ON THE UNIFORM CONVERGENCE OF DECONVOLUTION ESTIMATORS FROM REPEATED MEASUREMENTS

DAISUKE KURISU AND TAISUKE OTSU

Abstract. This paper studies the uniform convergence rates of Li and Vuong's (1998) nonparametric deconvolution estimator and its regularized version by Comte and Kappus (2015) for the classical measurement error model, where repeated noisy measurements on the error-free variable of interest are available. In contrast to Li and Vuong (1998), our assumptions allow unbounded supports for the error-free variable and measurement errors. Compared to Bonhomme and Robin (2010) specialized to the measurement error model, our assumptions do not require existence of the moment generating functions of the square and product of repeated measurements. Furthermore, by utilizing a maximal inequality for the multivariate normalized empirical characteristic function process, we derive the uniform convergence rates that are faster than the ones derived in these papers under such weaker conditions.

1. Introduction

This paper studies uniform convergence rates 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, based on Kotlarski's (1967) identity, a seminal work by Li and Vuong (1998, hereafter LV) developed a novel nonparametric estimator for the densities of the error-free variable of interest and the measurement errors. An 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 (Delaigle, Hall and Meister, 2008). The LV estimator has been applied in various contexts in econometrics, such as nonlinear errors-in-variables models (Li, 2002), panel data models (Evdokimov, 2010, Arellano and Bonhomme, 2012), generalized linear models (Li and Hsiao, 2004), auctions (e.g., Krasnokutskaya, 2011, and Athey and Haile, 2007, for a survey), identification of private information (Arcidiacono *et al.*, 2011), among others. See also Hu (2017) for a survey on various applications of measurement error models in economics.

In addition, for these econometric and statistical problems, the deconvolution estimators may not necessarily be the ultimate objects of interest, and may be intermediate objects to be plugged-in to obtain final estimators or test statistics. For example, Krasnokutskaya (2011) developed nonparametric estimators for individual bid functions and cost components in auction models with unobserved heterogeneity as certain functionals of the LV-type estimators. The semiparametric estimators by Li (2002) and Li and Hsiao (2004) are constructed as functionals of the LV estimator. Also, other nonparametric measurement error problems often call for estimation of the characteristic function of the measurement error, such as Adusumilli and Otsu (2018)

The authors would like to thank anonymous referees for helpful comments. Our research is supported by JSPS KAKENHI (JP17H02513, JP19K20881, JP20K13468) (Kurisu) and ERC Consolidator Grant (SNP 615882) (Otsu).

for nonparametric instrumental regression with errors-in-variables, Otsu and Taylor (2019) for specification testing on errors-in-variables regressions, and Adusumilli *et al.* (2020) for inference on distribution functions under measurement errors. For those purposes, the LV-type estimators play the same roles as primitive nonparametric estimators for semiparametric problems. Thus, it is crucial to establish uniform convergence rates for the LV-type estimators under widely applicable and mild conditions; this is the theme of the present paper.

In this paper, we derive uniform convergence rates for the LV estimator and its regularized version proposed by Comte and Kappus (2015, hereafter CK). CK modified the LV estimator by introducing a regularization factor to deal with small denominators and truncation to restrict the estimated characteristic function not to take values larger than one. They also established the L_2-convergence rates under weaker assumptions than the ones in LV. Importantly, CK dropped the bounded support conditions by LV on both the error-free variable of interest and the measurement errors. In contrast, we study uniform convergence rates and show that both the LV and CK estimators typically achieve faster uniform convergence rates under weaker assumptions (especially unbounded support).

In another important paper, Bonhomme and Robin (2010) considered a general latent multi-factor model, which includes the repeated measurements model as a special case, and established the uniform convergence rate for their nonparametric deconvolution estimator without assuming bounded support. Our convergence rates are faster than those given in Bonhomme and Robin (2010) under weaker assumptions. In particular, we do not require existence of the moment generating functions of the square and product of repeated measurements as in Bonhomme and Robin (2010). The relaxation of this assumption is achieved by showing a maximal inequality for the multivariate normalized empirical characteristic function process (Lemma 1 in Section 2.2), which is a multivariate version of Neumann and Reiss (2009, Theorem 4.1). This lemma may also be used in Bonhomme and Robin (2010) to relax their assumptions in other contexts, and thus is of independent interest.

The results of this paper are useful not only for extending the scope of empirical analysis for econometric objects identified by Kotlarski's identify, but also for addressing open questions on existing methods that involve the LV-type estimators. Although detailed analyses are beyond the scope of this paper, we mention such possibilities in our concluding remarks.

This paper is organized as follows. Section 2 presents our main results, the uniform convergence rates of the LV and CK estimators. Section 2.1 collects remarks on the main theorems, and Section 2.2 presents the maximal inequality for the multivariate normalized empirical characteristic function process. Section 3 concludes with some potential applications of our main results.

2. Main result

Consider a bivariate i.i.d. sample $\{Y_{j,1}, Y_{j,2}\}_{j=1}^n$ of (Y_1, Y_2), which is generated by

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

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.[1] We are interested in estimating the densities of X, ϵ_1, and ϵ_2.

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 $\varphi_X(\cdot) = E[e^{\mathrm{i}\cdot X}]$ and $\varphi_\epsilon(\cdot) = E[e^{\mathrm{i}\cdot\epsilon}]$ vanish nowhere, and $E[\epsilon] = 0$. $E[|Y_1|^{2+\eta}] < \infty$ for some $\eta > 0$.*

These assumptions are standard for the classical measurement error model (see, e.g., CK). The condition $E[\epsilon] = 0$ is considered as a normalization to identify the mean of X. However, they are weaker than other existing papers on the repeated measurements model, such as LV (which impose bounded support of f_X and f_ϵ), and Bonhomme and Robin (2010) (which require the existence of the moment generating functions of Y_1^2 and Y_1Y_2). See Remark 2 and Section 2.2 for detailed discussions.

This paper studies the uniform convergence rates of the LV and CK estimators for the densities and characteristic functions of X and ϵ. Let us first introduce the LV estimator. Define

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

Under the condition $E|Y_1| < \infty$, Kotlarski's identity gives us an explicit identification formula of φ_X, that is

$$\varphi_X(u) = \exp \int_0^u \frac{\partial\psi(0, u_2)/\partial u_1}{\psi(0, u_2)} du_2.$$

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

$$\hat{\varphi}_X(u) = \exp \int_0^u \frac{\partial\hat{\psi}(0, u_2)/\partial u_1}{\hat{\psi}(0, u_2)} du_2, \tag{2}$$

where $\hat{\psi}(u_1, u_2) = \frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}(u_1 Y_{j,1} + u_2 Y_{j,2})}$ and $\frac{\partial\hat{\psi}(u_1,u_2)}{\partial u_1} = \frac{1}{n}\sum_{j=1}^n \mathrm{i}Y_{j,1} e^{\mathrm{i}(u_1 Y_{j,1} + u_2 Y_{j,2})}$. Based on this estimator, the density f_X of X can be estimated by

$$\hat{f}_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ux}\hat{\varphi}_X(u)\varphi_K(hu)du, \tag{3}$$

where $\varphi_K(u) = \int_{\mathbb{R}} e^{\mathrm{i}ux}K(x)dx$ is the Fourier transform of a kernel function K and $h = h_n$ is a sequence of positive numbers (bandwidths) such that $h_n \to 0$ as $n \to \infty$.

Based on the expression $\varphi_\epsilon(u) = \psi(0, u)/\varphi_X(u)$, the characteristic function φ_ϵ of ϵ can also be estimated by

$$\hat{\varphi}_\epsilon(u) = \frac{\hat{\psi}(0, u)}{\hat{\varphi}_X(u)}. \tag{4}$$

The estimator $\hat{f}_\epsilon$ of the density f_ϵ is given by replacing $\hat{\varphi}_X$ in (3) with $\hat{\varphi}_\epsilon$.

[1]It is possible to extend to the case where more than two noisy measurements on X are available. However, for sake of simplicity and clarity, we concentrate on the two dimensional case.

We next introduce a regularized version of the LV estimator developed by CK. Their main idea is to regularize $\hat{\varphi}_X$ in (2) as

$$\tilde{\varphi}_X(u) = \frac{\tilde{\varphi}_X^{mod}(u)}{\max\{1, |\tilde{\varphi}_X^{mod}(u)|\}}, \tag{5}$$

where

$$\tilde{\varphi}_X^{mod}(u) = \exp \int_0^u \frac{\partial\hat{\psi}(0,u_2)/\partial u_1}{\tilde{\psi}(0,u_2)} du_2, \quad \text{with} \quad \tilde{\psi}(0,u_2) = \frac{\hat{\psi}(0,u_2)}{\min\{1, \sqrt{n}|\hat{\psi}(0,u_2)|\}}.$$

There are two differences between $\hat{\varphi}_X$ and $\tilde{\varphi}_X$. First, the reciprocal $1/\psi(0,u_2)$ is estimated by $1/\tilde{\psi}(0,u_2)$ instead of the empirical average $1/\hat{\psi}(0,u_2)$. The additional term, $\min\{1, \sqrt{n}|\hat{\psi}(0,u_2)|\}$, circumvents unfavorable effects caused by small values of the denominator. Second, the denominator of $\tilde{\varphi}_X$ in (5) is introduced to improve the quality of the estimator by imposing that the estimand is a characteristic function, which should not take values larger than one.

Based on this regularized estimator $\tilde{\varphi}_X$, the CK estimator $\tilde{f}_X$ of the density f_X is defined by replacing $\hat{\varphi}_X$ in (3) with $\tilde{\varphi}_X$. Also, the characteristic function φ_ϵ of ϵ can be estimated by

$$\tilde{\varphi}_\epsilon(u) = \frac{\hat{\psi}(0,u)}{\bar{\varphi}_X(u)}, \quad \text{where} \quad \bar{\varphi}_X(u) = \frac{\tilde{\varphi}_X(u)}{\min\{1, \sqrt{n}|\tilde{\varphi}_X(u)|\}}.$$

The estimator $\tilde{f}_\epsilon$ of the density f_ϵ is given by replacing $\hat{\varphi}_X$ in (3) with $\tilde{\varphi}_\epsilon$.

