ON LINEARIZATION OF NONPARAMETRIC DECONVOLUTION ESTIMATORS FOR REPEATED MEASUREMENTS MODEL

DAISUKE KURISU AND TAISUKE OTSU

Abstract. By utilizing intermediate Gaussian approximations, this paper establishes asymptotic linear representations of nonparametric deconvolution estimators for the classical measurement error model with repeated measurements. Our result is applied to derive confidence bands for the density and distribution functions of the error-free variable of interest and to establish faster convergence rates of the estimators than the ones obtained in the existing literature. Due to slower decay rates of the linearization errors, however, our bootstrap counterparts for confidence bands need to be constructed by subsamples.

1. Introduction

This paper establishes asymptotic linear representations of nonparametric deconvolution estimators for the classical measurement error model, where repeated noisy measurements on the error-free variable of interest are available. For this problem, a seminal work by Li and Vuong (1998, hereafter LV) developed a nonparametric estimator for the densities of the error-free variable of interest and the measurement errors, which are identified via Kotlarski's (1967) identity. A large body of the existing literature on nonparametric deconvolution methods (see Meister, 2009, for a review) requires perfect knowledge of the measurement error distribution, which is hardly available in practice. In contrast, the LV estimator circumvents such a requirement by utilizing information from repeated measurements. Another attractive feature of the LV estimator is that it does not require prior information on the shape of the measurement error density, such as symmetry (as in Delaigle, Hall and Meister, 2008) or auxiliary data drawn from the measurement error densities (as in Neumann, 2007).

Given this background, there is growing interest in the LV estimator and related methods. LV derived the uniform convergence rates for their estimators under bounded support conditions. Comte and Kappus (2015) studied a regularized version the LV estimator and established the L_2-convergence rates under weaker assumptions than the ones in LV, which allow unbounded data support. Bonhomme and Robin (2010) considered a general latent multi-factor model, which covers the repeated measurements model as a special case, and established the uniform convergence rate without assuming bounded support. Kurisu and Otsu (2021) derived faster uniform convergence rates than LV and Bonhomme and Robin (2010) under even weaker assumptions by utilizing a maximal inequality for the multivariate normalized empirical characteristic function process.

It should be noted that the existing literature mostly focuses on characterizing convergence rates of the LV-type estimators, and their further theoretical properties are largely unknown. For example, it is not clear how to construct confidence bands for the densities of the error-free variable of interest and the measurement errors based on those estimators. Also optimal convergence rates, adaptive bandwidth selection methods, and limit theorems for functionals of the LV-type estimators are open questions in this setup. A recent paper by Kato, Sasaki and Ura (2021) developed confidence bands for the densities

Kurisu acknowledges financial support by Grant-in-Aid for Early-Career Scientists (20K13468) from the JSPS.

in the repeated measurements model by exploring the implied moment conditions approximated by Hermite polynomial sieves. In contrast, this paper constructs confidence bands based on the LV estimator.

As an initial step toward filling this gap, this paper establishes (uniform) asymptotic linear approximations for the LV estimators for the densities of the error-free variable of interest and the measurement errors. Due to complicated structures of the LV estimators, there are at least two reasons that make our asymptotic analysis non-trivial. First, it is involving to characterize the empirical processes for the dominant terms of the characteristic function estimators by using intermediate Gaussian approximations (e.g., Chernozhukov, Chetverikov and Kato, 2016). Second, we need to apply intermediate Gaussian approximations again for the (regularized) Fourier inversions to establish the asymptotic linear forms for the resulting LV density estimators. To the best of our knowledge, our applications of intermediate Gaussian approximations seem new in the studies of the LV-type estimators.

Our asymptotic linear approximations immediately yield several important implications. First, our intermediate Gaussian approximation approach allows to derive faster uniform convergence rates than the existing ones, such as LV, Bonhomme and Robin (2010), and Kurisu and Otsu (2021). Second, by perturbing the obtained linear forms, we can develop Gaussian multiplier bootstrap confidence bands for the density and cumulative distribution functions of the error-free variable and measurement errors. Although this is the first paper establishing the confidence bands based on the LV estimators, a drawback of the proposed confidence bands is that our bootstrap counterparts need to be constructed based on subsamples because of the slower decay rates of the linearization errors. Finally, our intermediate Gaussian approximation approach provides bootstrap pointwise confidence intervals without knowing the limiting distributions of the LV-type estimators. Also we conjecture our linear approximations can serve as building blocks for further theoretical analyses on the LV-type estimators, such as optimal convergence rates.

In the context of constructing confidence bands for nonparametric measurement error problems, Bissantz *et al.* (2007) considered the case where the measurement error density is known to researchers, and developed confidence bands by showing that the supremum deviation of the deconvolution kernel density estimator converges in distribution to a Gumbel distribution. Kato and Sasaki (2018, 2019) studied the case where the measurement error density is unknown but auxiliary observations from the measurement error density are available so that the deconvolution kernel can be constructed by plugging in the empirical characteristic function of the measurement error distribution based on the auxiliary data. In this setup, Kato and Sasaki (2018, 2019) considered nonparametric density and regression estimation problems, respectively, established Gaussian intermediate approximations to their estimators with suitable normalizations, and proposed multiplier bootstrap confidence bands. Their setup covers repeated measurement models when the measurement error density is typically symmetric around zero (see also Delaigle, Hall and Meister, 2008). The major difference of the present work with Kato and Sasaki (2018, 2019) is that we do not impose such a shape constraint on the measurement error density and construct confidence bands based on the LV estimator, which takes a substantially different form from Kato and Sasaki's (2018) estimator so that theoretical developments are very different from theirs. See Remark 2 below for a further detail.

We also note that several empirical studies indicate asymmetric shapes of measurement error densities, which motivate the LV-type estimation methods. For example, Li, Perrigne and Vuong (2000) and Krasnokutskaya (2011) applied the LV estimator to auction data and reported asymmetric density estimates for the measurement errors. Bonhomme and Robin (2010) applied the LV-type estimator for multi-factor models to study earning dynamics in the US. Although their estimated densities (for the objects corresponding to the measurement error densities) are overall symmetric, such information on the shape of densities is typically unavaiable a priori. In this case, the LV-type estimator would be useful to motivate the symmetry or other shape restrictions on the measurement error densities and to proceed to more efficient estimation methods, such as Kato and Sasaki (2018).

This paper is organized as follows. Section 2 presents our main results, asymptotic linear approximations for the LV estimators. In Section 3, we discuss applications of our main results for refined convergence rates of the LV estimators (Section 3.1), confidence bands of the density functions of the error-free variable of interest and the measurement errors (Section 3.2), and confidence bands of the cumulative distribution functions of the error-free variable of interest and the measurement errors (Section 3.3).

Notation. Hereafter, we use the following notation. For any $a, b \in \mathbb{R}$, let $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. For any positive sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n \lesssim b_n$ if there is a positive constant C independent of n such that $a_n \le C b_n$ for all n, $a_n \sim b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$, and $a_n \ll b_n$ if $a_n/b_n \to 0$ as $n \to \infty$. For random variables X and Y, we write $X \stackrel{d}{=} Y$ if they have the same distribution.

2. Linearization lemmas

2.1. **Setup and estimators.** We first introduce our basic setup and define the density deconvolution estimators. Consider a bivariate i.i.d. sample $\{Y_{1,j}, Y_{2,j}\}_{j=1}^n$ of (Y_1, Y_2) generated by

$$\begin{aligned} Y_1 &= X + \epsilon_1, \qquad (1) \\ Y_2 &= X + \epsilon_2, \end{aligned}$$

where $(X, \epsilon_1, \epsilon_2)$ are unobservables. This setup is called the repeated measurements model, where X is an error-free variable of interest, (ϵ_1, ϵ_2) are measurement errors for X, and (Y_1, Y_2) are repeated noisy measurements on X. We are interested in estimating the densities of X, ϵ_1, and ϵ_2. For sake of simplicity and clarity, we hereafter concentrate on the bivariate case. It is possible to extend our method to the case where more than two noisy measurements on X are available.

Let $\mathrm{i} = \sqrt{-1}$. We impose the following assumptions on the model (1).

Assumption M. (ϵ_1, ϵ_2) *are independent copies of a random variable* ϵ*,* X *is independent of* (ϵ_1, ϵ_2)*,* X *and* ϵ *have square integrable Lebesgue densities* f_X *and* f_ϵ*, respectively, the characteristic functions* $\varphi_X(u) = E[e^{\mathrm{i}uX}]$ *and* $\varphi_\epsilon(u) = E[e^{\mathrm{i}u\epsilon}]$ *vanish nowhere,* $E[\epsilon] = 0$*, and* $E|Y_1|^4 < \infty$*.*

Although these assumptions are standard for the classical measurement error model (e.g., Comte and Kappus, 2015), they are weaker than some existing papers on the repeated measurements model, such as LV (which impose bounded supports of f_X and f_ϵ), Delaigle, Hall and Meister (2008) (which require f_ϵ to be symmetric around zero), and Bonhomme and Robin (2010) (which require the existence of the moment generating functions of Y_1^2 and $Y_1 Y_2$). See also Remark 2 below for a comparison with Kato and Sasaki (2018, 2019). The condition $E[\epsilon] = 0$ is a normalization to identify the densities f_X and f_ϵ.

This paper develops uniform confidence bands for the densities f_X and f_ϵ. To this end, we first introduce the LV estimator for f_X and f_ϵ. Define the characteristic function for the observables (Y_1, Y_2) as

$$\psi(u_1, u_2) = E[e^{\mathrm{i}(u_1 Y_1 + u_2 Y_2)}] = \varphi_X(u_1 + u_2)\varphi_\epsilon(u_1)\varphi_\epsilon(u_2),$$

and let $\psi_1(u_1, u_2) = \partial\psi(u_1, u_2)/\partial u_1 = \mathrm{i}E[Y_1 e^{\mathrm{i}(u_1 Y_1 + u_2 Y_2)}]$ be its derivative with respect to the first argument. Since $E|Y_1| < \infty$ under Assumption M, Kotlarski's identity gives us an explicit identification formula of φ_X, that is

$$\varphi_X(u) = \exp \int_0^u \frac{\psi_1(0, u_2)}{\psi(0, u_2)} du_2.$$

By taking the sample counterpart, LV proposed to estimate φ_X by

$$\hat{\varphi}_X(u) = \exp \int_0^u \frac{\hat{\psi}_1(0, u_2)}{\hat{\psi}(0, u_2)} du_2, \tag{2}$$

where $\hat{\psi}(u_1, u_2) = \frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}(u_1 Y_{1,j} + u_2 Y_{2,j})}$ and $\hat{\psi}_1(u_1, u_2) = \frac{\mathrm{i}}{n}\sum_{j=1}^n Y_{1,j} e^{\mathrm{i}(u_1 Y_{1,j} + u_2 Y_{2,j})}$. Also based on the expression $\varphi_\epsilon(u) = \psi(0, u)/\varphi_X(u)$, the characteristic function φ_ϵ of ϵ can be estimated by

$$\hat{\varphi}_\epsilon(u) = \frac{\hat{\psi}(0, u)}{\hat{\varphi}_X(u)}. \tag{3}$$

By taking the Fourier inversions with regularizations, the densities f_X and f_ϵ can be estimated by

$$\begin{aligned} \hat{f}_X(t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\varphi_K(hu)du, \qquad (4) \\ \hat{f}_\epsilon(t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_\epsilon(u)\varphi_K(hu)du, \end{aligned}$$

respectively, where $\varphi_K(u)$ is the Fourier transform of a kernel function K and $h = h_n$ is a sequence of positive numbers (bandwidths) such that $h \to 0$ as $n \to \infty$. We impose the following assumption on the kernel function.

Assumption K. *The kernel function K satisfies $\int_{\mathbb{R}} K(x)dx = 1$, $\int_{\mathbb{R}} x^\ell K(x)dx = 0$ for $\ell = 1, \ldots, p-1$, and $\int_{\mathbb{R}} |x|^p K(x)dx < \infty$ with a positive even integer p. Also, its Fourier transform φ_K satisfies $\varphi_K(u) = 0$ for any $|u| > 1$.*

This assumption requires the kernel function K to be a p-th order kernel, and its construction is typically done by specifying the Fourier transform φ_K. Let $\zeta : \mathbb{R} \to \mathbb{R}$ be an even function, which is supported on $[-1, 1]$, $(p+2)$-times continuously differentiable, and $\zeta^{(\ell)}(0) = 1$ for $\ell = 0$ and 0 for $\ell = 1, \ldots, p-1$. Then the function $K(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ux}\zeta(u)du$ is real-valued and satisfies $|K(x)| = o(|x|^{-p-2})$ as $|x| \to \infty$ (which follows from a change of variables) so that $(1 \vee |x|^p)K$ is integrable and

$$\int_{\mathbb{R}} x^\ell K(x)dx = \mathrm{i}^{-\ell}\zeta^{(\ell)}(0) = \begin{cases} 1 & \ell = 0, \\ 0 & \ell = 1, \ldots, p-1. \end{cases}$$

Since K is even, we have $\int_{\mathbb{R}} x^p K(x)dx = 0$ for even p. Examples of ζ include $\zeta(u) = (1-u^2)^k \mathbb{I}\{u \in [-1, 1]\}$ for $k \geq p+3$, and

$$\zeta(u) = \begin{cases} 1 & \text{if } |u| \leq c_0, \\ \exp\left\{\frac{-b\exp(-b/(|u|-c_0)^2)}{(|u|-1)^2}\right\} & \text{if } c_0 < |u| < 1, \\ 0 & \text{if } 1 \leq |u|, \end{cases}$$

for $0 < c_0 < 1$ and $b > 0$. For the latter case, ζ is infinitely differentiable with $\zeta^{(\ell)}(0) = 0$ for all $\ell \geq 1$, so that its inverse Fourier transform K, called a flat-top kernel, is of infinite order, i.e., $\int_{\mathbb{R}} x^\ell K(x)dx = 0$ for all integers $\ell \geq 1$ (McMurry and Politis, 2004). We also remark that the sinc kernel $K(x) = \sin(x)/x$ is another example of an infinite-order kernel and its Fourier transform is given by $\varphi_K(u) = \mathbb{I}\{u \in [-1, 1]\}$

To study the asymptotic properties of the estimators in (4) and develop confidence bands for f_X and f_ϵ, we split into two cases based on the density of f_X, i.e., the ordinary smooth (Section 2.2) and supersmooth (Section 2.3) cases. As we show below, the estimators in (4) exhibit rather different asymptotic behaviors.

