and the inequality $|z_1+z_2|^2 \le 2|z_1|^2 + 2|z_2|^2$ for all $z_1, z_2 \in \mathbb{C}$, we obtain

$$\int_{-\infty}^{\infty} |\varphi_{f_{X,2}}(t)|^2(1+t^2)^{\alpha_X}\mathrm{d}t \le 2\int_{-\infty}^{\infty}\mathrm{e}^{-2|t|}(1+t^2)^{\alpha_X}\mathrm{d}t + \frac{2\rho^2}{N}\int_1^2 |H(t)|^2(1+N^2t^2)^{\alpha_X}\mathrm{d}t$$
$$\lesssim 1 + \rho^2 N^{2\alpha_X - 1}.$$

From the latter estimate and the condition $\rho^2 N^{2\alpha_X-1} \lesssim 1$, there exists a constant $L_X > 0$ large enough, independent of n, so that $\int_{-\infty}^{\infty} |\varphi_{f_{X,2}}(t)|^2(1+t^2)^{\alpha_X}\mathrm{d}t \le L_X$. Of course, we also have $\int_{-\infty}^{\infty} |\varphi_{f_{X,1}}(t)|^2(1+t^2)^{\alpha_X}\mathrm{d}t \le L_X$.
Therefore, we conclude that $f_{X,1}, f_{X,2} \in \mathcal{S}(\alpha_X, L_X)$ for large n.
For brevity, we write $\widehat{F}_W(x)$ to represent $\widehat{F}_W(x; Y_1, \ldots, Y_n)$. For the densities $f_{X,1}$, $f_{X,2}$, we define $f_{W,\ell} := f_{X,\ell} * f_\delta$, $F_{W,\ell}(x) := \int_{-\infty}^{x} f_{W,\ell}(u)\mathrm{d}u$, with $\ell \in \{1,2\}$. Then

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \ge \frac{1}{2}\sum_{\ell=1}^{2}\mathbb{E}_{f_{X,\ell}*f_\varepsilon}|\widehat{F}_W(0) - F_{W,\ell}(0)|^2$$
$$\ge \frac{1}{4}|F_{W,2}(0) - F_{W,1}(0)|^2 \Delta, \tag{6.8}$$

with

$$\Delta := \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\min\left\{\prod_{j=1}^{n}(f_{X,1}*f_\varepsilon)(y_j); \prod_{j=1}^{n}(f_{X,2}*f_\varepsilon)(y_j)\right\}\mathrm{d}y_1\cdots\mathrm{d}y_n.$$

By the definitions of $F_{W,1}$, $F_{W,2}$ and the Fourier inversion formula, we have

$$F_{W,2}(0) - F_{W,1}(0) = \rho\int_{-\infty}^{0}(J(N\cdot)*f_\delta)(x)\mathrm{d}x$$
$$= \frac{\rho}{2\pi}\lim_{a\to-\infty}\int_a^0\left(\int_{-\infty}^{\infty}\varphi_{J(N\cdot)}(t)\varphi_\delta(t)\mathrm{e}^{-\mathrm{i}tx}\mathrm{d}t\right)\mathrm{d}x.$$

Using the Fubini theorem and the relation $\varphi_{J(N\cdot)}(t) = \mathrm{i}N^{-1}H(t/N)\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\varphi_\delta(t)}$ gives

$$\begin{aligned} F_{W,2}(0) - F_{W,1}(0) &= \frac{\rho}{2\pi\mathrm{i}} \lim_{a\to-\infty} \int_{-\infty}^{\infty} \varphi_{J(N\cdot)}(t)\varphi_\delta(t)\mathrm{e}^{-\mathrm{i}\mathrm{Arg}\varphi_\delta(t)} \frac{\mathrm{e}^{-\mathrm{i}ta}-1}{t}\mathrm{d}t \\ &= \frac{\rho}{2\pi N} \lim_{a\to-\infty} \int_{-\infty}^{\infty} H(t/N)|\varphi_\delta(t)| \frac{\mathrm{e}^{-\mathrm{i}ta}-1}{t}\mathrm{d}t. \end{aligned}$$

Note that the function $t \mapsto \frac{H(t/N)|\varphi_\delta(t)|}{t}$ is in $L^1(\mathbb{R})$. Thus, we apply the Riemann-Lebesgue lemma (see, e.g. Bochner and Chandrasekharan (1949), page 3) to obtain

$$F_{W,2}(0) - F_{W,1}(0) = -\frac{\rho}{2\pi N} \int_{-\infty}^{\infty} \frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t = -\frac{\rho}{\pi N} \int_1^2 \frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t. \tag{6.9}$$