For these regularized estimators, CK investigated the risk bounds and convergence rates for the L_2-loss function. In this paper, we study the uniform convergence rates of the CK estimators.

To estimate the densities by (3), we need to choose the kernel function K, and impose the following conditions.

Assumption K. *(i) 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, $\varphi_K(u) = 0$ for any $|u| > 1$. (ii) $p \geq \max\{\beta_x, \beta_\epsilon\}$. (iii) There exists $0 < c \leq 1$ such that $\varphi_K(x) = 1$ for $|x| \leq c$.*

Assumption K (i) says that K is a p-th order kernel function. See e.g., Tsybakov (2009, Section 1.2.2) for a construction of higher order kernels. Assumptions K (ii) and (iii) are used for the ordinary and super smooth cases, respectively.

To proceed, we adopt the terminology in Fan (1991) and consider two scenarios for the densities f_X and f_ϵ, called ordinary smooth and super smooth densities. In particular, we impose the following assumptions on the characteristic functions φ_X and φ_ϵ.

Assumption OS. *For some positive constants $\beta_x > 1$, $C_x \geq c_x$, ω_x, $\beta_\epsilon > 1$, $C_\epsilon \geq c_\epsilon$, and ω_ϵ, it holds*

$$c_x|u|^{-\beta_x} \leq |\varphi_X(u)| \leq C_x|u|^{-\beta_x} \quad \textit{for all } |u| \geq \omega_x,$$
$$c_\epsilon|u|^{-\beta_\epsilon} \leq |\varphi_\epsilon(u)| \leq C_\epsilon|u|^{-\beta_\epsilon} \quad \textit{for all } |u| \geq \omega_\epsilon.$$

In this case, f_X and f_ϵ are called ordinary smooth. The conditions $\beta_x, \beta_\epsilon > 1$ are introduced to guarantee the consistency of the density estimators. Since the estimators of the characteristic functions 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.

Assumption SS. *For some positive constants ρ_x, $C_x \geq c_x$, ω_x, μ_x, ρ_ϵ, $C_\epsilon \geq c_\epsilon$, ω_ϵ, μ_ϵ, and some constants $\beta_x, \beta_\epsilon \in \mathbb{R}$, it holds*

$$c_x|u|^{\beta_x}\exp(-|u|^{\rho_x}/\mu_x) \leq |\varphi_X(u)| \leq C_x|u|^{\beta_x}\exp(-|u|^{\rho_x}/\mu_x), \quad \text{for all } |u| \geq \omega_x.$$
$$c_\epsilon|u|^{\beta_\epsilon}\exp(-|u|^{\rho_\epsilon}/\mu_\epsilon) \leq |\varphi_\epsilon(u)| \leq C_\epsilon|u|^{\beta_\epsilon}\exp(-|u|^{\rho_\epsilon}/\mu_\epsilon), \quad \text{for all } |u| \geq \omega_\epsilon.$$

In this case, f_X and f_ϵ are called super smooth. Similar to Assumption OS, we use the lower and upper bounds to control estimation errors. A popular example of a super smooth density is the normal density.

Define the maximal deviations

$$\mathcal{D}_n^{\varphi,a} = \sup_{u\in[-T_n,T_n]} |\hat{\varphi}_a(u) - \varphi_a(u)|, \quad \text{and} \quad \mathcal{D}_n^{f,a} = \sup_{|x|\leq h^{-1}} |\hat{f}_a(u) - f_a(u)|,$$

for $a = X$ and ϵ. Under the above assumptions, the convergence rates of these maximal deviations are presented as follows. The proofs are presented in Appendices A-B.

First, we consider the case where both f_X and f_ϵ are ordinary smooth. Let $\varrho_T^{\text{o}} = T^{2\beta_x+2\beta_\epsilon+1}\log T$.

Theorem 1. *[OS f_X and OS f_ϵ] Suppose that Assumptions M and OS hold. Then*

$$\mathcal{D}_n^{\varphi,X} = O_p(n^{-1/2}\varrho_{T_n}^{\text{o}}) \quad \text{under } n^{-1/2}\varrho_{T_n}^{\text{o}} \to 0, \tag{6}$$
$$\mathcal{D}_n^{\varphi,\epsilon} = O_p(n^{-1/2}T_n^{\beta_x}\varrho_{T_n}^{\text{o}}) \quad \text{under } n^{-1/2}T_n^{\beta_x}\varrho_{T_n}^{\text{o}} \to 0. \tag{7}$$

Additionally suppose that Assumptions K (i)-(ii) hold. Then

$$\mathcal{D}_n^{f,X} = O_p(n^{-1/2}h^{-1}\varrho_{h^{-1}}^{\text{o}} + h^{\beta_x-1}) \quad \text{under } n^{-1/2}\varrho_{h^{-1}}^{\text{o}} \to 0, \tag{8}$$
$$\mathcal{D}_n^{f,\epsilon} = O_p(n^{-1/2}h^{-\beta_x-1}\varrho_{h^{-1}}^{\text{o}} + h^{\beta_\epsilon-1}) \quad \text{under } n^{-1/2}h^{-\beta_x}\varrho_{h^{-1}}^{\text{o}} \to 0. \tag{9}$$

See Section 2.1 below for detailed comparisons with existing results. Equations (6) and (7) characterize the uniform convergence rates of the characteristic function estimators $\hat{\varphi}_X$ and $\hat{\varphi}_\epsilon$, respectively. When $T_n \to \infty$, the convergence rate of $\hat{\varphi}_\epsilon$ is slower than that of $\hat{\varphi}_X$ due to the additional factor "$T_n^{\beta_x}$" in the right hand side of (7). Intuitively, this additional factor emerges from the linearization coefficient of (4) around $\hat{\varphi}_X(u) = \varphi_X(u)$. On the other hand, equations (8) and (9) characterize the convergence rates of the LV-type density estimators $\hat{f}_X$ and $\hat{f}_\epsilon$, respectively. The first terms in the right hand sides of (8) and (9) are orders of stochastic errors, and the second terms represent the bias terms. Compared to the first term of (6) (with setting $T_n = h^{-1}$), the first term of (8) contains an additional factor "h^{-1}", which is due to the regularization for the inverse Fourier transform (see (16) in Appendix), and thus converges slower. The same remark applies to the first terms of (7) and (9).

Second, we present the results for the case where both f_X and f_ϵ are super smooth. Let $\varrho_T^{\mathrm{s}} = T^{-2\beta_x-2\beta_\epsilon+1}(\log T)\exp\left(\frac{2T^{\rho_x}}{\mu_x}+\frac{2T^{\rho_\epsilon}}{\mu_\epsilon}\right)$,

$$\varsigma_{h,q}^x = h^{\frac{\rho_x}{q}-\beta_x-1}\exp\left(-\frac{c^{\rho_x}h^{-\rho_x}}{\mu_x}\right), \quad \text{and} \quad \varsigma_{h,q}^\epsilon = h^{\frac{\rho_\epsilon}{q}-\beta_\epsilon-1}\exp\left(-\frac{c^{\rho_\epsilon}h^{-\rho_\epsilon}}{\mu_\epsilon}\right),$$

where $q=1$ when $\beta_x,\beta_\epsilon>0$, and $q>1$ when $\beta_x,\beta_\epsilon\le 0$. We note that $\varsigma_{h,q}^x$ and $\varsigma_{h,q}^\epsilon$ are used to express the bias terms.

Theorem 2. *[SS f_X and SS f_ϵ] Suppose that Assumptions M and SS hold. Then*

$$\begin{aligned}\mathcal{D}_n^{\varphi,X} &= O_p(n^{-1/2}\varrho_{T_n}^{\mathrm{s}}) \quad \textit{under } n^{-1/2}\varrho_{T_n}^{\mathrm{s}} \to 0,\\ \mathcal{D}_n^{\varphi,\epsilon} &= O_p\left(n^{-1/2}T_n^{-\beta_x}e^{\frac{T_n^{\rho_x}}{\mu_x}}\varrho_{T_n}^{\mathrm{s}}\right) \quad \textit{under } n^{-1/2}T_n^{-\beta_x}e^{\frac{T_n^{\rho_x}}{\mu_x}}\varrho_{T_n}^{\mathrm{s}} \to 0.\end{aligned}$$

Additionally suppose that Assumptions K (i) and (iii) hold. Then

$$\begin{aligned}\mathcal{D}_n^{f,X} &= O_p(n^{-1/2}h^{-1}\varrho_{h^{-1}}^{\mathrm{s}}+\varsigma_{h,q}^x) \quad \textit{under } n^{-1/2}\varrho_{h^{-1}}^{\mathrm{s}} \to 0,\\ \mathcal{D}_n^{f,\epsilon} &= O_p\left(n^{-1/2}h^{\beta_x-1}e^{\frac{h^{-\rho_x}}{\mu_x}}\varrho_{h^{-1}}^{\mathrm{s}}+\varsigma_{h,q}^\epsilon\right) \quad \textit{under } n^{-1/2}h^{\beta_x}e^{\frac{h^{-\rho_x}}{\mu_x}}\varrho_{h^{-1}}^{\mathrm{s}} \to 0.\end{aligned}$$

Analogous comments to Theorem 1 apply. When $T_n\to\infty$, $\mathcal{D}_n^{\varphi,\epsilon}$ converges slower than $\mathcal{D}_n^{\varphi,X}$ due to the additional factor, $T_n^{-\beta_x}e^{\frac{T_n^{\rho_x}}{\mu_x}}$. The first and second terms of the rates of $\mathcal{D}_n^{f,X}$ and $\mathcal{D}_n^{f,\epsilon}$ correspond to stochastic errors and bias terms, respectively. Note that due to the exponents in these terms, we typically set the bandwidth h as a logarithmic decay rate.