2.2. **Ordinary smooth case.** In this subsection, we consider the case where the densities f_X and f_ϵ are ordinary smooth. In particular, we impose the following assumption.

Assumption OS. *For some constants $\beta_x, \beta_\epsilon > 1$, $\omega_x, \omega_\epsilon, \omega_{1x}, \delta > 0$, $C_x \geq c_x > 0$, $C_{1x} \geq c_{1x} > 0$, and $C_\epsilon \geq c_\epsilon > 0$, it holds*

$$\begin{aligned} &c_x|u|^{-\beta_x} \leq |\varphi_X(u)| \leq C_x|u|^{-\beta_x} \quad \text{for all } |u| \geq \omega_x, \\ &c_{1x}|u|^{-\delta} \leq \left|\frac{d\log\varphi_X(u)}{du}\right| \leq C_{1x}|u|^{-\delta} \quad \text{for all } |u| \geq \omega_{1x}, \\ &c_\epsilon|u|^{-\beta_\epsilon} \leq |\varphi_\epsilon(u)| \leq C_\epsilon|u|^{-\beta_\epsilon} \quad \text{for all } |u| \geq \omega_\epsilon. \end{aligned}$$

The conditions on φ_X and φ_ϵ are common in the literature of nonparametric deconvolution (see, e.g., Meister, 2009). The conditions $\beta_x, \beta_\epsilon > 1$ are introduced to guarantee consistency of the density estimators. Since the estimators of the characteristic functions in (2) and (3) are defined by the ratios of the (regularized) empirical averages, we need to use the lower and upper bounds of the characteristic functions to obtain suitable bounds of the stochastic and deterministic bias terms of the estimators. A popular example of an ordinary smooth density is the Laplace density. However, it should be noted that our assumption allows the (unknown) measurement error density f_ϵ to be asymmetric. The assumption on $d\log\varphi_X(u)/du$ is very mild. For example, Comte and Kappus (2015) assumed square integrability of $|d\log\varphi_X(u)/du|$.

Our goal is to develop confidence bands for the densities f_X and f_ϵ over a given compact set $\mathcal{T}$ based on Gaussian approximations for the density estimators $\hat{f}_X$ and $\hat{f}_\epsilon$. To this end, we first establish the asymptotic linear representations for $\hat{f}_X$ and $\hat{f}_\epsilon$.

Lemma 1. *[Asymptotic linear forms of $\hat{f}_X$ and $\hat{f}_\epsilon$ for ordinary smooth case] Suppose Assumptions M, K, and OS hold true.*

(i): *If*

$$n^{-\frac{1}{6}}(\log n)^{\frac{3}{2}} \vee \left(\frac{n}{\log n}\right)^{-\frac{1}{2\beta_\epsilon+3}} \vee \left(\frac{n}{(\log n)^3}\right)^{-\frac{1}{2\beta_x+2\beta_\epsilon}} \ll h \ll (n\log n)^{-\frac{1}{2\beta_x+2\beta_\epsilon+1}},$$

then it holds

$$\hat{f}_X(t) - f_X(t) = \frac{1}{n}\sum_{j=1}^n \{L_{X,j}(t) - E[L_{X,1}(t)]\} + o_p(n^{-\frac{1}{2}}h^{-\beta_\epsilon-\frac{3}{2}}(\log n)^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, *where*

$$L_{X,j}(t) = \frac{\mathrm{i}}{2\pi} \int_{\mathbb{R}} e^{-\mathrm{i}ut} \varphi_X(u) \left\{ \int_0^u \frac{Y_{1,j} e^{\mathrm{i}u_2 Y_{2,j}}}{\psi(0,u_2)} du_2 \right\} \varphi_K(hu) du. \tag{5}$$

(ii): *If*

$$n^{-\frac{1}{6}} (\log n)^{\frac{3}{2}} \vee \left(\frac{n}{\log n} \right)^{-\frac{1}{2\beta_x+3}} \vee \left(\frac{n}{(\log n)^3} \right)^{-\frac{1}{2\beta_x+2\beta_\epsilon}} \ll h \ll (n \log n)^{-\frac{1}{2\beta_x+2\beta_\epsilon+1}},$$

then it holds

$$\hat{f}_\epsilon(t) - f_\epsilon(t) = \frac{1}{n} \sum_{j=1}^{n} \{ L_{\epsilon,j}(t) - E[L_{\epsilon,1}(t)] \} + o_p(n^{-\frac{1}{2}} h^{-\beta_x - \frac{3}{2}} (\log n)^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, *where*

$$L_{\epsilon,j}(t) \;=\; \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\mathrm{i}ut} \left\{ \frac{e^{\mathrm{i}uY_{2,j}}}{\varphi_X(u)} - \mathrm{i}\varphi_\epsilon(u) \int_0^u \frac{Y_{1,j} e^{\mathrm{i}u_2 Y_{2,j}}}{\psi(0,u_2)} du_2 \right\} \varphi_K(hu) du. \tag{6}$$

Remark 1. In contrast to $L_{X,j}(t)$, the term $L_{\epsilon,j}(t)$ involves two components due to the influences from the denominator and numerator of $\hat{\varphi}_\epsilon$ in (3). The conditions on the bandwidth h are to control for the bias terms (by the upper bounds) and stochastic errors (by the lower bounds). In particular, the component $n^{-\frac{1}{6}} (\log n)^{\frac{3}{2}}$ in the lower bounds of h is to control errors for the Gaussian coupling by Chernozhukov, Chetverikov and Kato (2016, Theorem 2.1).

Remark 2. Kato and Sasaki (2018, 2019) developed bootstrap confidence bands for the density and regression functions of nonparametric measurement error models, respectively, for the case where the measurement error density is unknown but auxiliary observations from the measurement error density are available. Their method is applicable to the repeated measurements model in (1) as far as

$$\epsilon_1 + \epsilon_2 \text{ and } \epsilon_1 - \epsilon_2 \text{ have the same distribution,} \tag{7}$$

which basically requires symmetry of the distribution of ϵ_2 (given ϵ_1). In this scenario, the transformed data $(Y_1 - Y_2)/2$ can be regarded as the ones generated from f_ϵ so that the characteristic function φ_ϵ can be estimated by $\hat{\varphi}_\epsilon^{KS}(u) = \frac{1}{n} \sum_{j=1}^{n} e^{\mathrm{i}u(Y_{1,j} - Y_{2,j})/2}$. Kato and Sasaki (2018, 2019) plugged in $\hat{\varphi}_\epsilon^{KS}$ to the conventional deconvolution kernel density and regression estimators, respectively, provided intermediate Gaussian approximations for suitably normalized processes of their estimators, and developed valid multiplier bootstrap confidence bands.

In contrast to Kato and Sasaki (2018, 2019), this paper considers the LV estimator which is free from the assumption in (7). Therefore, our confidence bootstrap bands proposed in Section 3.2 below are more robust to the unknown form of f_ϵ. On the other hand, the estimators by Kato and Sasaki (2018, 2019) are more efficient than the LV-type estimators because they exploit the restriction in (7). For example, when both f_X and f_ϵ are ordinary smooth, the uniform convergence rate of Kato and Sasaki's (2018) estimator for f_X (obtained in their Corollary 2) is faster than the one of the LV estimator derived in Corollary 1 (i) below.

Furthermore, due to the different forms of the estimators, the theoretical developments of this paper are very different from Kato and Sasaki (2018, 2019). For example, even for establishing linear forms in the above lemma, we need to invoke strong approximation results as mentioned in Remark 1. Such technical arguments are not necessary for the estimators considered by Kato and Sasaki (2018, 2019) which take simpler forms.

Remark 3. We note that $L_{X,j}(t)$ and $L_{\epsilon,j}(t)$ are real-valued functions. For example, let $\bar{L}_{X,j}(t)$ be the complex conjugate of $L_{X,j}(t)$. Then a change of variables yields

$$\begin{aligned}\bar{L}_{X,j}(t) &= \int_{\mathbb{R}} e^{\mathrm{i}ut}\varphi_X(-u)\left\{-\mathrm{i}\int_0^u \frac{Y_{1,j}e^{-\mathrm{i}u_2Y_{2,j}}}{\psi(0,-u_2)}du_2\right\}\varphi_K(-hu)du \\ &= \int_{\mathbb{R}} e^{\mathrm{i}ut}\varphi_X(-u)\left\{\mathrm{i}\int_0^{-u} \frac{Y_{1,j}e^{\mathrm{i}u_2Y_{2,j}}}{\psi(0,u_2)}du_2\right\}\varphi_K(-hu)du = L_{X,j}(t).\end{aligned}$$

2.3. **Supersmooth case.** In this subsection, we consider the case where the densities f_X and f_ϵ are supersmooth. In particular, we impose the following assumption.

Assumption SS. *For constants $\beta_x, \beta_\epsilon \in \mathbb{R}$, $\rho_x, \omega_x, \omega_\epsilon, \omega_{1x}, \delta_1 > 0$, $\rho_\epsilon \geq 0$, $C_x \geq c_x > 0$, $C_{1x} \geq c_{1x} > 0$, and $C_\epsilon \geq c_\epsilon > 0$, it holds*

$$\begin{aligned}&c_x|u|^{\beta_x}\exp(-|u|^{\rho_x}/\mu_x) \leq |\varphi_X(u)| \leq C_x|u|^{\beta_x}\exp(-|u|^{\rho_x}/\mu_x), \quad \textit{for all } |u| \geq \omega_x, \\ &c_{1x}|u|^{\delta_1} \leq \left|\frac{d\log\varphi_X(u)}{du}\right| \leq C_{1x}|u|^{\delta_1} \quad \textit{for all } |u| \geq \omega_{1x}, \\ &c_\epsilon|u|^{\beta_\epsilon}\exp(-|u|^{\rho_\epsilon}/\mu_\epsilon) \leq |\varphi_\epsilon(u)| \leq C_\epsilon|u|^{\beta_\epsilon}\exp(-|u|^{\rho_\epsilon}/\mu_\epsilon), \quad \textit{for all } |u| \geq \omega_\epsilon.\end{aligned}$$

The conditions on φ_X and φ_ϵ are common conditions for supersmooth densities in the nonparametric deconvolution literature. Similar to the ordinary smooth case, we need lower and upper bounds of the characteristic functions. A popular example of a supersmooth density is the normal density. The assumption on $d\log\varphi_X(u)/du$ is mild and Schennach (2004) imposed a similar condition.

For the supersmooth case, the asymptotic linear representations of $\hat{f}_X$ and $\hat{f}_\epsilon$ are obtained as follows.

Lemma 2. *[Asymptotic linear forms of $\hat{f}_X$ and $\hat{f}_\epsilon$ for supersmooth case] Suppose Assumptions M, K, and SS hold true. Additionally, there exists $c \in (0,1]$ such that $\varphi_K(x) = 1$ for $|x| \leq c$.*

(i): *If*

$$\begin{aligned}\frac{1}{\sqrt{n\log h^{-1}}} &\gg h^{\frac{\rho_x}{q}-\beta_\epsilon-\beta_x+\frac{1}{2}+\delta_1}e^{-\frac{c^{\rho_x}h^{-\rho_x}}{\mu_x}-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \\ \sqrt{\frac{(\log h^{-1})^3}{n}} &\ll h^{-\beta_x-\beta_\epsilon+\delta_1}e^{-\frac{h^{-\rho_x}}{\mu_x}-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_\epsilon+\frac{3}{2}+\delta_1}e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}},\end{aligned}$$

for some $q \geq 1$, then it holds

$$\hat{f}_X(t) - f_X(t) = \frac{1}{n}\sum_{j=1}^n\{M_{X,j}(t) - E[M_{X,1}(t)]\} + o_p(n^{-\frac{1}{2}}h^{\beta_\epsilon-\frac{3}{2}-\delta_1}e^{\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}(\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where

$$M_{X,j}(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\varphi_X(u)\left\{\int_0^u \frac{\varphi_X'(u_2)}{\varphi_X(u_2)}\frac{e^{\mathrm{i}u_2Y_{2,j}}}{\psi(0,u_2)}du_2\right\}\varphi_K(hu)du. \tag{8}$$

(ii): *If*

$$\begin{aligned}\frac{1}{\sqrt{n\log h^{-1}}} &\gg h^{\frac{\rho_\epsilon}{q}-\beta_\epsilon-\beta_x+\frac{1}{2}+\delta_1}e^{-\frac{h^{-\rho_x}}{\mu_x}-\frac{c^{\rho_\epsilon}h^{-\rho_\epsilon}}{\mu_\epsilon}}, \\ \sqrt{\frac{(\log h^{-1})^3}{n}} &\ll h^{-\beta_x-\beta_\epsilon+\delta_1}e^{-\frac{h^{-\rho_x}}{\mu_x}-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_x+\frac{3}{2}+\delta_1}e^{-\frac{h^{-\rho_x}}{\mu_x}},\end{aligned}$$

for some $q \geq 1$, then it holds

$$\hat{f}_\epsilon(t) - f_\epsilon(t) = \frac{1}{n}\sum_{j=1}^{n}\{M_{\epsilon,j}(t) - E[M_{\epsilon,1}(t)]\} + o_p(n^{-\frac{1}{2}}h^{\beta_x - \frac{3}{2} - \delta_1}e^{\frac{h^{-\rho_x}}{\mu_x}}(\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where

$$M_{\epsilon,j}(t) = -\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\varphi_\epsilon(u)\left\{\int_0^u \frac{\varphi_X'(u_2)}{\varphi_X(u_2)}\frac{e^{\mathrm{i}u_2 Y_{2,j}}}{\psi(0,u_2)}du_2\right\}\varphi_K(hu)du. \tag{9}$$

Remark 4. The asymptotic linear representations in this lemma are different from the ones for the ordinary smooth case in Lemma 1. This is due to the fact that the dominant term in the decomposition in (10) is $\Delta_2(u)$ for the supersmooth case (instead of $\Delta_1(u)$ for the ordinary smooth case). Similar to Remark 3, we can see that $M_{X,j}(t)$ and $M_{\epsilon,j}(t)$ are real-valued functions. The ratio $\frac{\varphi_X'(u_2)}{\varphi_X(u_2)}$ in the definitions of $M_{X,j}(t)$ and $M_{\epsilon,j}(t)$ appears due to the equality $\frac{\varphi_X'(u_2)}{\varphi_X(u_2)} = \frac{\psi_1(0,u_2)}{\psi(0,u_2)}$.