It follows from (6.8) and (6.9) that

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \geq \frac{\rho^2}{4\pi^2 N^2} \left(\int_1^2 \frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t \right)^2 \Delta. \tag{6.10}$$

For the quantity Δ, we have by the LeCam inequality (see, e.g. Devroye and Lugosi (2001), page 43) that

$$\begin{aligned} \Delta &\geq \frac{1}{2} \left(\int_{-\infty}^{\infty} \sqrt{(f_{X,1} * f_\varepsilon)(y)} \sqrt{(f_{X,2} * f_\varepsilon)(y)} \mathrm{d}y \right)^{2n} \\ &\geq \frac{1}{2} \left(1 - \frac{1}{2} \int_{-\infty}^{\infty} \frac{\left[(f_{X,1} * f_\varepsilon)(y) - (f_{X,2} * f_\varepsilon)(y)\right]^2}{(f_{X,1} * f_\varepsilon)(y)} \mathrm{d}y \right)^{2n} = \frac{1}{2} \left(1 - \frac{1}{2} Q \right)^{2n} \end{aligned}$$

with

$$Q := \rho^2 \int_{-\infty}^{\infty} \frac{[(J(N\cdot) * f_\varepsilon)(y)]^2}{(f_{X,1} * f_\varepsilon)(y)} \mathrm{d}y.$$

Combining this estimate with (6.10) gives

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \geq \frac{\rho^2}{8\pi^2 N^2} \left(\int_1^2 \frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t \right)^2 \left(1 - \frac{1}{2} Q \right)^{2n}. \tag{6.11}$$

We now bound Q. Fix a constant $A > 0$ such that $\int_{-A}^{A} f_\varepsilon(u)\mathrm{d}u \geq 1/2$. Then

$$(f_{X,1} * f_\varepsilon)(y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f_\varepsilon(u)}{1+(y-u)^2}\mathrm{d}u \geq \frac{1}{\pi} \int_{-A}^{A} \frac{f_\varepsilon(u)}{1+2y^2+2u^2}\mathrm{d}u \geq \frac{1}{4\pi(1+A^2)(1+y^2)},$$

which gives

$$Q \lesssim \rho^2 \int_{-\infty}^{\infty} (1+y^2)[J(N\cdot) * f_\varepsilon)(y)]^2 \mathrm{d}y = \rho^2(Q_1 + Q_2), \tag{6.12}$$

where

$$Q_1 := \int_{-\infty}^{\infty} |(J(N\cdot) * f_\varepsilon)(y)|^2 \mathrm{d}y, \quad Q_2 := \int_{-\infty}^{\infty} |y(J(N\cdot) * f_\varepsilon)(y)|^2 \mathrm{d}y.$$

Applying the Parseval identity yields

$$\begin{aligned}
Q_1 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\varphi_{J(N\cdot)}(t)|^2 |\varphi_\varepsilon(t)|^2 \mathrm{d}t = \frac{1}{2\pi N^2} \int_{-\infty}^{\infty} |\varphi_J(t/N)|^2 |\varphi_\varepsilon(t)|^2 \mathrm{d}t \\
&= \frac{1}{2\pi N} \int_{-\infty}^{\infty} |\varphi_J(t)|^2 |\varphi_\varepsilon(Nt)|^2 \mathrm{d}t = \frac{1}{\pi N} \int_1^2 |H(t)|^2 |\varphi_\varepsilon(Nt)|^2 \mathrm{d}t, \\
Q_2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\varphi'_{J(N\cdot)*f_\varepsilon}(t)|^2 \mathrm{d}t = \frac{1}{2\pi N^2} \int_{-\infty}^{\infty} |[H(t/N)\mathrm{e}^{-\mathrm{i}a(t)}\varphi_\varepsilon(t)]'|^2 \mathrm{d}t \\
&= \frac{1}{2\pi N^2} \int_{-\infty}^{\infty} \left| \frac{1}{N} H'(t/N)\mathrm{e}^{-\mathrm{i}a(t)}\varphi_\varepsilon(t) - \mathrm{i}H(t/N)\varphi_\varepsilon(t)\mathrm{e}^{-\mathrm{i}a(t)}a'(t) \right. \\
&\quad \left. + H(t/N)\mathrm{e}^{-\mathrm{i}a(t)}\varphi'_\varepsilon(t) \right|^2 \mathrm{d}t \\
&\lesssim \frac{1}{N} \int_1^2 (|H(t)|^2 + |H'(t)|^2)(|\varphi_\varepsilon(Nt)|^2 + |\varphi'_\varepsilon(Nt)|^2) \mathrm{d}t.
\end{aligned}$$