Furthermore, we consider mixed cases, where f_X and f_ϵ belong to different categories of smoothness. Let $\varrho_T^{\mathrm{os}} = T^{2\beta_x-2\beta_\epsilon+1}(\log T)\exp\left(\frac{2T^{\rho_\epsilon}}{\mu_\epsilon}\right)$ and $\varrho_T^{\mathrm{so}} = T^{-2\beta_x+2\beta_\epsilon+1}(\log T)\exp\left(\frac{2T^{\rho_x}}{\mu_x}\right)$.

Theorem 3. *[OS f_X and SS f_ϵ] Suppose that Assumptions M, OS for φ_X, and SS for φ_ϵ hold true. Then*

$$\begin{aligned}\mathcal{D}_n^{\varphi,X} &= O_p(n^{-1/2}\varrho_{T_n}^{\mathrm{os}}) \quad \textit{under } n^{-1/2}\varrho_{T_n}^{\mathrm{os}} \to 0,\\ \mathcal{D}_n^{\varphi,\epsilon} &= O_p(n^{-1/2}T_n^{\beta_x}\varrho_{T_n}^{\mathrm{os}}) \quad \textit{under } n^{-1/2}T_n^{\beta_x}\varrho_{T_n}^{\mathrm{os}} \to 0.\end{aligned}$$

Additionally suppose that Assumptions K (i) and (iii) hold. Then

$$\begin{aligned}\mathcal{D}_n^{f,X} &= O_p(n^{-1/2}h^{-1}\varrho_{h^{-1}}^{\mathrm{os}}+h^{\beta_x-1}) \quad \textit{under } n^{-1/2}\varrho_{h^{-1}}^{\mathrm{os}} \to 0,\\ \mathcal{D}_n^{f,\epsilon} &= O_p(n^{-1/2}h^{-\beta_x-1}\varrho_{h^{-1}}^{\mathrm{os}}+\varsigma_{h,q}^\epsilon) \quad \textit{under } n^{-1/2}h^{-\beta_x}\varrho_{h^{-1}}^{\mathrm{os}} \to 0.\end{aligned}$$

Theorem 4. *[SS f_X and OS f_ϵ] Suppose that Assumptions M, SS for φ_X, and OS for φ_ϵ hold true. Then*

$$\begin{aligned}\mathcal{D}_n^{\varphi,X} &= O_p(n^{-1/2}\varrho_{T_n}^{\mathrm{so}}) \quad \textit{under } n^{-1/2}\varrho_{T_n}^{\mathrm{so}} \to 0,\\ \mathcal{D}_n^{\varphi,\epsilon} &= O_p\left(n^{-1/2}T_n^{-\beta_x}e^{\frac{T_n^{\rho_x}}{\mu_x}}\varrho_{T_n}^{\mathrm{so}}\right) \quad \textit{under } n^{-1/2}T_n^{-\beta_x}e^{\frac{T_n^{\rho_x}}{\mu_x}}\varrho_{T_n}^{\mathrm{so}} \to 0.\end{aligned}$$

Additionally suppose that Assumptions K (i) and (iii) hold. Then

$$\begin{aligned}
\mathcal{D}_n^{f,X} &= O_p(n^{-1/2}h^{-1}\varrho_{h^{-1}}^{\mathrm{so}} + \varsigma_{h,q}^{x}) \quad \textit{under } n^{-1/2}\varrho_{h^{-1}}^{\mathrm{so}} \to 0,\\
\mathcal{D}_n^{f,\epsilon} &= O_p\left(n^{-1/2}h^{\beta_x-1}e^{\frac{h^{-\rho_x}}{\mu_x}}\varrho_{h^{-1}}^{\mathrm{so}} + h^{\beta_\epsilon-1}\right) \quad \textit{under } n^{-1/2}h^{\beta_x}e^{\frac{h^{-\rho_x}}{\mu_x}}\varrho_{h^{-1}}^{\mathrm{so}} \to 0.
\end{aligned}$$

Similar comments to Theorem 1 apply. The bias terms associated with ordinary (or resp. super) smooth densities are polynomials (or resp. exponentials) of h.

Finally, we present the uniform convergence rates of the CK estimator.

Theorem 5. *[CK estimator] The same uniform convergence results in Theorems 1-4 hold true even if we replace the LV estimator $(\hat{\varphi}_X, \hat{\varphi}_\epsilon, \hat{f}_X, \hat{f}_\epsilon)$ with the CK estimator $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon, \tilde{f}_X, \tilde{f}_\epsilon)$.*

2.1. **Remarks on Theorems 1-5.**

Remark 1 (Comparison with LV)**.** We note that LV established the uniform convergence rates of their estimator under the assumption that both X and ϵ have bounded support. On the other hand, our theorems do not require such boundedness. Also, the convergence rates obtained in our theorems are typically faster than those obtained in LV. For example, if we set $T_n = O\left((n/\log\log n)^{\alpha/2(1+\beta_x+\beta_\epsilon)}\right)$ with $0 < \alpha < 1/2$ as in Lemma 3.1 of LV, our Theorem 1 implies that

$$\mathcal{D}_n^{\varphi,X} = O_p\left(\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\alpha-\frac{\alpha}{2(1+\beta_x+\beta_\epsilon)}}\right),$$

and this convergence rate is faster than that given in LV, i.e., $\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\alpha}$. Similar comments apply to other cases.

Remark 2. [Comparison with Bonhomme and Robin, 2010] The convergence rates in our theorems are also faster than those given in Bonhomme and Robin (2010). For example, under Assumption OS, Bonhomme and Robin (2010, Theorem 1) implies that

$$\begin{aligned}
\mathcal{D}_n^{\varphi,X} &= O_p(n^{-1/2}T_n^{3\beta_x+3\beta_\epsilon+2}\log T_n),\\
\mathcal{D}_n^{\varphi,\epsilon} &= O_p(n^{-1/2}T_n^{3\beta_x+3\beta_\epsilon+2}\log T_n).
\end{aligned}$$

In Bonhomme and Robin (2010, Footnote 20), they give a comment that if they focus on the LV estimator, their convergence rate can be improved. Therefore, our results can be interpreted as a theoretical justification of their comment. It should also be noted that our assumption on (Y_1, Y_2) is weaker than Assumption A4 in Bonhomme and Robin (2010) since we do not need the existence of the moment generating functions of Y_1^2 and Y_1Y_2. More precisely, the same convergence rate given in Lemma 1 of their paper can be obtained under weaker conditions by proving a maximal inequality for the multivariate empirical characteristic function processes (Lemma 1 below). See Section 2.2 for a detailed discussion.

Remark 3 (LV and CK estimators)**.** In Theorem 5, we show that the LV and CK estimators achieve the same uniform convergence rates. On the other hand, it is open whether the LV estimator can achieve the L_2 convergence rate in CK. To control the L_2 risk of the LV-type estimators which are defined by the ratios of the (regularized) empirical averages, it seems crucial to introduce some regularization as in CK.

Remark 4 (Generalization for non-identical distributions of ϵ_1 and ϵ_2)**.** Although this paper (and also LV and CK) considers the case where the distributions of the measurement errors ϵ_1 and ϵ_2 are identical, other papers including the original one by Kotlarski (1967) allow different distributions for the measurement errors. For example, under Assumption M with allowing $f_{\epsilon_1} \neq f_{\epsilon_2}$ for $E[\epsilon_1] \neq 0$, the proof of Evdokimov (2010, Lemma 1) implies that

$$\begin{aligned}
\varphi_{\epsilon_1}(u) &= \exp\left(\int_0^u \frac{\partial\psi(v,-v)/\partial u_1}{\psi(v,-v)}dv - \mathrm{i}uE[Y_1]\right), \\
\varphi_{\epsilon_2}(u) &= \psi(-u,u)/\varphi_{\epsilon_1}(-u), \\
\varphi_X(u) &= \psi(u,0)/\varphi_{\epsilon_1}(u).
\end{aligned}$$

Thus, these characteristic functions can be estimated by estimating ψ and $E[Y_1]$ with $\hat{\psi}$ and $\frac{1}{n}\sum_{j=1}^n Y_{j,1}$, respectively. Indeed, by extending the current proof, we can show that these estimators (say, $\bar{\varphi}_{\epsilon_1}$, $\bar{\varphi}_{\epsilon_2}$, and $\bar{\varphi}_X$) satisfy: if $\tau_n = n^{-1/2}T_n \log T_n \left(\inf_{|v|\leq T_n} |\psi(v,-v)|^2\right)^{-1} \to 0$, then

$$\sup_{|u|\leq T_n} |\bar{\varphi}_{\epsilon_1}(u) - \varphi_{\epsilon_1}(u)| = O_p(\tau_n),$$

and if $\tau_n \left(\inf_{|u|\leq T_n} |\varphi_{\epsilon_1}(u)|\right)^{-1} \to 0$, then

$$\begin{aligned}
\sup_{|u|\leq T_n} |\bar{\varphi}_{\epsilon_2}(u) - \varphi_{\epsilon_2}(u)| &= O_p\left(\tau_n\left(\inf_{|u|\leq T_n} |\varphi_{\epsilon_1}(u)|\right)^{-1}\right), \\
\sup_{|u|\leq T_n} |\bar{\varphi}_X(u) - \varphi_X(u)| &= O_p\left(\tau_n\left(\inf_{|u|\leq T_n} |\varphi_{\epsilon_1}(u)|\right)^{-1}\right).
\end{aligned}$$

Thus, the convergence rates of these characteristic functions and the associated density estimators will be obtained under specific assumptions on the tail behaviors of ψ and φ_{ϵ_1}.