Remark 5. In this lemma, we can set $q = 1$ when f_X satisfies Assumption SS with $\beta_x > 0$ (for Part (i)) and when f_ϵ satisfies Assumption SS with $\beta_\epsilon > 0$ (for Part (ii)). See the proof of Theorem in Kurisu and Otsu (2021) for details.

2.4. **Mixed case.** We now consider mixed cases, where (a) f_X is ordinary smooth and f_ϵ is supersmooth, or (b) f_X is supersmooth and f_ϵ is ordinary smooth. By adapting the proofs of Lemmas 1 and 2, the linearization results for the mixed cases are obtained as follows.

Lemma 3. *[Asymptotic linear forms of $\hat{f}_X$ and $\hat{f}_\epsilon$ for mixed cases] Suppose Assumptions M and K hold true. Additionally, there exists $c \in (0,1]$ such that $\varphi_K(x) = 1$ for $|x| \leq c$.*

(a-i): *Suppose that f_X satisfies Assumption OS and f_ϵ satisfies Assumption SS. If*

$$\frac{1}{\sqrt{n\log h^{-1}}} \gg h^{-\beta_\epsilon + \beta_x + \frac{1}{2}}e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}},$$
$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{\beta_x - \beta_\epsilon}e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_\epsilon + \frac{3}{2}}e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}},$$

for some $q \geq 1$, then it holds

$$\hat{f}_X(t) - f_X(t) = \frac{1}{n}\sum_{j=1}^{n}\{L_{X,j}(t) - E[L_{X,1}(t)]\} + o_p(n^{-\frac{1}{2}}h^{\beta_\epsilon - \frac{3}{2}}e^{\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}(\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where $L_{X,j}(t)$ is defined in (5).

(a-ii): *Suppose that f_X satisfies Assumption OS and f_ϵ satisfies Assumption SS. If*

$$\frac{1}{\sqrt{n\log h^{-1}}} \gg h^{\beta_\epsilon - \beta_x + \frac{1}{2}}e^{-\frac{h^{-\rho_x}}{\mu_x}},$$
$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{\beta_\epsilon - \beta_x}e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_x + \frac{3}{2}}e^{-\frac{h^{-\rho_x}}{\mu_x}},$$

for some $q \geq 1$, then it holds then it holds

$$\hat{f}_\epsilon(t) - f_\epsilon(t) = \frac{1}{n}\sum_{j=1}^{n}\{L_{\epsilon,j}(t) - E[L_{\epsilon,1}(t)]\} + o_p(n^{-\frac{1}{2}}h^{-\beta_x - \frac{3}{2}}(\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where where $L_{\epsilon,j}(t)$ is defined in (6).

(b-i): *Suppose that f_X satisfies Assumption SS and f_ϵ satisfies Assumption OS. If*

$$\frac{1}{\sqrt{n \log h^{-1}}} \gg h^{\frac{\rho_x}{q} - \beta_x + \beta_\epsilon + \frac{1}{2} + \delta_1} e^{-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}},$$

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x + \beta_\epsilon + \delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{\beta_\epsilon + \frac{3}{2} + \delta_1},$$

for some $q \geq 1$, then it holds

$$\hat{f}_X(t) - f_X(t) = \frac{1}{n} \sum_{j=1}^{n} \{M_{X,j}(t) - E[M_{X,1}(t)]\} + o_p(n^{-\frac{1}{2}} h^{-\beta_\epsilon - \frac{3}{2} - \delta_1} (\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where $M_{X,j}(t)$ is defined in (8).

(b-ii): *Suppose that f_X satisfies Assumption SS and f_ϵ satisfies Assumption OS. If*

$$\frac{1}{\sqrt{n \log h^{-1}}} \gg h^{\frac{\rho_\epsilon}{q} + \beta_x - \beta_\epsilon + \frac{1}{2} + \delta_1} e^{-\frac{c^{\rho_\epsilon} h^{-\rho_\epsilon}}{\mu_\epsilon}},$$

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x + \beta_\epsilon + \delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{\beta_x + \frac{3}{2} + \delta_1},$$

for some $q \geq 1$, then it holds

$$\hat{f}_\epsilon(t) - f_\epsilon(t) = \frac{1}{n} \sum_{j=1}^{n} \{M_{\epsilon,j}(t) - E[M_{\epsilon,1}(t)]\} + o_p(n^{-\frac{1}{2}} h^{\beta_x - \frac{3}{2} - \delta_1} e^{\frac{h^{-\rho_x}}{\mu_x}} (\log h^{-1})^{-\frac{1}{2}}),$$

uniformly over $t \in \mathcal{T}$, where $M_{\epsilon,j}(t)$ is defined in (9).

Remark 6. Note that the forms of the asymptotic linear terms for both $\hat{f}_X$ and $\hat{f}_\epsilon$ are determined by whether f_X is ordinary or super-smooth. For the case of ordinary smooth f_X and supersmooth f_ϵ, the linear terms of $\hat{f}_X$ and $\hat{f}_\epsilon$ are same as the ones in Lemma 1. For the case of supersmooth f_X and ordinary smooth f_ϵ, the linear terms of $\hat{f}_X$ and $\hat{f}_\epsilon$ are same as the ones in Lemma 2. Technically, this is due to the fact that the dominant linear terms are determined by the relative orders of the terms $\Delta_1(u)$ and $\Delta_2(u)$ in (10), which depend only on the tail behaviors of f_X. The conditions on the bandwidth are analogous to the ones in Lemmas 1 and 2.

3. Applications

3.1. **Refined convergence rates.** As direct applications of our linearization lemmas in the last section, we can derive the convergence rates of the density estimators $\hat{f}_X$ and $\hat{f}_\epsilon$, which are faster than the ones obtained in the existing literature. Inspections of the proofs of Lemmas 1 and 2 yield the following theorem.

Theorem 1. *Suppose Assumptions M and K hold true.*

(i): *Suppose Assumption OS holds true. If*

$$n^{-\frac{1}{6}} (\log n)^{\frac{3}{2}} \vee \left(\frac{n}{\log n}\right)^{-\frac{1}{2\beta_\epsilon + 3}} \vee \left(\frac{n}{(\log n)^3}\right)^{-\frac{1}{2\beta_x + 2\beta_\epsilon}} \ll h \ll 1,$$

then

$$\sup_{t \in \mathcal{T}} |\hat{f}_X(t) - f_X(t)| = O_p(n^{-\frac{1}{2}} h^{-\beta_\epsilon - \frac{3}{2}} (\log n)^{\frac{1}{2}} + h^{\beta_x - 1}).$$

Moreover, if

$$n^{-\frac{1}{6}}(\log n)^{\frac{3}{2}} \vee \left(\frac{n}{\log n}\right)^{-\frac{1}{2\beta_x+3}} \vee \left(\frac{n}{(\log n)^3}\right)^{-\frac{1}{2\beta_x+2\beta_\epsilon}} \ll h \ll 1,$$

then

$$\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| = O_p(n^{-\frac{1}{2}} h^{-\beta_x - \frac{3}{2}} (\log n)^{\frac{1}{2}} + h^{\beta_\epsilon - 1}).$$

(ii): *Suppose Assumption SS holds true. If $h \ll 1$ and*

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x-\beta_\epsilon+\delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x} - \frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_\epsilon + \frac{3}{2}+\delta_1} e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}},$$

then for $\varsigma_{h,q}^x = h^{\frac{\rho_x}{q}-\beta_x-1} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right)$, it holds

$$\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| = O_p(n^{-\frac{1}{2}} h^{\beta_\epsilon - \frac{3}{2} - \delta_1} e^{\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}} (\log h^{-1})^{\frac{1}{2}} + \varsigma_{h,q}^x).$$

Moreover, if $h \ll 1$ and

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x-\beta_\epsilon+\delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x} - \frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_x + \frac{3}{2}+\delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x}},$$

Then for $\varsigma_{h,q}^\epsilon = h^{\frac{\rho_\epsilon}{q}-\beta_\epsilon-1} \exp\left(-\frac{c^{\rho_\epsilon} h^{-\rho_\epsilon}}{\mu_\epsilon}\right)$, it holds

$$\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| = O_p(n^{-\frac{1}{2}} h^{\beta_x - \frac{3}{2} - \delta_1} e^{\frac{h^{-\rho_x}}{\mu_x}} (\log h^{-1})^{\frac{1}{2}} + \varsigma_{h,q}^\epsilon).$$

Remark 7. The uniform convergence rates in this theorem are faster than those given in Kurisu and Otsu (2021) even though the rates in Kurisu and Otsu (2021) are faster than the ones in LV or Bonhomme and Robin (2010). For example, under Assumptions M, K, and OS and $n^{-\frac{1}{4\beta_x+4\beta_\epsilon+2}}(\log n) \ll h \ll 1$ (for f_X) or $n^{-\frac{1}{6\beta_x+4\beta_\epsilon+2}}(\log n) \ll h \ll 1$ (for f_ϵ), the convergence rates in Kurisu and Otsu (2021) are

$$\begin{aligned}
\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| &= O_p(n^{-\frac{1}{2}} h^{-2\beta_x-2\beta_\epsilon-2}(\log n) + h^{\beta_x-1}), \\
\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| &= O_p(n^{-\frac{1}{2}} h^{-3\beta_x-2\beta_x-2}(\log n) + h^{\beta_\epsilon-1}).
\end{aligned}$$

A main reason for this refinement is that we employ intermediate Gaussian approximations for both the characteristic and density functions estimators instead of bounding those functions via maximal inequalities.

Similarly, the uniform convergence rates of the estimators for the mixed cases are obtained as follows.

Theorem 2. *Suppose Assumptions M and K hold true. Additionally, there exists $c \in (0,1]$ such that $\varphi_K(x) = 1$ for $|x| \le c$.*

(a-i): *Suppose that f_X satisfies Assumption OS and f_ϵ satisfies Assumption SS. If $h \ll 1$ and*

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{\beta_x-\beta_\epsilon} e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_\epsilon+\frac{3}{2}} e^{-\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}},$$

then

$$\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| = O_p\left(n^{-\frac{1}{2}} h^{\beta_\epsilon - \frac{3}{2}} e^{\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}} (\log h^{-1})^{\frac{1}{2}} + h^{\beta_x-1}\right).$$

(a-ii): *Suppose that f_X satisfies Assumption OS and f_ϵ satisfies Assumption SS. If $h \ll 1$ and*

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{\beta_\epsilon-\beta_x} e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{-\beta_x+\frac{3}{2}} e^{-\frac{h^{-\rho_x}}{\mu_x}},$$

then

$$\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| = O_p\left(n^{-\frac{1}{2}} h^{-\beta_x-\frac{3}{2}} (\log h^{-1})^{\frac{1}{2}} + \varsigma^{\epsilon}_{h,q}\right).$$

(b-i): *Suppose that f_X satisfies Assumption SS and f_ϵ satisfies Assumption OS. If $h \ll 1$ and*

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x+\beta_\epsilon+\delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{\beta_\epsilon+\frac{3}{2}+\delta_1},$$

then

$$\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| = O_p\left(n^{-\frac{1}{2}} h^{-\beta_\epsilon-\frac{3}{2}-\delta_1} (\log h^{-1})^{\frac{1}{2}} + \varsigma^{x}_{h,q}\right).$$

(b-ii): *Suppose that f_X satisfies Assumption SS and f_ϵ satisfies Assumption OS. If $h \ll 1$ and*

$$\sqrt{\frac{(\log h^{-1})^3}{n}} \ll h^{-\beta_x+\beta_\epsilon+\delta_1} e^{-\frac{h^{-\rho_x}}{\mu_x}}, \qquad \sqrt{\frac{\log h^{-1}}{n}} \ll h^{\beta_x+\frac{3}{2}+\delta_1},$$

then

$$\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| = O_p\left(n^{-\frac{1}{2}} h^{\beta_x-\frac{3}{2}-\delta_1} e^{\frac{h^{-\rho_x}}{\mu_x}} (\log h^{-1})^{\frac{1}{2}} + h^{\beta_\epsilon-1}\right).$$

The proof of this theorem is analogous to the one in Theorem 1, and similar comments to Remark 7 apply, i.e., our uniform convergence rates are typically faster than the ones obtained in the literature. To clarify this point, we choose the bandwidth h to balance the terms in the convergence rates in Theorems 1 and 2 and derive the uniform convergence rates depending only on n as follows. Let $l_n = n/(\log n)$ and $\ell_n = n/(\log\log n)$.

Corollary 1.