It follows from (6.12) and the estimates of Q_1 and Q_2 that

$$Q \lesssim \frac{\rho^2}{N} \int_1^2 (|H(t)|^2 + |H'(t)|^2)(|\varphi_\varepsilon(Nt)|^2 + |\varphi'_\varepsilon(Nt)|^2) \mathrm{d}t. \tag{6.13}$$

(a) Using (4.9) and the right inequality in (3.4), we obtain from (6.13) that

$$Q \lesssim \frac{\rho^2}{N} \int_1^2 (|H(t)|^2 + |H'(t)|^2)(1+Nt)^{-2\beta_\varepsilon} \mathrm{d}t \lesssim \rho^2 N^{-(2\beta_\varepsilon+1)}.$$

Letting $N = n^{1/(2\alpha_X+2\beta_\varepsilon)}$ and $\rho = N^{-(2\alpha_X-1)/2}$ implies $Q \lesssim 1/n$, which together with (6.11) yields

$$\mathcal{R}[\widehat{F}_W; \mathcal{S}(\alpha_X, L_X)] \gtrsim N^{-(2\alpha_X+1)} \left(\int_1^2 \frac{H(t)|\varphi_\delta(Nt)|}{t} \mathrm{d}t \right)^2.$$

Afterwards, using the condition (4.10) yields

$$\mathcal{R}[\widehat{F}_W;\mathcal{S}(\alpha_X,L_X)]\gtrsim N^{-(2\alpha_X+1)}\left(\int_1^2\frac{H(t)(1+Nt)^{-\beta_\delta}}{t}\mathrm{d}t\right)^2\gtrsim N^{-(2\alpha_X+2\beta_\delta+1)}$$
$$=n^{-(2\alpha_X+2\beta_\delta+1)/(2\alpha_X+2\beta_\varepsilon)}.$$

(b) Using (4.11) and the right inequality in (3.7), we derive from (6.13) that

$$Q\lesssim\frac{\rho^2}{N}\int_1^2(|H(t)|^2+|H'(t)|^2)\mathrm{e}^{-2k_\varepsilon(Nt)^{\gamma_\varepsilon}}\mathrm{d}t\lesssim\frac{\rho^2}{N}\mathrm{e}^{-2k_\varepsilon N^{\gamma_\varepsilon}}.$$

Letting $N=((\ln n)/(2k_\varepsilon))^{1/\gamma_\varepsilon}$ and $\rho=N^{-(2\alpha_X-1)/2}$ gives $Q\lesssim 1/n$, which combines with (6.11) to give

$$\mathcal{R}[\widehat{F}_W;\mathcal{S}(\alpha_X,L_X)]\gtrsim N^{-(2\alpha_X+1)}\left(\int_1^2\frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t\right)^2.$$

Afterwards, using the condition (4.12) yields

$$\mathcal{R}[\widehat{F}_W;\mathcal{S}(\alpha_X,L_X)]\gtrsim N^{-(2\alpha_X+1)}\left(\int_1^2\frac{H(t)\mathrm{e}^{-k_\delta(Nt)^{\gamma_\delta}}}{t}\mathrm{d}t\right)^2\gtrsim N^{-(2\alpha_X+1)}\mathrm{e}^{-2k_\delta(2N)^{\gamma_\delta}}$$
$$\gtrsim(\ln n)^{-(2\alpha_X+1)/\gamma_\varepsilon}\mathrm{e}^{-2^{1+\gamma_\delta}k_\delta\left(\frac{\ln n}{2k_\varepsilon}\right)^{\gamma_\delta/\gamma_\varepsilon}}.$$