Furthermore, Evdokimov and White (2012) extended the above identification result to more general setups, where φ_{ϵ_1} and φ_{ϵ_2} are allowed to have zeros. In this case, we conjecture that additional regularizations by some ridge parameters need to be introduced for estimation (see, Hall and Meister, 2007). Full investigations for these extensions are left for future research.

2.2. **Maximal inequality for multivariate empirical characteristic function processes.** One of the key features of the LV and CK estimators is that they involve the multivariate empirical characteristic functions and their derivatives. In particular, key ingredients for the proofs of our theorems are to establish the uniform convergence rates of the bivariate random functions $\hat{\psi}(0,u)$ and $\partial\hat{\psi}(0,u)/\partial u_1$ over $|u| \leq T_n$ as in Lemma 2 in Appendix C. For the univariate case, Neumann and Reiss (2009, Theorem 4.1) obtained a maximal inequality under the weighted sup-norm for the empirical characteristic function processes, which was used to study convergence rates of nonparametric estimators for Lévy processes. Here we extend their result to a multivariate setup to obtain the uniform convergence rates of $\hat{\psi}(0,u)$ and $\partial\hat{\psi}(0,u)/\partial u_1$.

To present our result, we need some notation. Let $\{\boldsymbol{Y}_j = (Y_{j,1}, \dots, Y_{j,d})'\}_{j=1}^n$ be $\mathbb{R}^d$-valued i.i.d. random variables. For $\boldsymbol{t} = (t_1, \dots, t_d)' \in \mathbb{R}^d$, define

$$\begin{aligned}
\psi(\boldsymbol{t}) &= E[e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}_1}], \quad \hat{\psi}(\boldsymbol{t}) = \frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}_j}, \quad \boldsymbol{t}\cdot\boldsymbol{Y}_j = \sum_{\ell=1}^d t_\ell Y_{j,\ell}, \\
C_n(\boldsymbol{t}) &= \frac{1}{\sqrt{n}}\sum_{j=1}^n (e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}_j} - E[e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}}]) = \sqrt{n}(\hat{\psi}(\boldsymbol{t}) - \psi(\boldsymbol{t})), \\
C_n^{(\boldsymbol{k})}(\boldsymbol{t}) &= \frac{1}{\sqrt{n}}\sum_{j=1}^n \frac{\partial^{|\boldsymbol{k}|}}{\partial t_1^{k_1}\cdots\partial t_d^{k_d}}(e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}_j} - E[e^{\mathrm{i}\boldsymbol{t}\cdot\boldsymbol{Y}}]), \quad \text{for } \boldsymbol{k} = (k_1, \dots, k_d)' \in \mathbb{N}^d, \ |\boldsymbol{k}| = \sum_{j=1}^d k_j, \\
E[\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)}] &= E\left[\sup_{\boldsymbol{t}\in\mathbb{R}^d}\left(w(\|\boldsymbol{t}\|)|C_n^{(\boldsymbol{k})}(\boldsymbol{t})|\right)\right],
\end{aligned}$$

where $w(t) = (\log(e + |t|))^{-1/2-\delta}$ is a weight function for some $\delta > 0$ and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^d$.

The multivariate version of Neumann and Reiss (2009, Theorem 4.1), a maximal inequality for $C_n^{(\boldsymbol{k})}(\boldsymbol{t})$, is obtained as follows. The proof is presented in Appendix D.

Lemma 1. *Assume* $E\left[\left(\prod_{j=1}^d |Y_{1,j}|^{(k_j \vee 1/2)}\right)^{2+\eta}\right] < \infty$ *for some* $\eta > 0$*. Then*

$$\sup_{n\geq 1} E[\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)}] < \infty.$$

Remark 5. By this lemma, the uniform convergence rate of $\frac{\partial^{|\boldsymbol{k}|}\hat{\psi}(\boldsymbol{u})}{\partial u_1^{k_1}\cdots\partial u_d^{k_d}}$ over $\|\boldsymbol{u}\| \leq T_n$ is obtained as follows. Since

$$\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)} \geq \sqrt{n}\sup_{\|\boldsymbol{u}\|\leq T_n}\left|\frac{\partial^{|\boldsymbol{k}|}}{\partial u_1^{k_1}\cdots\partial u_d^{k_d}}(\hat{\psi}(\boldsymbol{u}) - \psi(\boldsymbol{u}))\right| \inf_{\|\boldsymbol{u}\|\leq T_n} w(\|\boldsymbol{u}\|),$$

Lemma 1 and the definition of $w(\cdot)$ imply that

$$E\left[\sup_{\|\boldsymbol{u}\|\leq T_n}\left|\frac{\partial^{|\boldsymbol{k}|}}{\partial u_1^{k_1}\cdots\partial u_d^{k_d}}(\hat{\psi}(\boldsymbol{u}) - \psi(\boldsymbol{u}))\right|\right] \leq \frac{\sup_{n\geq 1} E[\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)}]}{\sqrt{n}\inf_{\|\boldsymbol{u}\|\leq T_n} w(\|\boldsymbol{u}\|)} = O(n^{-1/2}\log T_n).$$

Thus, Markov's inequality implies

$$\sup_{\|\boldsymbol{u}\|\leq T_n}\left|\frac{\partial^{|\boldsymbol{k}|}}{\partial u_1^{k_1}\cdots\partial u_d^{k_d}}(\hat{\psi}(\boldsymbol{u}) - \psi(\boldsymbol{u}))\right| = O_p(n^{-1/2}\log T_n).$$

Remark 6. The essential part of the proof of Lemma 1 is the way of applying the maximal inequality for empirical processes in van der Vaart (1998, Corollary 19.35). To this end, it is sufficient to compute the bracketing number of the class of functions $\mathbb{G}_{1,k} = \left\{\boldsymbol{y} \mapsto \frac{\partial^k}{\partial t_1}\cos(\boldsymbol{t}\cdot\boldsymbol{y}) : \boldsymbol{t} \in \mathbb{R}^d\right\} \cup \left\{\boldsymbol{y} \mapsto \frac{\partial^k}{\partial t_1}\sin(\boldsymbol{t}\cdot\boldsymbol{y}) : \boldsymbol{t} \in \mathbb{R}^d\right\}$ in (22) in Appendix D, and for the computation we only need polynomial moments of each component of $\boldsymbol{Y}$. This point is different from the proof of Bonhomme and Robin (2010, Lemma 1), which is based on applications of a maximal inequality for empirical processes and a general Chernoff bound that requires the existence of exponential moments of random variables (eqs. (A6) and (A9) in their paper, respectively).

Remark 7. Finally, we emphasize Lemma 1 could be applied to other contexts in econometrics and statistics. For example, it can be applied to examples discussed in Bonhomme and Robin (2010) and could also be used to extend the results in Kato and Kurisu (2020) and Kurisu (2019), which study nonparametric inference on univariate Lévy processes under high- and low-frequency observations, to multivariate setups.

3. Concluding remarks

In this paper, we derive the uniform convergence rates of the Li and Vuong (1998) and Comte and Kappus (2015) estimators for the classical measurement error model with repeated measurements, where its identification is achieved by Kotlarski's (1967) identity. The obtained convergence rates are faster than the ones derived in Li and Vuong (1998) and Bonhomme and Robin (2010), and also our assumptions are weaker than the ones in these papers. As a technical lemma of independent interest, we obtain a maximal inequality for the multivariate normalized empirical characteristic function process. We conjecture that our results are useful to address some open questions or new applications of econometric methods using Kotlarski's identify. We close this article by discussing such possibilities.

- **(i):** Adusumilli *et al.* (2020) already used an adapted version of our results to conduct inference on the cumulative distribution function of the error-free variable. To allow both unbounded support and non-existence of the moment generating function for observables, it is critical to employ our uniform convergence results. Furthermore, in Adusumilli *et al.* (2020, pp. 137-138), even our faster convergence rates are not sufficient to establish the asymptotic validity of the naive bootstrap inference for the distribution function of the error-free variable. Thus, Adusumilli *et al.* (2020) developed an alternative bootstrap procedure based on a modified statistic using a subsample. In this case, our faster convergence rates are useful to allow larger size of the subsample, which yields better power properties.
- **(ii):** Otsu and Taylor (2019) proposed a specification test for errors-in-variables models, where the measurement errors are required to be symmetrically distributed. To extend their approach to allow possibly asymmetric distributions on the measurement errors, one may plug-in the LV-type estimators for the characteristic functions of the measurement errors to their test statistic. In this case, if we wish to guarantee that the estimation errors for the plug-in LV-type estimators are dominated by the main term considered in Otsu and Taylor (2019), our faster convergence rates are useful to establish such asymptotic negligibility under weaker conditions. Even if the convergence rates of Otsu and Taylor's (2019) statistic are not sufficiently fast, we can adapt a subsample-based modification as in Adusumilli *et al.* (2020) to Otsu and Taylor's (2019) statistic so that our main theorems can be applied in an analogous way.
- **(iii):** For some existing methods, such as Li (2002) and Li and Hsiao (2004), only consistency is established in the literature and their convergence rates and limiting distributions remain open questions. The estimators by Li (2002) and Li and Hsiao (2004) can be considered as semiparametric M-estimators, where the nonparametric nuisance parameters

are estimated by the LV-type estimators. Thus, faster uniform convergence rates under weaker assumptions on the nonparametric plug-in components will be useful to achieve faster convergence rates for the estimators of finite-dimensional components and to derive the limiting distributions under mild regularity conditions.

Appendix A. Proof of Theorems 1 and 5 under Assumption OS

Here we present the proof of Theorem 5 for the CK estimator under Assumption OS (i.e., $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon, \tilde{f}_X, \tilde{f}_\epsilon)$ achieve the same uniform convergence rates as $(\hat{\varphi}_X, \hat{\varphi}_\epsilon, \hat{f}_X, \hat{f}_\epsilon)$ in Theorem 1). The proof of Theorem 1 on the LV estimator is its specialization (in particular, repeat the same arguments in Appendices A.1-A.4 by replacing $\tilde{\psi}(\cdot,\cdot)$ with $\hat{\psi}(\cdot,\cdot)$).