(i): *[Both f_X and f_ϵ are ordinary smooth] Under the assumptions of Theorem 1 (i), it holds*

$$\begin{aligned}
\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| &= O_p\left(l_n^{\frac{1-\beta_x}{2\beta_x+2\beta_\epsilon+1}}\right) \text{ by setting } h \sim l_n^{-\frac{1}{2\beta_x+2\beta_\epsilon+1}},\\
\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| &= O_p\left(l_n^{\frac{1-\beta_\epsilon}{2\beta_x+2\beta_\epsilon+1}}\right) \text{ by setting } h \sim l_n^{-\frac{1}{2\beta_x+2\beta_\epsilon+1}}.
\end{aligned}$$

(ii): *[Both f_X and f_ϵ are supersmooth] Under the assumptions of Theorem 1 (ii) with $c = 1$ and $\rho_x = \rho_\epsilon = \rho$, it holds*

$$\begin{aligned}
\sup_{t\in\mathcal{T}} |\hat{f}_X(t) - f_X(t)| &= O_p\left(\ell_n^{-\frac{\mu_\epsilon}{2\mu_x+2\mu_\epsilon}} (\log \ell_n)^{\max\left\{\frac{3/2+\delta_1-\beta_\epsilon}{\rho}, \frac{\beta_x+1-\rho_x/q}{\rho}\right\}} (\log\log n)^{1/2}\right),\\
\sup_{t\in\mathcal{T}} |\hat{f}_\epsilon(t) - f_\epsilon(t)| &= O_p\left(\ell_n^{-\frac{\mu_x}{2\mu_x+2\mu_\epsilon}} (\log \ell_n)^{\max\left\{\frac{3/2+\delta_1-\beta_x}{\rho}, \frac{\beta_\epsilon+1-\rho_\epsilon/q}{\rho}\right\}} (\log\log n)^{1/2}\right),
\end{aligned}$$

by setting $h \sim (2^{-1}\mu \log \ell_n)^{-1/\rho}$ with $\mu = \mu_x\mu_\epsilon/(\mu_x + \mu_\epsilon)$.

(iii): *[f_X is ordinary smooth and f_ϵ are supersmooth] Under the assumptions of Theorem 2 (a-i) (for $\hat{f}_X$) or (a-ii) (for $\hat{f}_\epsilon$) with $c = 1$, it holds*

$$\begin{aligned}\sup_{t\in\mathcal{T}}|\hat{f}_X(t) - f_X(t)| &= O_p\left(\ell_n^{-\frac{1}{2}+\frac{\alpha}{2}}(\log \ell_n)^{\frac{3/2-\beta_\epsilon}{\rho_\epsilon}}(\log\log n)^{1/2} + (\log \ell_n)^{\frac{1-\beta_x}{\rho_\epsilon}}\right),\\ \sup_{t\in\mathcal{T}}|\hat{f}_\epsilon(t) - f_\epsilon(t)| &= O_p\left(\ell_n^{-\frac{1}{2}}(\log \ell_n)^{\frac{3/2+\beta_x}{\rho_\epsilon}}(\log\log n)^{1/2} + \ell_n^{-\frac{\alpha}{2}}(\log \ell_n)^{\frac{\beta_\epsilon+1-\rho_\epsilon/q}{\rho_\epsilon}}\right),\end{aligned}$$

by setting $h \sim (2^{-1}\alpha\mu_\epsilon \log \ell_n)^{-1/\rho_\epsilon}$ for $\alpha \in (0, 1]$.

(iv): *[f_X is supersmooth and f_ϵ are ordinary smooth] Under the assumptions of Theorem 2 (b-i) (for $\hat{f}_X$) or (b-ii) (for $\hat{f}_\epsilon$) with $c = 1$, it holds*

$$\begin{aligned}\sup_{t\in\mathcal{T}}|\hat{f}_X(t) - f_X(t)| &= O_p\left(\ell_n^{-\frac{1}{2}}(\log \ell_n)^{\frac{3/2+\beta_\epsilon+\delta_1}{\rho_x}}(\log\log n)^{1/2} + \ell_n^{-\frac{\alpha}{2}}(\log \ell_n)^{\frac{\beta_x+1-\rho_x/q}{\rho_x}}\right),\\ \sup_{t\in\mathcal{T}}|\hat{f}_\epsilon(t) - f_\epsilon(t)| &= O_p\left(\ell_n^{-\frac{1}{2}+\frac{\alpha}{2}}(\log \ell_n)^{\frac{3/2+\delta_1-\beta_x}{\rho_x}}(\log\log n)^{1/2} + (\log \ell_n)^{\frac{1-\beta_\epsilon}{\rho_x}}\right),\end{aligned}$$

by setting $h \sim (2^{-1}\alpha\mu_x \log \ell_n)^{-1/\rho_x}$ for $\alpha \in (0, 1]$.

The additional assumption $\rho_x = \rho_\epsilon = \rho$ in Part (ii) of this corollary is imposed to simplify the presentation and may be relaxed. The above uniform convergence rates are faster than those in Li and Vuong (1998) and Kurisu and Otsu (2021) except for the cases of Part (iii) on $\hat{f}_X$ and Part (iv) on $\hat{f}_\epsilon$, where the rates are same as theirs.

3.2. **Confidence bands for density functions.** In this subsection, we apply our linearization lemmas to construct confidence bands for the densities f_X and f_ϵ. In particular, we develop Gaussian multiplier bootstrap approximations by perturbing the sample counterparts of the linear terms in Lemmas 1 and 2. Let $(\hat{\varphi}_X, \hat{\varphi}_\epsilon, \hat{\psi}, \hat{\psi}_1)$ be the estimators defined in Section 2 based on the full sample of size n. Then we define the sample counterparts of the asymptotic linear terms as

$$\begin{aligned}\hat{L}_{X,j}(t) &= \frac{\mathrm{i}}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left\{\int_0^u \frac{Y_{1,j}e^{\mathrm{i}u_2Y_{2,j}}}{\hat{\psi}(0,u_2)}du_2\right\}\varphi_K(hu)du,\\ \hat{L}_{\epsilon,j}(t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\left\{\frac{e^{\mathrm{i}u_2Y_{2,j}}}{\hat{\varphi}_X(u)} - \mathrm{i}\hat{\varphi}_\epsilon(u)\int_0^u \frac{Y_{1,j}e^{\mathrm{i}u_2Y_{2,j}}}{\hat{\psi}(0,u_2)}du_2\right\}\varphi_K(hu)du,\\ \hat{M}_{X,j}(t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left\{\int_0^u \widehat{\left(\frac{\varphi_X'(u_2)}{\varphi_X(u_2)}\right)}\frac{e^{\mathrm{i}u_2Y_{2,j}}}{\hat{\psi}(0,u_2)}du_2\right\}\varphi_K(hu)du,\\ \hat{M}_{\epsilon,j}(t) &= -\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_\epsilon(u)\left\{\int_0^u \widehat{\left(\frac{\varphi_X'(u_2)}{\varphi_X(u_2)}\right)}\frac{e^{\mathrm{i}u_2Y_{2,j}}}{\hat{\psi}(0,u_2)}du_2\right\}\varphi_K(hu)du,\end{aligned}$$

where $\widehat{\left(\frac{\varphi_X'(u_2)}{\varphi_X(u_2)}\right)} = \widehat{\left(\frac{\psi_1(0,u_2)}{\psi(0,u_2)}\right)} = \frac{\hat{\psi}_1(0,u_2)}{\hat{\psi}(0,u_2)}$. Although it is natural to apply the multiplier bootstrap these sample counterparts, the approximation errors for the linear terms decay too slow to construct the bootstrap counterparts by using the full sample. Therefore, we propose to approximate the distributions of (suprema of) $\hat{f}_X - f_X$ and $\hat{f}_\epsilon - f_\epsilon$ by using the subsample-based bootstrap counterparts of the linearization terms:

$$\hat{L}_X^\xi(t) = \frac{1}{m}\sum_{j=1}^m \xi_j\left\{\hat{L}_{X,j}(t) - \frac{1}{m}\sum_{k=1}^m \hat{L}_{X,k}(t)\right\}, \qquad \hat{L}_\epsilon^\xi(t) = \frac{1}{m}\sum_{j=1}^m \xi_j\left\{\hat{L}_{\epsilon,j}(t) - \frac{1}{m}\sum_{k=1}^m \hat{L}_{\epsilon,k}(t)\right\},$$

for the ordinary smooth case, and

$$\hat{M}_X^{\xi}(t) = \frac{1}{m}\sum_{j=1}^{m}\xi_j\left\{\hat{M}_{X,j}(t) - \frac{1}{m}\sum_{k=1}^{m}\hat{M}_{X,k}(t)\right\}, \qquad \hat{M}_\epsilon^{\xi}(t) = \frac{1}{m}\sum_{j=1}^{m}\xi_j\left\{\hat{M}_{\epsilon,j}(t) - \frac{1}{m}\sum_{k=1}^{m}\hat{M}_{\epsilon,k}(t)\right\},$$

for the super smooth case, where $m < n$ is the subsample size, and $\xi_1, \ldots, \xi_m \sim N(0,1)$ are independent from the data $\mathcal{Y}_n = \{Y_{1,j}, Y_{2,j}\}_{j=1}^n$.

To show validity of our bootstrap approximations, we impose the following assumptions in this subsection.

Assumption OSB.

(i): *Assumptions in Lemma 1 hold true by replacing n with m.*

(ii): *[Undersmoothing] Let*

$$\sigma_{X,m}^2(t) = \mathit{Var}(L_{X,1}(t)), \quad s_{X,m}^2 = \inf_{t\in\mathcal{T}}\sigma_{X,m}^2(t),$$
$$\sigma_{\epsilon,m}^2(t) = \mathit{Var}(L_{\epsilon,1}(t)), \quad s_{\epsilon,m}^2 = \inf_{t\in\mathcal{T}}\sigma_{\epsilon,m}^2(t).$$

Assume that

$$\sqrt{m}s_{X,m}^{-1}h^{\beta_x-1} = o(\log^{-\frac{1}{2}} m) \quad \text{for } f_X,$$
$$\sqrt{m}s_{\epsilon,m}^{-1}h^{\beta_\epsilon-1} = o(\log^{-\frac{1}{2}} m) \quad \text{for } f_\epsilon.$$

(iii): *[Variance estimation] Define $\sigma_{X,m}(t) = \sqrt{\sigma_{X,m}^2(t)}$ and $\sigma_{\epsilon,m}(t) = \sqrt{\sigma_{\epsilon,m}^2(t)}$. There exist estimators $\hat{\sigma}_{X,m}^2(t)$ and $\hat{\sigma}_{\epsilon,m}^2(t)$ such that*

$$\sup_{t\in\mathcal{T}}|\hat{\sigma}_{X,m}(t)/\sigma_{X,m}(t) - 1| = o_p(\log^{-1} m) \quad \text{for } f_X,$$
$$\sup_{t\in\mathcal{T}}|\hat{\sigma}_{\epsilon,m}(t)/\sigma_{\epsilon,m}(t) - 1| = o_p(\log^{-1} m) \quad \text{for } f_\epsilon,$$

where $\hat{\sigma}_{X,m}(t) = \sqrt{\hat{\sigma}_{X,m}^2(t)}$ and $\hat{\sigma}_{\epsilon,m}(t) = \sqrt{\hat{\sigma}_{\epsilon,m}^2(t)}$.

(iv): *[Bandwidth and subsample size] As $n \to \infty$, it holds $\sqrt{m/n} = o((\log n)^{-\frac{1}{2}})$,*

$$m^{-\frac{1}{4}}s_{X,m}^{-1}h^{-\beta_\epsilon-2}(\log m)^{\frac{5}{4}} = o((\log m)^{-\frac{1}{2}}),$$
$$n^{-\frac{1}{2}}s_{X,m}^{-1}h^{-\beta_\epsilon-2}(\log n)(\log m)^{\frac{1}{2}} = o(1), \text{ and}$$
$$\left(\frac{m}{n}\right)^{\frac{1}{2}}s_{X,m}^{-1}h^{-\beta_\epsilon-2}\left(\frac{\log n}{\log m}\right) = o((\log m)^{-\frac{1}{2}}) \text{ for } f_X,$$
$$m^{-\frac{1}{4}}s_{\epsilon,m}^{-1}h^{-\beta_x-2}(\log m)^{\frac{5}{4}} = o((\log m)^{-\frac{1}{2}}),$$
$$n^{-\frac{1}{2}}s_{\epsilon,m}^{-1}h^{-\beta_x-2}(\log n)(\log m)^{\frac{1}{2}} = o(1), \text{ and}$$
$$\left(\frac{m}{n}\right)^{\frac{1}{2}}s_{\epsilon,m}^{-1}h^{-\beta_x-2}\left(\frac{\log n}{\log m}\right) = o((\log m)^{-\frac{1}{2}}) \text{ for } f_\epsilon.$$

Condition (ii) is an undersmoothing condition. Condition (iii) is on approximation error of $\sigma_{X,m}(t)$ and $\sigma_{\epsilon,m}(t)$ by $\hat{\sigma}_{X,m}(t)$ and $\hat{\sigma}_{\epsilon,m}(t)$, respectively. We need the condition to approximate $\frac{\hat{f}_X(t)-f_X(t)}{\hat{\sigma}_{X,m}(t)}$ (or $\frac{\hat{f}_\epsilon(t)-f_\epsilon(t)}{\hat{\sigma}_{\epsilon,m}(t)}$) by $\frac{\hat{f}_X(t)-f_X(t)}{\sigma_{X,m}(t)}$ (or $\frac{\hat{f}_\epsilon(t)-f_\epsilon(t)}{\sigma_{\epsilon,m}(t)}$). Condition (iv) is a set of other technical assumptions. Indeed, for f_X, we need the first assumption in Condition (iv) to approximate the supremum of $\frac{\sigma_{X,m}^{-1}(t)}{n}\sum_{j=1}^n\{L_{X,j}(t) - E[L_{X,1}(t)]\}$ by the supremum of a Gaussian random variable. We also need the second and third assumptions to show asymptotic validity of our bootstrap-based uniform confidence bands. Precisely, we can replace $\hat{L}_{X,j}(t)$ (or $\hat{L}_{\epsilon,j}(t)$) for $j = 1, \ldots, m$ in the definition of $\hat{L}^{\xi}_{X}(t)$ (or $\hat{L}^{\xi}_{\epsilon}(t)$) with $L_{X,j}(t)$ (or $L_{\epsilon,j}(t)$) for $j = 1, \ldots, m$ under the second and third assumptions in Condition (iv). The same comment applies to Assumption SSB.

Assumption SSB.