(c) From (6.13), (4.11) and the right inequality in (3.7), we derive $Q\lesssim\rho^2N^{-1}\mathrm{e}^{-2k_\varepsilon N^{\gamma_\varepsilon}}$. Letting $N=(d\ln n)^{1/\gamma_\varepsilon}$ (for $d>1/(2k_\varepsilon)$) and $\rho=N^{(1-2\alpha_X)/2}$ gives $Q\lesssim 1/n$, which together with (6.11) implies

$$\mathcal{R}[\widehat{F}_W;\mathcal{S}(\alpha_X,L_X)]\gtrsim N^{-(2\alpha_X+1)}\left(\int_1^2\frac{H(t)|\varphi_\delta(Nt)|}{t}\mathrm{d}t\right)^2\gtrsim N^{-(2\alpha_X+2\beta_\delta+1)}$$
$$\gtrsim(\ln n)^{-(2\alpha_X+2\beta_\delta+1)/\gamma_\varepsilon},$$

where we have used the condition (4.10). □

Acknowledgements We would like to thank the reviewer and the editors for fruitful comments and suggestions which help to significantly improve the paper.

Declaration

Conflict of interest The authors state that there is no conflict of interest.

References

Berkson J (1950) Are there two regressions? J Am Stat Assoc 45(250):164–180
Bochner S, Chandrasekharan K (1949) Fourier transforms, 19th edn. Princeton University Press, New Jersey

Carroll RJ, Hall P (1988) Optimal rates of convergence for deconvolving a density. J Am Stat Assoc 83(404):1184–1186
Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM (2006) Measurement error in nonlinear models: A modern perspective. CRC Press, Boca Raton
Carroll RJ, Delaigle A, Hall P (2009) Nonparametric prediction in measurement error models. J Am Stat Assoc 104(487):993–1003
Comte F, Lacour C (2013) Anisotropic adaptive kernel deconvolution. Annales de l-IHP Probabilités et statistiques 49(2):569–609
Comte F, Rozenholc Y, Taupin ML (2006) Penalized contrast estimator for adaptive density deconvolution. Can J Stat 34(3):431–452
Cordy CB, Thomas DR (1997) Deconvolution of a distribution function. J Am Stat Assoc 92(440):1459–1465
Dattner I, Reiser B (2013) Estimation of distribution functions in measurement error models. J Stat Plan Inference 143(3):479–493
Dattner I, Goldenshluger A, Juditsky A (2011) On deconvolution of distribution functions. Annals Stat 39(5):2477–2501
Delaigle A (2007) Nonparametric density estimation from data with a mixture of Berkson and classical errors. Can J Stat 35(1):89–104
Delaigle A (2014) Nonparametric kernel methods with errors-in-variables: constructing estimators, computing them, and avoiding common mistakes. Aust New Zealand J Stat 56(2):105–124
Delaigle A, Hall P (2006) On optimal kernel choice for deconvolution. Stat Probab Lett 76(15):1594–1602
Delaigle A, Meister A (2011) Rate-optimal nonparametric estimation in classical and berkson errors-in-variables problems. J Stat Plan Inference 141(1):102–114
Delaigle A, Hall P, Qiu P (2006) Nonparametric methods for solving the Berkson errors-in-variables problem. J Royal Stat Soc: Series B (Stat Methodol) 68(2):201–220
Devroye L, Lugosi G (2001) Combinatorial methods in density estimation. Springer, New York
Fan J (1991) On the optimal rates of convergence for nonparametric deconvolution problems. Annals Stat 19(3):1257–1272
Fuller WA (1987) Measurement error models. Wiley, New York
Gaffey WR (1959) A consistent estimator of a component of a convolution. Annals Math Stat 30(1):198–205
Gil-Pelaez J (1951) Note on the inversion theorem. Biometrika 38(3–4):481–482
Hall P, Meister A (2007) A ridge-parameter approach to deconvolution. Annals Stat 35(4):1535–1558
Kawata T (1972) Fourier analysis in probability theory. Academic Press, New York
Meister A (2009) Deconvolution problems in nonparametric statistics. Springer, Berlin
Pensky M, Vidakovic B (1999) Adaptive wavelet estimator for nonparametric density deconvolution. Annals Stat 27(6):2033–2053
Rimal R, Pensky M (2019) Density deconvolution with small Berkson errors. Math Methods Stat 28(3):208–227
Stefanski LA, Carroll RJ (1990) Deconvolving kernel density estimators. Stat 21(2):169–184
Trong DD, Phuong CX (2019) Deconvolution of a cumulative distribution function with some non-standard noise densities. Vietnam J Math 47(2):327–353
Wand M (1998) Finite sample performance of deconvolving density estimators. Stat Probab Lett 37(2):131–139