We use the following notation.

$$
\begin{aligned}
\Delta(u) &= \log\left(\frac{\tilde{\varphi}_X^{mod}(u)}{\varphi_X(u)}\right) = \int_0^u \left(\frac{\partial\hat{\psi}(0,u_2)/\partial u_1}{\tilde{\psi}(0,u_2)} - \frac{\partial\psi(0,u_2)/\partial u_1}{\psi(0,u_2)}\right) du_2, \\
R_1(u) &= \frac{1}{\psi(0,u)} - \frac{1}{\tilde{\psi}(0,u)}, \qquad R_2(u) = \frac{\partial\hat{\psi}(0,u)}{\partial u_1} - \frac{\partial\psi(0,u)}{\partial u_1}.
\end{aligned}
$$

A.1. **Proof for $\tilde{\varphi}_X$.** The definition of $\tilde{\varphi}_X$ implies

$$
\begin{aligned}
|\tilde{\varphi}_X(u) - \varphi_X(u)| &= |\tilde{\varphi}_X^{mod}(u) - \varphi_X(u)|\mathbb{I}\{|\tilde{\varphi}_X^{mod}(u)| \le 1\} + \left|\frac{\tilde{\varphi}_X^{mod}(u)}{|\tilde{\varphi}_X^{mod}(u)|} - \varphi_X(u)\right|\mathbb{I}\{|\tilde{\varphi}_X^{mod}(u)| > 1\} \\
&\le |\tilde{\varphi}_X^{mod}(u) - \varphi_X(u)|, \qquad (10)
\end{aligned}
$$

where the inequality (for the case of $|\tilde{\varphi}_X^{mod}(u)| > 1$) follows from the facts that $\varphi_X(u)$ is inside the unit circle on $\mathbb{C}$ but $\tilde{\varphi}_X^{mod}(u)$ is outside. Thus, we obtain

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

where the second inequality follows from (10) and the fact that $|\tilde{\varphi}_X(u) - \varphi_X(u)| \le 2$, the equality follows from the definitions of $\tilde{\varphi}_X^{mod}(u)$ and $\Delta(u)$, and the third inequality follows from the fact that $|1 - e^z| \le 2|z|$ for $z \in \mathbb{C}$ with $|z| \le 1$. Thus, it is sufficient for the conclusion to derive the rate of $\sup_{u\in[-T_n,T_n]} |\Delta(u)|$.

Decompose

$$
\begin{aligned}
\Delta(u) &= \int_0^u \frac{R_2(u_2)}{\psi(0,u_2)} du_2 + \int_0^u \frac{\partial\psi(0,u_2)}{\partial u_1} R_1(u_2) du_2 + \int_0^u R_1(u_2)R_2(u_2)du_2 \\
&:= \Delta_1(u) + \Delta_2(u) + \Delta_3(u),
\end{aligned}
$$

which are bounded as

$$
\begin{aligned}
\sup_{u\in[-T_n,T_n]} |\Delta_1(u)| &\leq \sup_{u\in[-T_n,T_n]} |R_2(u)| \left(\int_0^{T_n} \frac{1}{|\psi(0,u_2)|} du_2 \right), \\
\sup_{u\in[-T_n,T_n]} |\Delta_2(u)| &\leq \sup_{u\in[-T_n,T_n]} |R_1(u)| \left(\int_0^{T_n} \left| \frac{\partial \psi(0,u_2)}{\partial u_1} \right| du_2 \right) \\
&= \sup_{u\in[-T_n,T_n]} |R_1(u)| \left(\int_0^{T_n} |E[Y_1 e^{iu_2Y_2}]| du_2 \right) \leq T_n E[|Y_1|] \sup_{u\in[-T_n,T_n]} |R_1(u)|, \\
\sup_{u\in[-T_n,T_n]} |\Delta_3(u)| &\leq T_n \sup_{u\in[-T_n,T_n]} |R_1(u)| \sup_{u\in[-T_n,T_n]} |R_2(u)|.
\end{aligned}
$$

Therefore, the conclusion follows from Lemmas 2 and 3.

A.2. **Proof for $\tilde{\varphi}_\epsilon$.** Note that

$$
\begin{aligned}
\sup_{u\in[-T_n,T_n]} \left| \frac{\hat{\psi}(0,u) - \psi(0,u)}{\psi(0,u)} \right| &= O_p\left(n^{-1/2} T_n^{\beta_x+\beta_\epsilon} \log T_n \right), \\
\sup_{u\in[-T_n,T_n]} \left| \frac{\tilde{\varphi}_X(u) - \varphi_X(u)}{\varphi_X(u)} \right| &= O_p\left(n^{-1/2} T_n^{3\beta_x+2\beta_\epsilon+1} \log T_n \right).
\end{aligned}
$$

We also note that

$$
\begin{aligned}
& |\bar{\varphi}_X(u) - \varphi_X(u)| \leq |\bar{\varphi}_X(u) - \tilde{\varphi}_X(u)| + |\tilde{\varphi}_X(u) - \varphi_X(u)| \\
\leq\; & |\bar{\varphi}_X(u) - \tilde{\varphi}_X(u)| + 4|\Delta(u)| \\
=\; & |\bar{\varphi}_X(u) - \tilde{\varphi}_X(u)| \mathbb{I}\{\sqrt{n}|\tilde{\varphi}_X(u)| \leq 1\} + |\bar{\varphi}_X(u) - \tilde{\varphi}_X(u)| \mathbb{I}\{\sqrt{n}|\tilde{\varphi}_X(u)| > 1\} + 4|\Delta(u)| \\
\leq\; & (1/\sqrt{n} + |\tilde{\varphi}_X(u)|) \mathbb{I}\{\sqrt{n}|\tilde{\varphi}_X(u)| \leq 1\} + 0 + 4|\Delta(u)| \\
\leq\; & 2/\sqrt{n} + 4|\Delta(u)|,
\end{aligned}
$$

and this implies that

$$
\sup_{u\in[-T_n,T_n]} |\bar{\varphi}_X(u) - \varphi_X(u)| = O\left(\sup_{u\in[-T_n,T_n]} |\tilde{\varphi}_X(u) - \varphi_X(u)| \right). \tag{12}
$$

First we show

$$
\sup_{u\in[-T_n,T_n]} |\log \tilde{\varphi}_\epsilon(u) - \log \varphi_\epsilon(u)| = O_p\left(n^{-1/2} T_n^{3\beta_x+2\beta_\epsilon+1} \log T_n \right).
$$

Let $F(y) = \log(1+y)$, and $\zeta(u) = (\hat{\psi}(0,u) - \psi(0,u))/\psi(0,u)$. Observe that for any $|u| \leq T_n$,

$$
\begin{aligned}
(F \circ \zeta)(u) &= F(0) + F'(\theta_1 \zeta(u))\zeta(u) = (F'(0) + \theta_1 F''(\theta_2 \zeta(u))\zeta(u))\zeta(u) \\
&= \zeta(u) + \theta_1 F''(\theta_2 \zeta(u))\zeta^2(u),
\end{aligned}
$$

for some $\theta_1, \theta_2 \in [0,1]$. Then we have

$$
\sup_{u\in[-T_n,T_n]} \left| \log\left(\frac{\hat{\psi}(0,u)}{\psi(0,u)} \right) - \frac{\hat{\psi}(0,u) - \psi(0,u)}{\psi(0,u)} \right| \leq O\left(\sup_{u\in[-T_n,T_n]} |\zeta(u)|^2 \right) = O_p\left(n^{-1} T_n^{2\beta_x+2\beta_\epsilon} (\log T_n)^2 \right),
$$

which yields

$$\sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\psi(0,u))\right| = O\left(\sup_{u\in[-T_n,T_n]}|\zeta(u)|\right) = O_p\left(n^{-1/2}T_n^{\beta_x+\beta_\epsilon}\log T_n\right). \tag{13}$$

Likewise, we can show that

$$\sup_{u\in[-T_n,T_n]}|\log(\tilde{\varphi}_X(u)/\varphi_X(u))| = O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\tilde{\varphi}_X(u)-\varphi_X(u)}{\varphi_X(u)}\right|\right) = O_p\left(n^{-1/2}T_n^{3\beta_x+2\beta_\epsilon+1}\log T_n\right). \tag{14}$$

Together with (12), (13), and (14), we have that

$$\begin{aligned}
&\sup_{u\in[-T_n,T_n]}|\log\tilde{\varphi}_\epsilon(u)-\log\varphi_\epsilon(u)| \\
=\;& \sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\bar{\varphi}_X(u)) - \log(\psi(0,u)/\varphi_X(u))\right| \\
\le\;& \sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\psi(0,u))\right| + \sup_{u\in[-T_n,T_n]}|\log(\bar{\varphi}_X(u)/\varphi_X(u))| \\
=\;& O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\hat{\psi}(0,u)-\psi(0,u)}{\psi(0,u)}\right|\right) + O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\tilde{\varphi}_X(u)-\varphi_X(u)}{\varphi_X(u)}\right|\right) \\
=\;& O_p\left(n^{-1/2}T_n^{3\beta_x+2\beta_\epsilon+1}\log T_n\right) = o_p(1). \qquad (15)
\end{aligned}$$

On the other hand, since $|\varphi_\epsilon(u)|\le 1$ and $|e^z-1|\le|z|$ for $z\in\mathbb{C}$ with $|z|<1$, a Taylor expansion of $\tilde{\varphi}_\epsilon(u)-\varphi_\epsilon(u)$ gives that

$$\sup_{u\in[-T_n,T_n]}|\tilde{\varphi}_\epsilon(u)-\varphi_\epsilon(u)| \le O\left(\sup_{u\in[-T_n,T_n]}|\log\tilde{\varphi}_\epsilon(u)-\log\varphi_\epsilon(u)|\right),$$

provided $\sup_{u\in[-T_n,T_n]}|\log\tilde{\varphi}_\epsilon(u)-\log\varphi_\epsilon(u)|<1$. Therefore, (15) yields the desired result.