(i): *Assumptions in Lemma 2 hold true by replacing n with m.*

(ii): *[Undersmoothing] Let*

$$\sigma^2_{X,m}(t) = Var(M_{X,1}(t)), \quad s^2_{X,m} = \inf_{t\in\mathcal{T}} \sigma^2_{X,m}(t),$$
$$\sigma^2_{\epsilon,m}(t) = Var(M_{\epsilon,1}(t)), \quad s^2_{\epsilon,m} = \inf_{t\in\mathcal{T}} \sigma^2_{\epsilon,m}(t).$$

Assume that

$$\sqrt{m} s^{-1}_{X,m} h^{\frac{\rho_x}{q}-\beta_x-1} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right) = o(\log^{-\frac{1}{2}} m) \quad \textit{for } f_X,$$
$$\sqrt{m} s^{-1}_{\epsilon,m} h^{\frac{\rho_\epsilon}{q}-\beta_\epsilon-1} \exp\left(-\frac{c^{\rho_\epsilon} h^{-\rho_\epsilon}}{\mu_\epsilon}\right) = o(\log^{-\frac{1}{2}} m) \quad \textit{for } f_\epsilon.$$

(iii): *[Variance estimation] Define $\sigma_{X,m}(t) = \sqrt{\sigma^2_{X,m}(t)}$ and $\sigma_{\epsilon,m}(t) = \sqrt{\sigma^2_{\epsilon,m}(t)}$. There exist estimators $\hat{\sigma}^2_{X,m}(t)$ and $\hat{\sigma}^2_{\epsilon,m}(t)$ such that*

$$\sup_{t\in\mathcal{T}} |\hat{\sigma}_{X,m}(t)/\sigma_{X,m}(t) - 1| = o_p(\log^{-1} m) \quad \textit{for } f_X,$$
$$\sup_{t\in\mathcal{T}} |\hat{\sigma}_{\epsilon,m}(t)/\sigma_{\epsilon,m}(t) - 1| = o_p(\log^{-1} m) \quad \textit{for } f_\epsilon,$$

where $\hat{\sigma}_{X,m}(t) = \sqrt{\hat{\sigma}^2_{X,m}(t)}$ and $\hat{\sigma}_{\epsilon,m}(t) = \sqrt{\hat{\sigma}^2_{\epsilon,m}(t)}$.

(iv): *[Bandwidth and subsample size] $\sqrt{m/n} = o((\log n)^{-\frac{1}{2}})$,*

$$m^{-\frac{1}{4}} s^{-1}_{X,m} h^{\beta_\epsilon-2} \exp\left(\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}\right) (\log m)^{\frac{5}{4}} = o((\log m)^{-\frac{1}{2}}),$$
$$n^{-\frac{1}{2}} s^{-1}_{X,m} h^{\beta_\epsilon-2-\delta_1} \exp\left(\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}\right) (\log n)(\log m)^{\frac{1}{2}} = o(1), \ \textit{and}$$
$$\left(\frac{m}{n}\right)^{\frac{1}{2}} s^{-1}_{X,m} h^{\beta_\epsilon-2-\delta_1} \exp\left(\frac{h^{-\rho_\epsilon}}{\mu_\epsilon}\right) \left(\frac{\log n}{\log m}\right) = o((\log m)^{-\frac{1}{2}}) \ \textit{for } f_X,$$
$$m^{-\frac{1}{4}} s^{-1}_{\epsilon,m} h^{\beta_x-2} \exp\left(\frac{h^{-\rho_x}}{\mu_x}\right) (\log m)^{\frac{5}{4}} = o((\log m)^{-\frac{1}{2}}),$$
$$n^{-\frac{1}{2}} s^{-1}_{\epsilon,m} h^{\beta_x-2-\delta_1} \exp\left(\frac{h^{-\rho_x}}{\mu_x}\right) (\log n)(\log m)^{\frac{1}{2}} = o(1), \ \textit{and}$$
$$\left(\frac{m}{n}\right)^{\frac{1}{2}} s^{-1}_{\epsilon,m} h^{\beta_x-2-\delta_1} \exp\left(\frac{h^{-\rho_x}}{\mu_x}\right) \left(\frac{\log n}{\log m}\right) = o((\log m)^{-\frac{1}{2}}) \ \textit{for } f_\epsilon.$$

For the variance estimation, one may use

$$\hat{\sigma}^2_{X,m}(t) = \begin{cases} \frac{1}{m}\sum_{j=1}^m \hat{L}^2_{X,j}(t) - \left(\frac{1}{m}\sum_{k=1}^m \hat{L}_{X,k}(t)\right)^2 & \text{under Assumption OSB} \\ \frac{1}{m}\sum_{j=1}^m \hat{M}^2_{X,j}(t) - \left(\frac{1}{m}\sum_{k=1}^m \hat{M}_{X,k}(t)\right)^2 & \text{under Assumption SSB} \end{cases},$$

$$\hat{\sigma}^2_{\epsilon,m}(t) = \begin{cases} \frac{1}{m}\sum_{j=1}^m \hat{L}^2_{\epsilon,j}(t) - \left(\frac{1}{m}\sum_{k=1}^m \hat{L}_{\epsilon,k}(t)\right)^2 & \text{under Assumption OSB} \\ \frac{1}{m}\sum_{j=1}^m \hat{M}^2_{\epsilon,j}(t) - \left(\frac{1}{m}\sum_{k=1}^m \hat{M}_{\epsilon,k}(t)\right)^2 & \text{under Assumption SSB} \end{cases}.$$

Theorem 3. *[Bootstrap approximations] Suppose Assumptions M, K, and OSB or SSB hold true. Then as $n \to \infty$,*

$$\sup_{z\in\mathbb{R}}\left|\Pr\left\{\sqrt{m}\sup_{t\in\mathcal{T}}\left|\frac{\hat{f}_X(t)-f_X(t)}{\hat{\sigma}_{X,m}(t)}\right|\le z\right\}-\Pr\left\{\sqrt{m}\sup_{t\in\mathcal{T}}\left|\frac{\hat{B}^{\xi}_X(t)}{\hat{\sigma}_{X,m}(t)}\right|\le z\middle|\mathcal{Y}_n\right\}\right| \xrightarrow{p} 0.$$
$$\sup_{z\in\mathbb{R}}\left|\Pr\left\{\sqrt{m}\sup_{t\in\mathcal{T}}\left|\frac{\hat{f}_\epsilon(t)-f_\epsilon(t)}{\hat{\sigma}_{\epsilon,m}(t)}\right|\le z\right\}-\Pr\left\{\sqrt{m}\sup_{t\in\mathcal{T}}\left|\frac{\hat{B}^{\xi}_\epsilon(t)}{\hat{\sigma}_{\epsilon,m}(t)}\right|\le z\middle|\mathcal{Y}_n\right\}\right| \xrightarrow{p} 0,$$

where $(\hat{B}^{\xi}_X,\hat{B}^{\xi}_\epsilon)=(\hat{L}^{\xi}_X,\hat{L}^{\xi}_\epsilon)$ under Assumption OSB and $(\hat{B}^{\xi}_X,\hat{B}^{\xi}_\epsilon)=(\hat{M}^{\xi}_X,\hat{M}^{\xi}_\epsilon)$ under Assumption SSB.

Let $\hat{c}_X^{1-\tau}$and $\hat{c}_\epsilon^{1-\tau}$ be the conditional $(1-\tau)$-th quantiles of $\sqrt{m}\sup_{t\in\mathcal{T}}|\hat{L}^{\xi}_X(t)/\hat{\sigma}_{X,m}(t)|$ (or $\sqrt{m}\sup_{t\in\mathcal{T}}|\hat{M}^{\xi}_X(t)/\hat{\sigma}_{X,m}(t)|$) and $\sqrt{m}\sup_{t\in\mathcal{T}}|\hat{L}^{\xi}_\epsilon(t)/\hat{\sigma}_{\epsilon,m}(t)|$ (or $\sqrt{m}\sup_{t\in\mathcal{T}}|\hat{M}^{\xi}_\epsilon(t)/\hat{\sigma}_{\epsilon,m}(t)|$) given the data $\mathcal{Y}_n$, respectively. Then the confidence bands of f_X and f_ϵ over $\mathcal{T}$ are constructed as

$$\begin{aligned}\hat{\mathcal{C}}_X(t) &= [\hat{f}_X(t)-\hat{\sigma}_{X,m}(t)\hat{c}_X^{1-\tau}/\sqrt{m},\hat{f}_X(t)+\hat{\sigma}_{X,m}(t)\hat{c}_X^{1-\tau}/\sqrt{m}],\\ \hat{\mathcal{C}}_\epsilon(t) &= [\hat{f}_\epsilon(t)-\hat{\sigma}_{\epsilon,m}(t)\hat{c}_\epsilon^{1-\tau}/\sqrt{m},\hat{f}_\epsilon(t)+\hat{\sigma}_{\epsilon,m}(t)\hat{c}_\epsilon^{1-\tau}/\sqrt{m}],\end{aligned}$$

for $t\in\mathcal{T}$, respectively. For completeness, we present the asymptotic validity of these confidence bands.

Proposition 1. *Suppose Assumptions M, K, and OSB or SSB hold true. Then* $\Pr\{f_X(t)\in\hat{\mathcal{C}}_X(t)$ *for all* $t\in\mathcal{T}\}\to 1-\tau$ *and* $\Pr\{f_\epsilon(t)\in\hat{\mathcal{C}}_\epsilon(t)$ *for all* $t\in\mathcal{T}\}\to 1-\tau$ *as* $n\to\infty$.

3.3. **Confidence bands for distribution functions.** Adusumilli *et al.* (2020, Theorem 2) proposed a bootstrap confidence band for the distribution function of X. Their theoretical development relies upon the uniform convergence rate in Kurisu and Otsu (2021), which restricts the growth rate of the subsample size to construct the bootstrap counterpart.

Based on the faster convergence rates obtained in Theorem 1, we can relax the requirements on the bandwidth in Adusumilli *et al.* (2020, Theorem 2). In particular, we can replace Assumptions OS' (v) and SS' (iv) in Adusumilli *et al.* (2020) by $n^{-\frac{1}{2}}(h^{-\gamma-\beta}+h^{-3}+h^{-\gamma-\frac{3}{2}})(\log h^{-1})^5\to 0$ and $n^{-\frac{1}{2}}\left(h^{\lambda_{0x}+\lambda-\delta_1}\exp(\frac{h^{-\lambda_x}}{\mu_x}+\frac{h^{-\lambda}}{\mu})+h^{\lambda_{0x}-\frac{3}{2}-\delta_1}\exp(\frac{h^{-\lambda_x}}{\mu_x})\right)(\log h^{-1})^{\frac{3}{2}}\to 0$ as $n\to\infty$, respectively, in their notations. These weaker conditions on the bandwidth in turn allow faster growth rates for the subsample size m in their notation.

Appendix A. Proofs

Notation. Hereafter, we use the following notation. For an arbitrary set T, let $\ell^\infty(T)$ denote the space of all bounded functions $T\to\mathbb{C}$, equipped with the uniform norm $\sup_{t\in T}|f(t)|$. For a probability measure Q on a measurable space $(S,\mathcal{S})$ and a class of measurable functions $\mathcal{F}$ on S such that $\mathcal{F}\subset L^2(Q)$, let $N(\mathcal{F},\|\cdot\|_{Q,2},\epsilon)$ denote the ϵ-covering number for $\mathcal{F}$ with respect to the $L^2(Q)$-seminorm $\|\cdot\|_{Q,2}$. See Section 2.1 in van der Vaart and Wellner (1996) for details. Let $\mathbb{G}_n(f)=\frac{1}{\sqrt{n}}\sum_{j=1}^n\{f(Y_{1,j},Y_{2,j})-E[f(Y_1,Y_2)]\}$ be the empirical process, and

$$\begin{aligned}\Delta(u) &= \log\left(\frac{\hat{\varphi}_X(u)}{\varphi_X(u)}\right)=\int_0^u\left(\frac{\hat{\psi}_1(0,u_2)}{\hat{\psi}(0,u_2)}-\frac{\psi_1(0,u_2)}{\psi(0,u_2)}\right)du_2.\\ R_1(u) &= \hat{\psi}_1(0,u)-\psi_1(0,u),\qquad R_2(u)=\frac{1}{\psi(0,u)}-\frac{1}{\hat{\psi}(0,u)}.\end{aligned}$$

We decompose $\Delta(u)$ as

$$\begin{aligned}\Delta(u) &= \int_0^u \frac{R_1(u_2)}{\psi(0,u_2)} du_2 + \int_0^u \psi_1(0,u_2) R_2(u_2) du_2 + \int_0^u R_1(u_2) R_2(u_2) du_2 \\ &:= \Delta_1(u) + \Delta_2(u) + \Delta_3(u). \end{aligned} \tag{10}$$

A.1. **Proof of Lemma 1.**

Proof of (i). **Step 1:** Linearization of $\hat{\varphi}_X(u) - \varphi_X(u)$.