A.3. **Proof for $\tilde{f}_X$.** Note that for all x,

$$\begin{aligned}
&|\tilde{f}_X(x)-f_X(x)| \\
=\;& \left|\frac{1}{2\pi}\int_{-h^{-1}}^{h^{-1}} e^{-\mathrm{i}ux}\{\tilde{\varphi}_X(u)-\varphi_X(u)\}\varphi_K(hu)du + \frac{1}{2\pi}\int e^{-\mathrm{i}ux}\varphi_X(u)\{\varphi_K(hu)-1\}du\right| \\
\le\;& C_K\frac{h^{-1}}{\pi}\sup_{u\in[-h^{-1},h^{-1}]}|\tilde{\varphi}_X(u)-\varphi_X(u)| + \frac{1}{2\pi}\int|\varphi_X(u)||\varphi_K(hu)-1|du, \qquad (16)
\end{aligned}$$

where the inequality follows from $|e^{-\mathrm{i}ux}|=1$ and $\sup_{x\in\mathbb{R}}|\varphi_K(x)|\le C_K$ for some positive constant $C_K<\infty$. By the first part of this theorem, the first term of (16) is of order $O_p\left(n^{-1/2}h^{-2\beta_x-2\beta_\epsilon-2}\log h^{-1}\right)$.

Since K is a p-th order kernel, the p-th order Taylor expansion of $\varphi_K(x)$ around $x=0$ yields $\varphi_K(x)=1+m(x)x^p$ for all $x\in[-1,1]$, where m is some continuous function on $[-1,1]$ (for $|x|>1$, Assumption K requires $\varphi_K(x)=0$). Therefore, the second term of (16) satisfies

$$\begin{aligned}
&\int|\varphi_X(u)||\varphi_K(hu)-1|du \le C_x\int|u|^{-\beta_x}|\varphi_K(hu)-1|du \\
\le\;& C_x\sup_{v\in[-1,1]}|m(v)|h^p\int_{-h^{-1}}^{h^{-1}}|u|^{-\beta_x+p}du + 2C_x\int_{h^{-1}}^{\infty}|u|^{-\beta_x}du = O(h^{\beta_x-1}), \qquad (17)
\end{aligned}$$

where the first inequality follows from Assumption OS, and the second inequality follows from the assumption on K. Therefore, the conclusion follows.

A.4. **Proof for $\tilde{f}_\epsilon$.** The proof is similar to the one in Appendix A.3. Replace "$\tilde{f}_X, f_X, \tilde{\varphi}_X, \varphi_X$" with "$\tilde{f}_\epsilon, f_\epsilon, \tilde{\varphi}_\epsilon, \varphi_\epsilon$" and proceed in the same manner by using the uniform convergence rate of $\tilde{\varphi}_\epsilon$ obtained in Appendix A.2.

Appendix B. Proofs for other cases

B.1. **Proof of Theorems 2 and 5 under Assumption SS.** The proof for $\tilde{\varphi}_X$ is exactly same as the one in Appendix A.1 except for the last sentence. Under Assumption SS, the conclusion follows from Lemmas 2 and 4 (instead of Lemma 3 under Assumption OS), and characterization of $\int_0^{T_n} \frac{1}{|\psi(0,u_2)|} du_2$ using Assumption SS.

The proof for $\tilde{\varphi}_\epsilon$ proceeds in the same way as the one in Appendix A.2. The only difference is to insert the orders for $\psi(0,u)$, $\varphi_X(u)$, and $\varphi_\epsilon(u)$ implied from Assumption SS.

The proof for $\tilde{f}_X$ proceeds in the same way as the one in Appendix A.3. For the first term of (16), replace the order of $\sup_{u\in[-h^{-1},h^{-1}]} |\tilde{\varphi}_X(u) - \varphi_X(u)|$ with the one derived from the first part of this theorem by using Lemma 4 (instead of Lemma 3 under Assumption OS). For the second term of (16) (i.e., the bias term), under Assumptions SS and K with $\varphi_K(x) = 1$, $|x| \le c$ for some $0 < c \le 1$, there exists a positive constant C_0 such that

$$\begin{aligned} \int |\varphi_X(u)||\varphi_K(hu) - 1| du &\le C_0 C_x \int_{ch^{-1}}^{\infty} |u|^{\beta_x} e^{-|u|^{\rho_x}/\mu_x} du \\ &= \begin{cases} O\left(h^{-\beta_x - 1 + \rho_x/q} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right)\right) & \text{if } \beta_x \le 0, \\ O\left(h^{-\beta_x - 1 + \rho_x} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right)\right) & \text{if } \beta_x > 0, \end{cases} \end{aligned} \tag{18}$$

where q is any constant with $q > 1$. In fact, when $\beta_x \le 0$, by using Lemma 4.2 in LV and Hölder's inequality, we have that

$$\begin{aligned} \int_{ch^{-1}}^{\infty} |u|^{\beta_x} e^{-|u|^{\rho_x}/\mu_x} du &\le \left(\int_{ch^{-1}}^{\infty} |u|^{q_1(\beta_x - 1)} du\right)^{1/q_1} \left(\int_{ch^{-1}}^{\infty} |u|^{q_2} e^{-q_2 |u|^{\rho_x}/\mu_x} du\right)^{1/q_2} \\ &= O\left(h^{-\beta_x + 1 - 1/q_1}\right) \times O\left(h^{\rho_x/q_2 - 1 - 1/q_2} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right)\right) \\ &= O\left(h^{\rho_x/q_2 - \beta_x - 1} \exp\left(-\frac{c^{\rho_x} h^{-\rho_x}}{\mu_x}\right)\right), \end{aligned}$$

where q_1 and q_2 are constants with $1/q_1 + 1/q_2 = 1$ and $q_1, q_2 > 1$. Therefore, the conclusion follows.

Finally, the proof for $\tilde{f}_\epsilon$ is similar to the one for $\tilde{f}_X$ above. Replace "$\tilde{f}_X, f_X, \tilde{\varphi}_X, \varphi_X$" with "$\tilde{f}_\epsilon, f_\epsilon, \tilde{\varphi}_\epsilon, \varphi_\epsilon$" and proceed in the same manner by using the uniform convergence rate of $\tilde{\varphi}_\epsilon$. The bias term is bounded as in (18) by replacing "φ_X, β_x, ρ_x, μ_x, C_x" with "φ_ϵ, β_ϵ, ρ_ϵ, μ_ϵ, C_ϵ".

B.2. **Proof of Theorems 3 and 5 with OS f_X and SS f_ϵ.** The proof for $\tilde{\varphi}_X$ is exactly same as the one in Appendix A.1 except for the last sentence. For the case of OS f_X and SS f_ϵ, the conclusion follows from Lemmas 2 and 5 (instead of Lemma 3 under Assumption OS), and characterization of $\int_0^{T_n} \frac{1}{|\psi(0,u_2)|} du_2$ using Assumptions OS for f_X and SS f_ϵ.

The proof for $\tilde{\varphi}_\epsilon$ proceeds in the same way as the one in Appendix A.2. The only difference is to insert the orders for $\psi(0,u)$, $\varphi_X(u)$, and $\varphi_\epsilon(u)$ implied from Assumptions OS for f_X and SS f_ϵ.

The proof for $\tilde{f}_X$ proceeds in the same way as the one in Appendix A.3. For the first term of (16), replace the order of $\sup_{u\in[-h^{-1},h^{-1}]}|\tilde{\varphi}_X(u)-\varphi_X(u)|$ with the one derived from the first part of this theorem by using Lemma 5 (instead of Lemma 3 under Assumption OS). For the second term of (16) (i.e., the bias term), we can apply the same order in (17) because f_X satisfies Assumption OS.

Finally, the proof for $\tilde{f}_\epsilon$ is similar to the one for $\tilde{f}_X$ above. Replace "$\tilde{f}_X, f_X, \tilde{\varphi}_X, \varphi_X$" with "$\tilde{f}_\epsilon, f_\epsilon, \tilde{\varphi}_\epsilon, \varphi_\epsilon$" and proceed in the same manner by using the uniform convergence rate of $\tilde{\varphi}_\epsilon$. Also, the bias term is bounded as in (18) by replacing "φ_X, β_x, ρ_x, μ_x, C_x" with "φ_ϵ, β_ϵ, ρ_ϵ, μ_ϵ, C_ϵ".

B.3. **Proof of Theorems 4 and 5 with SS f_X and OS f_ϵ.** The proof for $\tilde{\varphi}_X$ is exactly same as the one in Appendix A.1 except for the last sentence. For the case of SS f_X and OS f_ϵ, the conclusion follows from Lemmas 2 and 6 (instead of Lemma 3 under Assumption OS), and characterization of $\int_0^{T_n}\frac{1}{|\psi(0,u_2)|}du_2$ using Assumptions SS for f_X and OS f_ϵ.

The proof for $\tilde{\varphi}_\epsilon$ proceeds in the same way as the one in Appendix A.2. The only difference is to insert the orders for $\psi(0,u)$, $\varphi_X(u)$, and $\varphi_\epsilon(u)$ implied from Assumptions SS for f_X and OS f_ϵ.

The proof for $\tilde{f}_X$ proceeds in the same way as the one in Appendix A.3. For the first term of (16), replace the order of $\sup_{u\in[-h^{-1},h^{-1}]}|\tilde{\varphi}_X(u)-\varphi_X(u)|$ with the one derived from the first part of this theorem by using Lemma 6 (instead of Lemma 3 under Assumption OS). For the second term of (16) (i.e., the bias term), we can apply the same order in (18) because f_X satisfies Assumption SS.

Finally, the proof for $\tilde{f}_\epsilon$ is similar to the one for $\tilde{f}_X$ above. Replace "$\tilde{f}_X, f_X, \tilde{\varphi}_X, \varphi_X$" with "$\tilde{f}_\epsilon, f_\epsilon, \tilde{\varphi}_\epsilon, \varphi_\epsilon$" and proceed in the same manner by using the uniform convergence rate of $\tilde{\varphi}_\epsilon$. Also, the bias term is bounded as in (17) because f_ϵ satisfies Assumption OS.

Appendix C. Lemmas

Lemma 2. *Assume $E[|Y_1|^{2+\eta}]<\infty$ for some $\eta>0$. Then we have that*

$$\sup_{u\in[-T_n,T_n]}|\hat{\psi}(0,u)-\psi(0,u)|=O_p(n^{-1/2}\log T_n),\qquad \sup_{u\in[-T_n,T_n]}|R_2(u)|=O_p(n^{-1/2}\log T_n).$$

Proof. Since $\hat{\psi}(0,u)=\frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}uY_{j,2}}$ is the empirical characteristic function of $\psi(0,u)$, we can apply Lemma 1 with $d=1$ and $\boldsymbol{k}=0$. This yields that

$$\sup_{u\in[-T_n,T_n]}|\hat{\psi}(0,u)-\psi(0,u)|=O_p(n^{-1/2}\log T_n).$$

An application of Lemma 1 with $d=2$, $\boldsymbol{k}=(1,0)'$ also yields that

$$\sup_{u\in[-T_n,T_n]}|R_2(u)|=O_p(n^{-1/2}\log T_n).$$

□

Lemma 3. *Suppose Assumption OS holds. Assume $E[|Y_1|^{2+\eta}] < \infty$ for some $\eta > 0$ and $n^{-1/2}T_n^{\beta_x+\beta_\epsilon}\log T_n \to 0$ as $n \to \infty$. Then we have that*

$$\sup_{u\in[-T_n,T_n]} |R_1(u)| = O_p(n^{-1/2}T_n^{2\beta_x+2\beta_\epsilon}\log T_n).$$

Proof. Note that

$$|R_1(u)| \le \frac{|\tilde{\psi}(0,u)-\psi(0,u)|}{|\psi(0,u)||\tilde{\psi}(0,u)|} \le \frac{|\tilde{\psi}(0,u)-\psi(0,u)|}{|\psi(0,u)||\hat{\psi}(0,u)|}.$$

Here, we used the fact $|\tilde{\psi}(0,u)| \le |\hat{\psi}(0,u)|$. By the definition of $\tilde{\psi}(0,u)$, we have that

$$\begin{aligned}
|\tilde{\psi}(0,u)-\psi(0,u)| &\le |\tilde{\psi}(0,u)-\hat{\psi}(0,u)| + |\hat{\psi}(0,u)-\psi(0,u)| \\
&\le |\tilde{\psi}(0,u)-\hat{\psi}(0,u)|\mathbb{I}\{|\hat{\psi}(0,u)| > 1/\sqrt{n}\} + |\tilde{\psi}(0,u)-\hat{\psi}(0,u)|\mathbb{I}\{|\hat{\psi}(0,u)| \le 1/\sqrt{n}\} \\
&\quad +|\hat{\psi}(0,u)-\psi(0,u)| \\
&= |n^{-1/2}-\hat{\psi}(0,u)|\mathbb{I}\{|\hat{\psi}(0,u)| \le 1/\sqrt{n}\} + |\hat{\psi}(0,u)-\psi(0,u)| \\
&\le \frac{1}{\sqrt{n}}\mathbb{I}\{|\hat{\psi}(0,u)| \le 1/\sqrt{n}\} + |\hat{\psi}(0,u)|\mathbb{I}\{|\hat{\psi}(0,u)| \le 1/\sqrt{n}\} + |\hat{\psi}(0,u)-\psi(0,u)| \\
&\le \frac{2}{\sqrt{n}} + |\hat{\psi}(0,u)-\psi(0,u)|.
\end{aligned}$$

Combining this with Lemma 2, we have

$$\sup_{u\in[-T_n,T_n]} |\tilde{\psi}(0,u)-\psi(0,u)| = O_p(n^{-1/2}\log T_n). \tag{19}$$

Assumption OS and Lemma 2 using the condition $n^{-1/2}T_n^{\beta_x+\beta_\epsilon}\log T_n \to 0$ also imply that

$$\inf_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)| \ge \inf_{u\in[-T_n,T_n]} |\psi(0,u)| - \sup_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)-\psi(0,u)| = O_p(T_n^{-\beta_x-\beta_\epsilon}). \tag{20}$$

Combining (19) and (20), we finally obtain that

$$\sup_{u\in[-T_n,T_n]} |R_1(u)| \le \frac{\sup_{u\in[-T_n,T_n]} |\tilde{\psi}(0,u)-\psi(0,u)|}{\inf_{u\in[-T_n,T_n]} |\psi(0,u)| \inf_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)|} = O_p(n^{-1/2}T^{2\beta_x+2\beta_\epsilon}\log T_n).$$

□

Lemma 4. *Suppose Assumption SS holds. Assume $E[|Y_1|^{2+\eta}] < \infty$ for some $\eta > 0$ and $n^{-1/2}T_n^{-\beta_x-\beta_\epsilon}(\log T_n)\exp\left(\frac{T_n^{\rho_x}}{\mu_x} + \frac{T_n^{\rho_\epsilon}}{\mu_\epsilon}\right) \to 0$ as $n \to \infty$. Then we have that*

$$\sup_{u\in[-T_n,T_n]} |R_1(u)| = O_p\left(n^{-1/2}T_n^{-2\beta_x-2\beta_\epsilon}(\log T_n)\exp\left(\frac{2T_n^{\rho_x}}{\mu_x} + \frac{2T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\right).$$

Proof. The proof is same as Lemma 3 up to (19). Then Assumption SS and Lemma 2 combined with the condition $n^{-1/2}T_n^{-\beta_x-\beta_\epsilon}(\log T_n)\exp\left(\frac{T_n^{\rho_x}}{\mu_x} + \frac{T_n^{\rho_\epsilon}}{\mu_\epsilon}\right) \to 0$ also imply that

$$\begin{aligned}
\inf_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)| &\ge \inf_{u\in[-T_n,T_n]} |\psi(0,u)| - \sup_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)-\psi(0,u)| \\
&= O_p\left(T_n^{\beta_x+\beta_\epsilon}\exp\left(-\frac{T_n^{\rho_x}}{\mu_x} - \frac{T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\right). \qquad (21)
\end{aligned}$$

By (19) and (21), we finally obtain that

$$\begin{aligned}\sup_{u\in[-T_n,T_n]}|R_1(u)| &\leq \frac{\sup_{u\in[-T_n,T_n]}|\tilde{\psi}(0,u)-\psi(0,u)|}{\inf_{u\in[-T_n,T_n]}|\psi(0,u)|\inf_{u\in[-T_n,T_n]}|\hat{\psi}(0,u)|}\\ &= O_p\left(n^{-1/2}T_n^{-2\beta_x-2\beta_\epsilon}(\log T_n)\exp\left(\frac{2T_n^{\rho_x}}{\mu_x}+\frac{2T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\right).\end{aligned}$$

□

Analogous lemmas can be obtained for the mixed cases. Since the proofs are similar to the ones in Lemmas 3 and 4, they are omitted.

Lemma 5. *Suppose Assumption OS holds for f_X and Assumption SS holds for f_ϵ. Assume $E[|Y_1|^{2+\eta}]<\infty$ for some $\eta>0$ and $n^{-1/2}T_n^{\beta_x-\beta_\epsilon}(\log T_n)\exp\left(\frac{T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\to 0$ as $n\to\infty$. Then we have that*

$$\sup_{u\in[-T_n,T_n]}|R_1(u)| = O_p\left(n^{-1/2}T_n^{2\beta_x-2\beta_\epsilon}(\log T_n)\exp\left(\frac{2T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\right).$$

Lemma 6. *Suppose Assumption SS holds for f_X and Assumption OS holds for f_ϵ. Assume $E[|Y_1|^{2+\eta}]<\infty$ for some $\eta>0$ and $n^{-1/2}T_n^{-\beta_x+\beta_\epsilon}(\log T_n)\exp\left(\frac{T_n^{\rho_x}}{\mu_x}\right)\to 0$ as $n\to\infty$. Then we have that*

$$\sup_{u\in[-T_n,T_n]}|R_1(u)| = O_p\left(n^{-1/2}T_n^{-2\beta_x+2\beta_\epsilon}(\log T_n)\exp\left(\frac{2T_n^{\rho_x}}{\mu_x}\right)\right).$$

Appendix D. Proof of Lemma 1

We prove the case when $\boldsymbol{k}=(k,0,\ldots,0)'$. In this case, we rewrite $C_n^{(\boldsymbol{k})}$ as $C_n^{(k)}$. Other cases can be proved similarly. We follow the notations used in the proof of Neumann and Reiss (2009, Theorem 4.1). Given two functions $\ell,u:\mathbb{R}^d\to\mathbb{R}$ the bracket $[\ell,u]$ denotes the set of functions f with $\ell\leq f\leq u$. For a set G of functions the L^2-bracketing number $N_{[\cdot]}(\epsilon,G)$ is the minimum number of brackets $[\ell_j,u_j]$, satisfying $E[(u_j(\boldsymbol{Y})-\ell_j(\boldsymbol{Y}))^2]\leq\epsilon^2$, that are needed to cover G. The associated bracketing number is defined as

$$J_{[\cdot]}(\delta,G)=\int_0^\delta\sqrt{\log(N_{[\cdot]}(\epsilon,G))}d\epsilon.$$