Observe that

$$\begin{aligned}|\hat{\varphi}_X(u) - \varphi_X(u)| &= |\hat{\varphi}_X(u) - \varphi_X(u)| \mathbb{I}\{|\Delta(u)| \le 1\} + |\hat{\varphi}_X(u) - \varphi_X(u)| \mathbb{I}\{|\Delta(u)| > 1\} \\ &\le |\hat{\varphi}_X(u) - \varphi_X(u)| \mathbb{I}\{|\Delta(u)| \le 1\} + |\hat{\varphi}_X(u) - \varphi_X(u)||\Delta(u)| \mathbb{I}\{|\Delta(u)| > 1\} \\ &= |\varphi_X(u)||1 - e^{\Delta(u)}| \mathbb{I}\{|\Delta(u)| \le 1\} + |\hat{\varphi}_X(u) - \varphi_X(u)||\Delta(u)| \mathbb{I}\{|\Delta(u)| > 1\} \\ &\le 2|\varphi_X(u)||\Delta(u)| \mathbb{I}\{|\Delta(u)| \le 1\} + |\hat{\varphi}_X(u) - \varphi_X(u)||\Delta(u)| \mathbb{I}\{|\Delta(u)| > 1\} \\ &\le 2|\varphi_X(u)||\Delta(u)| + |\hat{\varphi}_X(u) - \varphi_X(u)||\Delta(u)|^p, \end{aligned} \tag{11}$$

for $p > 1$, where the first inequality follows from the fact that $|\hat{\varphi}_X(u) - \varphi_X(u)| \le 2$, the second equality follows from the definitions of $\hat{\varphi}_X(u)$ and $\Delta(u)$, and the second inequality follows from the fact that $|1 - e^z| \le 2|z|$ for $z \in \mathbb{C}$ with $|z| \le 1$. Note that we will show $\sup_{|u| \le h^{-1}} |\Delta(u)| = O_p(n^{-1/2} h^{-\beta_x - \beta_\epsilon} (\log n)^{1/2}) = o_p((\log n)^{-1})$ below. Then for sufficiently large $p > 1$, we have

$$\begin{aligned}\sup_{|u| \le h^{-1}} |\hat{\varphi}_X(u) - \varphi_X(u)||\Delta(u)|^p &\le \left(\sup_{|u| \le h^{-1}} |\hat{\varphi}_X(u) - \varphi_X(u)| \right) \left(\sup_{|u| \le h^{-1}} |\Delta(u)| \right)^p \\ &= \left(\sup_{|u| \le h^{-1}} |\hat{\varphi}_X(u) - \varphi_X(u)| \right) \times o_p((\log n)^{-p}), \end{aligned}$$

which implies that $(\hat{\varphi}_X(u) - \varphi_X(u)) \mathbb{I}\{|\Delta(u)| > 1\}$ does not contribute to the uniform convergence rate of $\hat{\varphi}_X$ and

$$\sup_{|u| \le h^{-1}} |\hat{\varphi}_X(u) - \varphi_X(u)| = O_p\left(\sup_{|u| \le h^{-1}} |\varphi_X(u)||\Delta(u)| \right) = O_p\left(\sum_{\ell=1}^3 \sup_{|u| \le h^{-1}} |\varphi_X(u)||\Delta_\ell(u)| \right).$$

Now we investigate stochastic orders of $|\varphi_X(u)||\Delta_\ell(u)|$ for $\ell = 1, 2, 3$. Define

$$\mathcal{G}_h = \left\{ g_u(\cdot) : (y_1, y_2) \mapsto \mathrm{i} h^{\beta_\epsilon} |\varphi_X(u)| \int_0^u \frac{y_1 e^{\mathrm{i} u_2 y_2}}{\psi(0,u_2)} du_2, u \in [-h^{-1}, h^{-1}] \right\}.$$

Then we can write as

$$n^{1/2} h^{\beta_\epsilon} |\varphi_X(u)| \Delta_1(u) = \mathbb{G}_n(f) \quad \text{for } f \in \mathcal{G}_h.$$

For any $g_{v_1}(\cdot), g_{v_2}(\cdot) \in \mathcal{G}_h$with $v_1, v_2 \in [-h^{-1}, h^{-1}]$, we can show that $|g_{v_1}(\cdot) - g_{v_2}(\cdot)| \lesssim |v_1 - v_2|$. Therefore, Andrews (1994, Theorem 2) implies that $\mathcal{G}_h$ is a Vapnik-Chervonenkis (VC) type class with envelope function $G_h(y_1, y_2) = D_0 h^{-1} |y_1|$ for some positive constant D_0, that is, there exist constants $A_1, v_1 > 0$ independent of n such that

$$\sup_Q N(\mathcal{G}_h, \|\cdot\|_{Q,2}, \epsilon \|G_h\|_{Q,2}) \le (A_1/\epsilon)^{v_1}, 0 < \forall \epsilon \le 1,$$

where $\sup_Q$ is taken over all finitely discrete distributions on $\mathbb{R}^2$. See also Pakes and Pollard (1989, Lemma 2.13). Furthermore, $\mathcal{G}_h$ satisfies Assumptions (A)-(C) in Chernozhukov, Chetverikov and Kato

(2016) with $B(f)=0$, $A\sim 1$, $v\sim 1$, $\sigma=b\sim h^{-1}$, and $K_n\sim\log n$. Let $\mathbb{U}_n$ be a tight Gaussian random variable in $\ell^\infty(\mathcal{G}_h)$. Then applying Chernozhukov, Chetverikov and Kato (2016, Theorem 2.1) with $q=4$ and $\gamma=1/\log n$ yields that there exists a random variable U_n with $U_n\overset{d}{=}\sup_{f\in\mathcal{G}_h}|\mathbb{U}_n(f)|$ such that

$$\left|\sup_{f\in\mathcal{G}_h}|\mathbb{G}_n(f)|-U_n\right|=O_p\left(\frac{(\log n)^{1+1/4}}{n^{1/4}h}+\frac{\log n}{n^{1/6}h}\right)=o_p((\log n)^{-1/2}). \tag{12}$$

Moreover, Dudley's entropy integral bound (van der Vaart and Wellner (1996, Corollary 2.2.8)) guarantees

$$E\left[\sup_{f\in\mathcal{G}_h}|\mathbb{U}_n(f)|\right]\lesssim\int_0^1\sqrt{1+\log(1/\epsilon h)}d\epsilon\lesssim(\log h^{-1})^{1/2}\lesssim(\log n)^{1/2}. \tag{13}$$

By (12) and (13), we have $\sup_{f\in\mathcal{G}_h}|\mathbb{G}_n(f)|=O_p(\sup_{f\in\mathcal{G}_h}|\mathbb{U}_n(f)|)=O_p((\log n)^{1/2})$, and thus

$$\sup_{|u|\le h^{-1}}|\varphi_X(u)||\Delta_1(u)|=n^{-1/2}h^{-\beta_\epsilon}\sup_{f\in\mathcal{G}_h}|\mathbb{G}_n(f)|=O_p(n^{-1/2}h^{-\beta_\epsilon}(\log n)^{1/2}). \tag{14}$$

Similarly, we can show that

$$\sup_{|u|\le h^{-1}}|\varphi_X(u)||\Delta_2(u)|=O_p(n^{-1/2}h^{-\beta_\epsilon+\delta}(\log h^{-1})^{1/2})=o_p(n^{-1/2}h^{-\beta_\epsilon}(\log n)^{-1/2}),$$
$$\sup_{|u|\le h^{-1}}|\varphi_X(u)||\Delta_3(u)|=O_p(n^{-1}h^{-\beta_x-2\beta_\epsilon}(\log h^{-1})^{1/2})=o_p(n^{-1/2}h^{-\beta_\epsilon}(\log n)^{-1/2}).$$

Combining these results, we obtain

$$\sup_{|u|\le h^{-1}}|\Delta(u)|=O_p(n^{-1/2}h^{-\beta_x-\beta_\epsilon}(\log n)^{1/2})=o_p((\log n)^{-1/2}).$$

Therefore,

$$\hat\varphi_X(u)-\varphi_X(u)=\varphi_X(u)\Delta_1(u)+o_p(n^{-1/2}h^{-\beta_\epsilon}(\log n)^{-1/2}), \tag{15}$$

uniformly on $u\in[-h^{-1},h^{-1}]$.

Step 2: Linearization of $\hat f_X(t)-f_X(t)$.

Let $\tilde f_X(t)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\varphi_K(hu)du$. Then (15) yields the following asymptotic linear representation of $\hat f_X(t)-\tilde f_X(t)$ uniformly on $t\in\mathcal{T}$:

$$\begin{aligned}&\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ux}\varphi_X(u)\Delta_1(u)\varphi_K(hu)du\\ =\;&\frac1n\sum_{j=1}^n\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ux}\varphi_X(u)\left\{\int_0^u\frac{Y_{1,j}e^{\mathrm{i}u_2Y_{2,j}}}{\psi(0,u_2)}du_2-\int_0^u\frac{E[Y_1e^{\mathrm{i}u_2Y_2}]}{\psi(0,u_2)}du_2\right\}\varphi_K(hu)du.\end{aligned} \tag{16}$$

Let

$$\mathcal{H}=\left\{(y_1,y_2)\mapsto h^{\beta_\epsilon+3/2}\frac{\mathrm{i}}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u\frac{y_1e^{\mathrm{i}u_2y_2}}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du:t\in\mathcal{T}\right\}.$$

A similar argument to show (14) yields

$$\sup_{t\in\mathcal{T}}|\hat f_X(t)-\tilde f_X(t)|=n^{-1/2}h^{-\beta_\epsilon-3/2}\sup_{f\in\mathcal{H}}|\mathbb{G}_n(f)|=O_p(n^{-1/2}h^{-\beta_\epsilon-3/2}(\log n)^{1/2}).$$

Note that

$$\sup_{t\in\mathcal{T}}|\hat f_X(t)-f_X(t)|\le\sup_{t\in\mathcal{T}}|\hat f_X(t)-\tilde f_X(t)|+\sup_{t\in\mathcal{T}}|\tilde f_X(t)-f_X(t)|,$$

and

$$\sup_{t\in\mathcal{T}}|\tilde{f}_X(t)-f_X(t)| \leq \frac{1}{2\pi}\int_{\mathbb{R}}|\varphi_X(u)||1-\varphi_K(hu)|du \lesssim h^{-1}\int_{[-1,1]^c}(u/h)^{-\beta_x}du \lesssim h^{\beta_x-1}.$$

Thus, the conclusion follows.

Proof of (ii). By the definition of $\hat{\varphi}_\epsilon$, we decompose

$$\begin{aligned}\hat{\varphi}_\epsilon(u)-\varphi_\epsilon(u) &= \frac{1}{\hat{\varphi}_X(u)}(\hat{\psi}(0,u)-\psi(0,u)) - \frac{\psi(0,u)}{\hat{\varphi}_X(u)}\left(\frac{\hat{\varphi}_X(u)-\varphi_X(u)}{\varphi_X(u)}\right)\\ &:= \Theta_1(u)-\Theta_2(u).\end{aligned}$$

From the results in Part (i) of this lemma, we can show that both $\sup_{|u|\leq h^{-1}}|\Theta_1(u)|$ and $\sup_{|u|\leq h^{-1}}|\Theta_2(u)|$ are of order $O_p(n^{-1/2}h^{-\beta_x}(\log n)^{1/2})$. From (15) in Part (i) of this lemma, the asymptotic linear representation of $\hat{\varphi}_X(u)-\varphi_X(u)$ is given by $\varphi_X(u)\Delta_1(u)$, and thus

$$\hat{\varphi}_\epsilon(u)-\varphi_\epsilon(u) = \frac{\hat{\psi}(0,u)-\psi(0,u)}{\varphi_X(u)} - \varphi_\epsilon(u)\Delta_1(u) + o_p(n^{-1/2}h^{-\beta_x}(\log n)^{-1/2}), \tag{17}$$

uniformly on $u\in[-h^{-1},h^{-1}]$. This implies

$$\begin{aligned}\hat{f}_\epsilon(t)-\tilde{f}_\epsilon(t) &= \frac{1}{n}\sum_{j=1}^n\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\frac{e^{\mathrm{i}uY_{2,j}}-E[e^{\mathrm{i}uY_2}]}{\varphi_X(u)}\varphi_K(hu)du\\ &\quad -\frac{1}{n}\sum_{j=1}^n\frac{\mathrm{i}}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_\epsilon(u)\left\{\int_0^u\left(\frac{Y_{1,j}e^{\mathrm{i}u_2Y_{2,j}}}{\psi(0,u_2)}-\frac{E[Y_1e^{\mathrm{i}u_2Y_2}]}{\psi(0,u_2)}\right)du_2\right\}\varphi_K(hu)du\\ &\quad +o_p(n^{-1/2}h^{-\beta_x-3/2}(\log n)^{-1/2}),\end{aligned}$$

uniformly on $t\in\mathcal{T}$, where $\tilde{f}_\epsilon(t)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_\epsilon(u)\varphi_K(hu)du$, and a similar argument to the proof of Part (i) of this lemma yields the conclusion.

A.2. **Proof of Lemma 2.** The proof is similar to that of Lemma 1. The only differences are: (i) the term $\sup_{|u|\leq h^{-1}}|\varphi_X(u)||\Delta_2(u)|$ will be dominant, and (ii) the bias term for $\sup_{t\in\mathcal{T}}|\tilde{f}_X(t)-f_X(t)|$ will be evaluated as in Kurisu and Otsu (2021).

A.3. **Proof of Theorem 3.** We only give the proof of the bootstrap approximation for $\hat{f}_X$ when f_X and f_ϵ are ordinary smooth (i.e. under Assumption OSB) since the proof of other cases ($\hat{f}_\epsilon$ under Assumption OSB, and $\hat{f}_X$ and $\hat{f}_\epsilon$ under Assumption SSB) are similar. The same comment applies to the proof of Proposition 1.

Proof of (i). Define $L_X(t)=\frac{1}{m}\sum_{j=1}^m\{L_{X,j}(t)-E[L_{X,1}(t)]\}$ for $t\in\mathcal{T}$.

Step 1: Gaussian approximation to L_X.

Letting

$$f_t(y_1,y_2)=y_1\frac{\mathrm{i}}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\left\{\int_0^u\frac{e^{\mathrm{i}u_2y_2}}{\psi(0,u_2)}du_2\right\}\varphi_K(hu)du,$$

it can be written as $\sqrt{m}\sup_{t\in\mathcal{T}}|L_X(t)/\sigma_{X,m}(t)|=\sup_{t\in\mathcal{T}}|\mathbb{G}_m(f_t)|$. For each $t_1,t_2\in\mathcal{T}$, we can show that

$$|f_{t_1}(y_1,y_2)-f_{t_2}(y_1,y_2)|\lesssim\frac{h^{-2}}{|\varphi_\epsilon(h^{-1})|}|y_1||t_1-t_2|,$$

for all y_1 and y_2. Therefore, by Andrews (1994, Theorem 2) and a similar argument to Step1 in the proof of Lemma 1 (i), $\tilde{\mathcal{F}}_n=\{f_t: t\in\mathcal{T}\}$ is a VC-type class with envelop function $F_h(y_1,y_2)=$

$Dh^{-2}|\varphi_\epsilon(h^{-1})|^{-1}|y_1|$ for a positive constant D. Let $\mathcal{F}_m = \{f_t/\sigma_{X,m}(t) : t \in \mathcal{T}\}$. Note that the set $\{1/\sigma_{X,m}(t) : t \in \mathcal{T}\}$ is bounded with $\sup_{t\in\mathcal{T}} |\sigma_{X,m}^{-1}(t)| \leq s_{X,m}^{-1}$. Then from Chernozhukov, Chetverikov and Kato (2014, Corollary A.1), there exist constants $A', v' > 0$ independent of n such that

$$\sup_Q N(\mathcal{F}_m, \|\cdot\|_{Q,2}, \epsilon D h^{-2} s_{X,m}^{-1}/|\varphi_\epsilon(h^{-1})|) \leq (A'/\epsilon)^{v'},$$

for all $0 < \epsilon \leq 1$. Furthermore, $\mathcal{F}_n$ satisfies Assumptions (A)-(C) in Chernozhukov, Chetverikov and Kato (2016) with $B(f) = 0$, $A = A'$, $v = v'$, $\sigma = 1$, $b = Dh^{-2}s_{X,m}^{-1}/|\varphi_\epsilon(h^{-1})|$, and $K_n \sim \log m$. Let $\mathbb{Z}_m$ be a tight Gaussian random variable in $\ell^\infty(\mathcal{F}_m)$ with mean zero and the same covariance function as $\mathbb{G}_m$. By applying Chernozhukov, Chetverikov and Kato (2016, Theorem 2.1) with $q = 4$ and $\gamma = 1/\log m$, there exists a random variable V_m with $V_m \stackrel{d}{=} \sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)|$ such that

$$\left|\sup_{t\in\mathcal{T}} |\mathbb{G}_m(f_t)| - V_m\right| = O_p\left(\frac{(\log m)^{5/4}}{m^{1/4}h^2 s_{X,m}|\varphi_\epsilon(h^{-1})|} + \frac{\log m}{m^{1/6}h^{2/3}s_{X,m}^{1/3}|\varphi_\epsilon(h^{-1})|^{1/3}}\right) = o_p((\log m)^{-1/2}).$$

Therefore, Chernozhukov, Chetverikov and Kato (2016, Lemma 2.1) guarantees

$$\begin{aligned} & \sup_{z\in\mathbb{R}} \left| \Pr\left\{\sup_{t\in\mathcal{T}} |\mathbb{G}_m(f_t)| \leq z\right\} - \Pr\left\{\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| \leq z\right\}\right| \\ \leq\ & \sup_{z\in\mathbb{R}} \Pr\left\{\left|\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| - z\right| \leq \delta_m (\log m)^{-1/2}\right\} + o(1), \end{aligned}$$

for some sequence $\delta_m \to 0$as $m \to \infty$. Now the anti-concentration inequality for the supremum of a Gaussian process yields

$$\sup_{z\in\mathbb{R},\delta>0} \frac{1}{\delta} \Pr\left\{\left|\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| - z\right| \leq \delta\right\} \lesssim E\left[\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)|\right] \lesssim (\log m)^{1/2}, \tag{18}$$

where the second inequality follows from Dudley's entropy integral bound. Combining these results, we obtain

$$\sup_{z\in\mathbb{R}} \left|\Pr\left\{\sqrt{m}\sup_{t\in\mathcal{T}} |L_X(t)/\sigma_{X,m}(t)| \leq z\right\} - \Pr\left\{\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| \leq z\right\}\right| \to 0. \tag{19}$$

Step 2: Approximate $\sqrt{m}\sup_{t\in\mathcal{T}} |\hat{L}_X^\xi(t)/\hat{\sigma}_{X,m}(t)|$ by $\sqrt{m}\sup_{t\in\mathcal{T}} |L_X^\xi(t)/\sigma_{X,m}(t)|$.

Define $L_X^\xi(t) = \frac{1}{m}\sum_{j=1}^m \xi_j \left\{L_{X,j}(t) - \frac{1}{m}\sum_{k=1}^m L_{X,k}(t)\right\}$ for $t \in \mathcal{T}$. In this step, we show

$$\begin{aligned} & \frac{1}{\sigma_{X,m}(t)} \sum_{j=1}^m \xi_j \left\{\hat{L}_{X,j}(t) - \frac{1}{m}\sum_{k=1}^m \hat{L}_{X,k}(t)\right\} \\ =\ & \frac{1}{\hat{\sigma}_{X,m}(t)} \sum_{j=1}^m \xi_j \left\{L_{X,j}(t) - \frac{1}{m}\sum_{k=1}^m L_{X,k}(t)\right\} + o_p(m^{1/2}(\log m)^{-1/2}), \end{aligned} \tag{20}$$

uniformly in $t \in \mathcal{T}$. Let

$$\begin{aligned} g_t &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u \frac{\hat{\psi}_1(0,u_2)}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du, \\ \hat{g}_t &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left(\int_0^u \frac{\hat{\psi}_1(0,u_2)}{\hat{\psi}(0,u_2)}du_2\right)\varphi_K(hu)du. \end{aligned}$$

Then we have

$$\sup_{t\in\mathcal{T}}\left|\frac{1}{\sigma_{X,m}(t)}\frac{1}{m}\sum_{j=1}^{m}\{\hat{L}_{X,j}(t)-L_{X,j}(t)\}\right|\leq\frac{1}{s_{X,m}}\sup_{t\in\mathcal{T}}|\hat{g}_t-g_t|.$$

Note that

$$\begin{aligned}
\hat{g}_t-g_t &= \left\{\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left(\int_0^u\frac{\psi_1(0,u_2)}{\hat{\psi}(0,u_2)}du_2\right)\varphi_K(hu)du\right.\\
&\quad\left.-\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u\frac{\psi_1(0,u_2)}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du\right\}\\
&\quad+\left\{\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left(\int_0^u\frac{\hat{\psi}_1(0,u_2)-\psi_1(0,u_2)}{\hat{\psi}(0,u_2)}du_2\right)\varphi_K(hu)du\right.\\
&\quad\left.-\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u\frac{\hat{\psi}_1(0,u_2)-\psi_1(0,u_2)}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du\right\}\\
&=: \mathbb{I}_n+\mathbb{III}_n.
\end{aligned}$$

Define $R_{\varphi_X}(u)=\hat{\varphi}_X(u)-\varphi_X(u)$, $R_\psi(u)=1/\hat{\psi}(0,u)-1/\psi(0,u)$, and $R'_\psi(u)=\hat{\psi}(0,u)-\psi(0,u)$. The term $\mathbb{I}_n$ can be further decomposed as

$$\begin{aligned}
\mathbb{I}_n &= \frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}R_{\varphi_X}(u)\left(\int_0^u\frac{\psi_1(0,u_2)}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du\\
&\quad+\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u\psi_1(0,u_2)R_\psi(u)du_2\right)\varphi_K(hu)du\\
&\quad+\frac{1}{2\pi}\int_{\mathbb{R}}e^{-\mathrm{i}ut}R_{\varphi_X}(u)\left(\int_0^u\psi_1(0,u_2)R_\psi(u)du_2\right)\varphi_K(hu)du\\
&=: \mathbb{I}_{1,n}+\mathbb{I}_{2,n}+\mathbb{I}_{3,n},
\end{aligned}$$

and these terms are bounded as

$$\begin{aligned}
\sup_{t\in\mathcal{T}}|\mathbb{I}_{1,n}| &\lesssim \sup_{|u|\leq h^{-1}}|R_{\varphi_X}(u)|\int_{-h^{-1}}^{h^{-1}}\left(\int_0^{|u|}|\frac{\psi_1(0,u_2)}{\psi(0,u_2)}|du_2\right)du\\
&= O_p(n^{-1/2}h^{\delta-2}|\varphi_\epsilon(h^{-1})|^{-1}(\log n)^{1/2})\\
\sup_{t\in\mathcal{T}}|\mathbb{I}_{2,n}| &\lesssim \sup_{|u|\leq h^{-1}}|R_\psi(u)|\int_{-h^{-1}}^{h^{-1}}|\varphi_X(u)|\left(\int_0^{|u|}|\frac{\psi_1(0,u_2)}{\psi^2(0,u_2)}|du_2\right)du\\
&= O_p(n^{-1/2}h^{\delta-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n)\\
\sup_{t\in\mathcal{T}}|\mathbb{I}_{3,n}| &\lesssim \sup_{|u|\leq h^{-1}}|R_{\varphi_X}(u)|\sup_{|u|\leq h^{-1}}|R'_\psi(u)|\int_{-h^{-1}}^{h^{-1}}\left(\int_0^{|u|}|\frac{\psi_1(0,u_2)}{\psi^2(0,u_2)}|du_2\right)du\\
&= O_p(n^{-1}h^{\delta-2}|\varphi_X(h^{-1})|^{-1}|\varphi_\epsilon(h^{-1})|^{-1}(\log n)^{3/2})=o_p(n^{-1/2}h^{\delta-2}(\log n)^{1/2}),
\end{aligned}$$

which implies $\sup_{t\in\mathcal{T}}|\mathbb{I}_n| = O_p(n^{-1/2}h^{\delta-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n)$. Likewise, we can show that $\sup_{t\in\mathcal{T}}|\mathbb{II}_n| = O_p(n^{-1/2}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n)$. Combining these results,

$$
\begin{aligned}
\sup_{t\in\mathcal{T}}\left|\frac{1}{\sigma_{X,m}(t)}\frac{1}{m}\sum_{j=1}^{m}\{\hat{L}_{X,j}(t)-L_{X,j}(t)\}\right| &\leq \frac{1}{s_{X,m}}\sup_{t\in\mathcal{T}}|\hat{g}_t-g_t| \leq \frac{1}{s_{X,m}}\left(\sup_{t\in\mathcal{T}}|\mathbb{I}_n|+\sup_{t\in\mathcal{T}}|\mathbb{II}_n|\right) \\
&= O_p(n^{-1/2}s_{X,m}^{-1}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n) \\
&= O_p\left(\left(\frac{m}{n}\right)^{1/2}m^{-1/2}s_{X,m}^{-1}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n\right),
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\sup_{t\in\mathcal{T}}\left|\left(\sum_{j=1}^{m}\xi_j\right)\frac{1}{\sigma_{X,m}(t)}\frac{1}{m}\sum_{j=1}^{m}\{\hat{L}_{X,j}(t)-L_{X,j}(t)\}\right| \\
&= O_p\left(\left(\frac{m}{n}\right)^{1/2}s_{X,m}^{-1}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\log n\right) = o_p(m^{1/2}(\log m)^{-1/2}).
\end{aligned}
$$

Now define

$$
\begin{aligned}
g_t(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\varphi_X(u)\left(\int_0^u \frac{e^{iu_2y}}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du, \\
\hat{g}_t(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\hat{\varphi}_X(u)\left(\int_0^u \frac{e^{iu_2y}}{\hat{\psi}(0,u_2)}du_2\right)\varphi_K(hu)du.
\end{aligned}
$$

We decompose

$$
\begin{aligned}
&\sum_{j=1}^{m}\xi_jY_{1,j}\{\hat{g}_t(Y_{2,j})-g_t(Y_{2,j})\} \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}R_{\varphi_X}(u)\left(\sum_{j=1}^{m}\xi_jY_{1,j}\int_0^u\frac{e^{iu_2Y_{2,j}}}{\psi(0,u_2)}du_2\right)\varphi_K(hu)du \\
&\quad +\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}\varphi_X(u)\left(\sum_{j=1}^{m}\xi_jY_{1,j}\int_0^u e^{iu_2Y_{2,j}}R_\psi(u)du_2\right)\varphi_K(hu)du \\
&\quad +\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ut}R_{\varphi_X}(u)\left(\sum_{j=1}^{m}\xi_jY_{1,j}\int_0^u e^{iu_2Y_{2,j}}R_\psi(u)du_2\right)\varphi_K(hu)du \\
&=: A_{1,n}+A_{2,n}+A_{3,n}.
\end{aligned}
$$

For $A_{1,n}$, the Cauchy-Schwarz inequality yields

$$
|A_{1,n}| \lesssim h^{-1}\left(\int|R_{\varphi_X}(u/h)|^2|\varphi_K(u)|du\right)^{1/2}\left(\int\left|\sum_{j=1}^{m}\xi_jY_{1,j}\int_0^{u/h}\frac{e^{iu_2Y_{2,j}}}{\psi(0,u_2)}du_2\right|^2|\varphi_K(u)|du\right)^{1/2}.
$$

Since

$$
\begin{aligned}
E\left[\left|\sum_{j=1}^{m}\xi_jY_{1,j}\int_0^{u/h}\frac{e^{iu_2Y_{2,j}}}{\psi(0,u_2)}du_2\right|^2\right] &= \sum_{j=1}^{m}E\left[\xi_1^2Y_{1,1}^2\left|\int_0^{u/h}\frac{e^{iu_2Y_{2,j}}}{\psi(0,u_2)}du_2\right|^2\right] \\
&\lesssim mh^{-2}|\varphi_X(h^{-1})|^{-2}|\varphi_\epsilon(h^{-1})|^{-2},
\end{aligned}
$$

we obtain

$$
\begin{aligned}
|A_{1,n}| &\lesssim h^{-1}O_p(n^{-1/2}|\varphi_\epsilon(h^{-1})|^{-1}(\log n)^{1/2}) \times O_p(m^{1/2}h^{-1}|\varphi_X(h^{-1})|^{-1}|\varphi_\epsilon(h^{-1})|^{-1}) \\
&= O_p\left(\left(\frac{m}{n}\right)^{1/2} m^{1/2}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\left(\frac{(\log n)^{1/2}}{\log m}\right)\right).
\end{aligned}
$$