Moreover, a function F is called an envelop function for G if $|f|\leq F$ holds for all $f\in G$. We decompose $C_n^{(k)}$ into its real and imaginary parts and introduce the set of functions

$$\mathbb{G}_{1,k}:=\left\{\boldsymbol{y}\mapsto\frac{\partial^k}{\partial t_1}\cos(\boldsymbol{t}\cdot\boldsymbol{y}):\boldsymbol{t}\in\mathbb{R}^d\right\}\cup\left\{\boldsymbol{y}\mapsto\frac{\partial^k}{\partial t_1}\sin(\boldsymbol{t}\cdot\boldsymbol{y}):\boldsymbol{t}\in\mathbb{R}^d\right\}=:\mathbb{G}_{1,k}^{(c)}\cup\mathbb{G}_{1,k}^{(s)}. \quad (22)$$

Since an envelop function of $\mathbb{G}_{1,k}$ is given by $F_k=|y_1|^k$ and $E[|Y_1|^{2k}]<\infty$, an application of van der Vaart (1998, Corollary 19.35) yields that

$$\sup_{n\geq 1}E[\|C_n^{(k)}\|_{L_\infty(w)}]\leq CJ_{[\cdot]}\left(\sqrt{E[Y_1^{2k}]},\mathbb{G}_{1,k}\right),$$

for a universal constant C. Define

$$M:=M(\epsilon,k):=\inf\left\{m>0:E[|Y_1|^{2k}1\{|Y_1|\geq m\}]\right\}\leq\epsilon^2.$$

Furthermore, let $\|w\|_\infty = \sup_{x\in\mathbb{R}} |w(x)|$ and we set for grid points $\boldsymbol{t} \in \mathbb{R}^d$,

$$\begin{aligned} g_j^{\pm}(\boldsymbol{y}) &= \left(w(\|\boldsymbol{t}_j\|) \frac{\partial^k}{\partial t_1} \cos(\boldsymbol{t}\cdot\boldsymbol{y}) \pm \epsilon |y_1|^k \right) \mathbb{I}\{y_1 \in [-M, M]\} \pm \|w\|_\infty |y_1|^k \mathbb{I}\{y_1 \in [-M, M]^c\}, \\ h_j^{\pm}(\boldsymbol{y}) &= \left(w(\|\boldsymbol{t}_j\|) \frac{\partial^k}{\partial t_1} \sin(\boldsymbol{t}\cdot\boldsymbol{y}) \pm \epsilon |y_1|^k \right) \mathbb{I}\{y_1 \in [-M, M]\} \pm \|w\|_\infty |y_1|^k \mathbb{I}\{y_1 \in [-M, M]^c\}. \end{aligned}$$

For these functions, we can show that

$$\begin{aligned} E[(g_j^+(\boldsymbol{Y}) - g_j^-(\boldsymbol{Y}))^2] &\le 4\epsilon^2 (E[|Y_1|^{2k}] + \|w\|_\infty^2), \\ E[(h_j^+(\boldsymbol{Y}) - h_j^-(\boldsymbol{Y}))^2] &\le 4\epsilon^2 (E[|Y_1|^{2k}] + \|w\|_\infty^2). \end{aligned}$$

To obtain above inequalities, we used the definition of $M = M(\epsilon, k)$. It remains to choose the grid points $\boldsymbol{t}_j$ in such a way that the brackets cover the set $\mathbb{G}_{1,k}$. Let $Lip(w)$ be the Lipschitz constant of the weight function w. For an arbitrary $\boldsymbol{t} \in \mathbb{R}^d$ and any grid point $\boldsymbol{t}_j$ we have that

$$\begin{aligned} & \left| w(\|\boldsymbol{t}\|) \frac{\partial^k}{\partial t_1} \cos(\boldsymbol{t}\cdot\boldsymbol{y}) - w(\|\boldsymbol{t}_j\|) \frac{\partial^k}{\partial t_{j,1}} \cos(\boldsymbol{t}_j\cdot\boldsymbol{y}) \right| \\ \le\ & |y_1|^k Lip(w) |\|\boldsymbol{t}\| - \|\boldsymbol{t}_j\|| + \|w\|_\infty |y_1|^{k+1} \|\boldsymbol{t} - \boldsymbol{t}_j\| \le |y_1|^k (Lip(w) + \|w\|_\infty |y_1|) \|\boldsymbol{t} - \boldsymbol{t}_j\|, \end{aligned}$$

and also have that

$$\left| w(\|\boldsymbol{t}\|) \frac{\partial^k}{\partial t_1} \cos(\boldsymbol{t}\cdot\boldsymbol{y}) - w(\|\boldsymbol{t}_j\|) \frac{\partial^k}{\partial t_{j,1}} \cos(\boldsymbol{t}_j\cdot\boldsymbol{y}) \right| \le |y_1|^k (w(\|\boldsymbol{t}\|) + w(\|\boldsymbol{t}_j\|)). \tag{23}$$

Therefore, the function $\boldsymbol{y} \mapsto w(\|\boldsymbol{t}\|) \frac{\partial^k}{\partial t_1} \cos(\boldsymbol{t}\cdot\boldsymbol{y})$ is contained in the bracket $[g_j^-, g_j^+]$ if

$$(Lip(w) + \|w\|_\infty M) \|\boldsymbol{t} - \boldsymbol{t}_j\| \le \epsilon.$$

Consequently, we choose the grid points as

$$\boldsymbol{t}_j = \epsilon \boldsymbol{z}_j / (Lip(w) + \|w\|_\infty M(\epsilon, k)),\ \ \boldsymbol{z}_j \in \mathbb{Z}^d,$$

for $\|\boldsymbol{z}_j\| \le J(\epsilon)$, where $J(\epsilon)$ is the smallest integer such that $\epsilon J(\epsilon)/(Lip(w) + \|w\|_\infty M(\epsilon, k))$ is greater than or equal to

$$U(\epsilon) = \inf\{a > 0 : \sup_{|v|\ge a} w(v) \le \epsilon/2\}.$$

Together with this and (23) yield that $N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(c)}) \le (2J(\epsilon) + 1)^d$ (we can show the same bound for $N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(s)})$). Then we have that

$$N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}) \le N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(c)}) + N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(s)}) \le 2(2J(\epsilon) + 1)^d.$$

If follows from the Markov inequality that

$$M(\epsilon, k) \le (E[|Y_1|^{2k+\eta}]/\epsilon^2)^{1/\eta}.$$

From the definition of $J(\epsilon)$, have that

$$\frac{\epsilon J(\epsilon)}{2(Lip(w) + \|w\|_\infty M(\epsilon, k))} \le U(\epsilon).$$

Therefore we obtain the inequality

$$J(\epsilon) \leq 2U(\epsilon)(Lip(w) + \|w\|_\infty M(\epsilon, k))/\epsilon + 1.$$

Then we have that $\log(N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k})) = O(\log J(\epsilon)) = O(\epsilon^{-(\delta+1/2)^{-1}} + \log(\epsilon^{-1-2/\eta})) = O(\epsilon^{-\kappa})$ for $\kappa = (\delta + 1/2)^{-1} < 2$. This implies that $J(\delta, \mathbb{G}_{1,k}) < \infty$ as required.

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] Adusumilli, K. and T. Otsu (2018) Nonparametric instrumental regression with errors in variables, *Econometric Theory*, 34, 1256-1280.

[3] Arcidiacono, P., Aucejo, E. M., Fang, H. and K. I. Spenner (2011) Does affirmative action lead to mismatch? A new test and evidence, *Quantitative Economics*, 2, 303-333.

[4] Arellano, M. and S. Bonhomme (2012) Identifying distributional characteristics in random coefficients panel data models, *Review of Economic Studies*, 79, 987-1020.

[5] Athey, S. and P. A. Haile (2007) Nonparametric approaches to auctions, in Heckman, J. J. and E. E. Leamer (eds.), *Handbook of Econometrics*, vol. 6A, 3847-3965.

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

[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] Evdokimov, K. (2010) Identification and estimation of a nonparametric panel data model with unobserved heterogeneity, Working paper.

[10] Evdokimov, K. and H. White (2012) Some extensions of a lemma of Kotlarski, *Econometric Theory*, 2012, 28, 925-932.

[11] Fan, J. (1991) On the optimal rates of convergence for non-parametric deconvolution problems, *Annals of Statistics*, 19, 1257-1272.

[12] Hall, P. and A. Meister (2007) A ridge-parameter approach to deconvolution, *Annals of Statistics*, 35, 1535-1558.

[13] Hu, Y. (2017) The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics, *Journal of Econometrics*, 200, 154-168.

[14] Kato, K. and D. Kurisu (2020) Bootstrap confidence bands for spectral estimation of Lévy densities under high-frequency observations, *Stochastic Processes and their Applications*, 130, 1159-1205.

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

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

[17] Kurisu, D. (2019) Nonparametric inference on Lévy measures of compound Poisson-driven Ornstein-Uhlenbeck processes under macroscopic discrete observations, *Electronic Journal of Statistics*, 13, 2521-2565.

[18] Li, T. (2002) Robust and consistent estimation of nonlinear errors-in-variables models, *Journal of Econometrics*, 110, 1-26.

[19] Li, T. and C. Hsiao (2004) Robust estimation of generalized linear models with measurement errors, *Journal of Econometrics*, 118, 51-65.

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

[21] Neumann, M. and M. Reiss (2009) Nonparametric estimation for Lévy processes from low-frequency observations, *Bernoulli*, 15, 223-248.

[22] Otsu, T. and L. N. Taylor (2019) Specification testing for errors-in-variables models, forthcoming in *Econometric Theory*.

[23] Tsybakov, A. B. (2009) *Introduction to Nonparametric Estimation*, Springer.

[24] van der Vaart, A. W. (1998) *Asymptotic Statistics*, Cambridge University Press.

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`