Similarly, for $A_{2,n}$, we have

$$
\begin{aligned}
|A_{2,n}| &\lesssim h^{-1}\left(\int|\varphi_X(u/h)||\varphi_K(u)|du\right)^{1/2} \\
&\quad \times\left(\int|\varphi_X(u/h)|\left|\int_0^{u/h}\sum_{j=1}^n \xi_j Y_{1,j}e^{iu_2Y_{2,j}}R_\psi(u_2)du_2\right|^2|\varphi_K(u)|du\right)^{1/2} \\
&\lesssim h^{-1}\left(\int|\varphi_X(u/h)|\left(\int_0^{|u|/h}\left|\sum_{j=1}^n \xi_j Y_{1,j}e^{iu_2Y_{2,j}}\right|^2 du_2\right)^{1/2}\right. \\
&\quad \left.\times\left(\int_0^{|u|/h}|R_\psi(u_2)|^2du_2\right)^{1/2}|\varphi_K(u)|du\right).
\end{aligned}
$$

Since

$$
E\left[\left|\sum_{j=1}^m \xi_j Y_{1,j}e^{iu_2Y_{2,j}}\right|^2\right] = \sum_{j=1}^m E[\xi_1^2Y_{1,1}^2] \lesssim m,
$$

we obtain

$$
\begin{aligned}
|A_{2,n}| &\lesssim h^{-1}O_p(m^{1/2}h^{-1/2}) \times O_p(n^{-1/2}h^{-1/2}|\varphi_X(h^{-1})|^{-1}|\varphi_\epsilon(h^{-1})|^{-2}(\log n)) \\
&= O_p\left(\left(\frac{m}{n}\right)^{1/2} m^{1/2}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\left(\frac{\log n}{\log m}\right)\right).
\end{aligned}
$$

Likewise, for $A_{3,n}$, it holds

$$
|A_{3,n}| = O_p\left(\left(\frac{m}{n}\right) m^{1/2}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\left(\frac{(\log n)^{3/2}}{(\log m)^2}\right)\right).
$$

Combining these results,

$$
\begin{aligned}
\sup_{t\in\mathcal{T}}\left|\sigma_{X,m}^{-1}(t)\sum_{j=1}^m \xi_j Y_{1,j}\{\hat{g}_t(Y_{2,j})-g_t(Y_{2,j})\}\right| &\leq s_{X,m}^{-1}\left(\sup_{t\in\mathcal{T}}|A_{1,n}|+\sup_{t\in\mathcal{T}}|A_{2,n}|+\sup_{t\in\mathcal{T}}|A_{2,n}|\right) \\
&= O_p\left(\left(\frac{m}{n}\right)^{1/2} m^{1/2}s_{X,m}^{-1}h^{-2}|\varphi_\epsilon(h^{-1})|^{-1}\left(\frac{\log n}{\log m}\right)\right).
\end{aligned}
$$

Since $\sup_{t\in\mathcal{T}}|\hat{\sigma}_{X,m}(t)/\sigma_{X,m}(t)-1| = o_p((\log m)^{-1/2})$, we obtain (20).

Step 3: Conditional approximation of $\sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)|$ by $\sqrt{m}\sup_{t\in\mathcal{T}}|L_X^\xi(t)/\sigma_{X,m}(t)|$.

By applying Chernozhukov, Chetverikov and Kato (2016, Theorem 2.2) with $q=4$ and $\gamma = 1/\log m$, there exists a random variable V_m^ξ with $V_m^\xi|\mathcal{Y}_n \stackrel{d}{=} \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)|$ such that

$$
\left|\sqrt{m}\sup_{t\in\mathcal{T}}|L_X^\xi(t)/\sigma_{X,m}(t)| - V_m^\xi\right| = o_p((\log m)^{-1/2}).
$$

Therefore, there exists a sequence $\delta_m \to 0$ such that

$$\begin{aligned}
\Pr\left\{ \sqrt{m}\sup_{t\in\mathcal{T}}|L_X^{\xi}(t)/\sigma_{X,m}(t)| \le z \,\middle|\, \mathcal{Y}_n \right\} &= \Pr\left\{ V_m^{\xi} \le z + \delta_m(\log m)^{-1/2} \,\middle|\, \mathcal{Y}_n \right\} + o_p(1) \\
&= \Pr\left\{ \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z + \delta_m(\log m)^{-1/2} \right\} + o_p(1) \\
&\le \Pr\left\{ \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z \right\} + o_p(1),
\end{aligned}$$

uniformly in $z \in \mathbb{R}$, where the inequality follows from (18). Similarly, we can show that $\Pr\left\{ \sqrt{m}\sup_{t\in\mathcal{T}}|L_X^{\xi}(t)/\sigma_{X,m}(t)| \le z \,\middle|\, \mathcal{Y}_n \right\} \ge \Pr\left\{\sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z\right\} - o_p(1)$, and thus

$$\sup_{z\in\mathbb{R}}\left| \Pr\left\{ \sqrt{m}\sup_{t\in\mathcal{T}}|L_X^{\xi}(t)/\sigma_{X,m}(t)| \le z \,\middle|\, \mathcal{Y}_n \right\} - \Pr\left\{ \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z \right\} \right| = o_p(1). \tag{21}$$

Step 4: Proof of the theorem and asymptotic validity of the uniform confidence bands.

Observe that

$$\begin{aligned}
\sqrt{m}(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t) &= \sqrt{m}(1 + o_p((\log m)^{-1}))(\hat{f}_X(t) - f_X(t))/\sigma_{X,m}(t) \\
&= (1 + o_p((\log m)^{-1}))(\sqrt{m}L_X(t)/\sigma_{X,m}(t) + o_p((\log m)^{-1/2})) \\
&= \sqrt{m}L_X(t)/\sigma_{X,m}(t) + o_p((\log m)^{-1/2})
\end{aligned}$$

uniformly on $t \in \mathcal{T}$. Combining this and the result of Step 1 in the proof of 3, and using the anti-consentration inequality, we can show that

$$\sup_{z\in\mathbb{R}}\left| \Pr\left\{ \sqrt{m}\sup_{t\in\mathcal{T}}|(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le z \right\} - \Pr\left\{ \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z \right\} \right| \to 0.$$

Therefore, the conclusion follows from (19) and (21).

Proof of (ii). The proof is similar to Part (i) of this theorem.

A.4. **Proof of Proposition 1.** We wish to show that $\Pr\{f_X(t) \in \hat{C}_X \ \forall t \in \mathcal{T}\} \to 1 - \tau$. Note that

$$f_X(t) \in \hat{C}_X \ \forall x \in \mathcal{T} \Leftrightarrow \sup_{t\in\mathcal{T}}|(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le \hat{c}_X^{1-\tau}.$$

Together with the result in Step 1 in the proof of Theorem 3 and $E[\sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)|] \lesssim (\log m)^{1/2}$, we have $\sqrt{m}\sup_{t\in\mathcal{T}}|L_X(t)/\sigma_{X,m}(t)| = O_p((\log m)^{1/2})$. Observe that

$$\begin{aligned}
\sqrt{m}(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t) &= \sqrt{m}(1 + o_p((\log m)^{-1}))(\hat{f}_X(t) - f_X(t))/\sigma_{X,m}(t) \\
&= (1 + o_p((\log m)^{-1}))(\sqrt{m}L_X(t)/\sigma_{X,m}(t) + o_p((\log m)^{-1/2})) \\
&= \sqrt{m}L_X(t)/\sigma_{X,m}(t) + o_p((\log m)^{-1/2}),
\end{aligned}$$

uniformly on $t \in \mathcal{T}$. The result of Step 4 in the proof of Theorem 3 implies that there exists a sequence of constants $\epsilon_{n,1} \to 0$ such that

$$\sup_{z\in\mathbb{R}}\left| \Pr\left\{ \sqrt{m}\sup_{t\in\mathcal{T}}|(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le z \right\} - \Pr\left\{ \sup_{t\in\mathcal{T}}|\mathbb{Z}_m(f_t)| \le z \right\} \right| \le \epsilon_{n,1}.$$

Moreover, the results of Steps 2 and 3 in the proof of Theorem 3 yields that there exists a sequence of constants $\epsilon_{n,2} \to 0$ such that

$$\sup_{z\in\mathbb{R}} \left| \Pr\left\{ \sqrt{m} \sup_{t\in\mathcal{T}} |\hat{L}_X^{\xi}(t)/\hat{\sigma}_{X,m}(t)| \le z \,\middle|\, \mathcal{Y}_n \right\} - \Pr\left\{ \sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| \le z \right\} \right| \le \epsilon_{n,2}.$$

Let Ω_n denote the event on which these inequalities hold and let $c(1-\tau)$ denote the $(1-\tau)$-th quantile of $\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)|$. Note that $\Pr\{\Omega_n\} \to 1$ as $n \to \infty$. Define $\epsilon_n' = \epsilon_{n,1} \vee \epsilon_{n,2} (\to 0)$. Then on Ω_n, we have

$$\Pr\left\{ \sqrt{m} \sup_{t\in\mathcal{T}} |\hat{L}_X^{\xi}(t)/\hat{\sigma}_{X,m}(t)| \le c(1-\tau+\epsilon_n') \,\middle|\, \mathcal{Y}_n \right\} \ge \Pr\left\{ \sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| \le c(1-\tau+\epsilon_n') \right\} - \epsilon_n' = 1-\tau.$$

We used the continuity of the distribution of $\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)|$ to obtain the last equation (this follows from the anti-concentration inequality). This yields that on Ω_n,

$$\hat{c}_X^{1-\tau} \le c(1-\tau+\epsilon_n').$$

Likewise, we can show that $c(1-\tau-\epsilon_n') \le \hat{c}_X^{1-\tau}$ on Ω_n. Then we have

$$\begin{aligned}
& \Pr\left\{ \sqrt{m} \sup_{t\in\mathcal{T}} |(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le \hat{c}_X^{1-\tau} \right\} \\
\le\ & \Pr\left\{ \sqrt{m} \sup_{t\in\mathcal{T}} |(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le c(1-\tau+\epsilon_n') \right\} + o(1) \\
=\ & \Pr\left\{ \sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)| \le c(1-\tau+\epsilon_n') \right\} + o(1) = 1-\tau+\epsilon_n' + o(1) = 1-\tau+o(1).
\end{aligned}$$

To obtain the third equation, we used the continuity of the distribution of $\sup_{t\in\mathcal{T}} |\mathbb{Z}_m(f_t)|$. Likewise, we have

$$\Pr\left\{ \sqrt{m} \sup_{t\in\mathcal{T}} |(\hat{f}_X(t) - f_X(t))/\hat{\sigma}_{X,m}(t)| \le \hat{c}_X^{1-\tau} \right\} \ge 1-\tau-o(1).$$

Therefore, the conclusion follows.

References

[1] Adusumilli, K., Kurisu, D., Otsu, T and Y. -J. Whang (2020) Inference on distribution functions under measurement error, *Journal of Econometrics*, 215, 131-164.

[2] Andrews, D. W. K. (1994) Empirical process methods in econometrics, *Handbook of Econometrics*, 4, 2247-2294.

[3] Bissantz, N., Dümbgen, L., Holzmann, H. and A. Munk (2007) Non-parametric confidence bands in deconvolution density estimation, *Journal of the Royal Statistical Society*, B 69, 483-506.

[4] Bonhomme, S. and J.-M. Robin (2010) Generalized non-parametric deconvolution with an application to earnings dynamics, *Review of Economic Studies*, 77, 491-533.

[5] Chernozhukov, V., Chetverikov, D. and K. Kato (2014) Gaussian approximation of suprema of empirical processes, *Annals of Statistics*, 42, 1564-1597.

[6] Chernozhukov, V., Chetverikov, D. and K. Kato (2016) Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related Gaussian couplings. *Stochastic Processes and their Applications*, 126, 3632-3651.

[7] Comte, F. and J. Kappus (2015) Density deconvolution from repeated measurement without symmetry assumption on the errors, *Journal of Multivariate Analysis*, 140, 31-46.

[8] Delaigle, A., Hall, P. and A. Meister (2008) On deconvolution with repeated measurements, *Annals of Statistics*, 36, 665-685.

[9] Kato, K. and Y. Sasaki (2018) Uniform confidence bands in deconvolution with unknown error distribution, *Journal of Econometrics*, 207, 129-161.

[10] Kato, K. and Y. Sasaki (2019) Uniform confidence bands for nonparametric errors-in-variables regression, *Journal of Econometrics*, 213, 516-555.

[11] Kato, K., Sasaki, Y. and T. Ura (2021) Robust inference in deconvolution, *Quantitative Economics*, 12, 109-142.

[12] Kotlarski, I. (1967) On characterizing the gamma and the normal distribution, *Pacific Journal of Mathematics*, 20, 69-76.

[13] Krasnokutskaya, E. (2011) Idenfication and estimation of auction models with unobserved heterogeneity, *Review of Economic Studies*, 78, 293-327.

[14] Kurisu, D. and T. Otsu (2021) On the uniform convergence of deconvolution estimators from repeated measurements, forthcoming in *Econometric Theory*.

[15] Li, T., Perrigne I. and Q. Vuong (2000) Conditionally independent private information in OCS wildcat auctions, *Journal of Econometrics*, 98, 129-161.

[16] Li, T. and Q. Vuong (1998) Nonparametric estimation of the measurement error model using multiple indicators, *Journal of Multivariate Analysis*, 65, 139-165.

[17] McMurry, T. L. and D. N. Politis (2004) Nonparametric regression with infinite order flat-top kernels, *Journal of Nonparametric Statistics*, 16, 549-562.

[18] Meister, A. (2009) *Deconvolution Problems in Nonparametric Statistics*, Springer.

[19] Neumann, M. H. (2007) Deconvolution from panel data with unknown error distribution, *Journal of Multivariate Analysis*, 98, 1955-1968.

[20] Pakes, A. and D. Pollard (1989) Simulation and the asymptotics of optimization estimators, *Econometrica*, 57, 1027-1057.

[21] Schennach, S. M. (2004) Nonparametric regression in the presence of measurement error. *Econometric Theory*. 20, 1046-1093.

[22] van der Vaart, A. W. and J. A. Wellner (1996) *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer.

Department of Industrial Engineering and Economics, School of Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan.

Email address: `kurisu.d.aa@m.titech.ac.jp`

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.

Email address: `t.otsu@lse.ac.uk`