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 measurements are available. Our assumptions are weaker than existing results, such as Li and Vuong (1998) which requires bounded support, and a specialization of Bonhomme and Robin (2010) which requires the existence of moment generating functions of certain observables. Moreover, our uniform convergence rates are typically faster than those obtained in these papers.

1. Introduction

This paper studies uniform convergence rates of nonparametric deconvolution estimators for the classical measurement error model, where repeated measurements 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 final object 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 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) for nonparametric instrumental regression with errors-in-variables, and Otsu and Taylor (2018) for specification testing on errors-in-variables regressions. For those purposes, the LV-type estimators play the same role as primitive nonparametric estimators for semiparametric problems. Thus, it is crucial

Our research is supported by Grant-in-Aid for Research Activity Start-up (19K20881) from the JSPS (Kurisu) and ERC Consolidator Grant (SNP 615882) (Otsu).

to equip uniform convergence rate results for 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). Comte and Kappus (2015) 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 rate under weaker assumptions than the ones in LV. Importantly, Comte and Kappus (2015) 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 Comte and Kappus (2015) 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. Since our theoretical development is a specialization for the repeated measurements model, our convergence rates are faster than those given in Bonhomme and Robin (2010) under weaker assumptions. In particular, we do not require the existence of moment generating functions of certain observables as in Bonhomme and Robin (2010). The relaxation of this assumption is achieved by a multivariate version of Neumann and Reiss (2009, Theorem 4.1) on the normalized empirical characteristic function process (Lemma 3 in Appendix C below). This lemma can also be used in Bonhomme and Robin (2010) to relax their assumptions in other contexts, and thus is of independent interest.

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 $E[\epsilon] = 0$.*

[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.

These assumptions are standard for the classical measurement error model (e.g., Comte and Kappus, 2015). 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 remarks for our main theorem below for a detailed discussion.

This paper studies the uniform convergence rates of the LV and Comte and Kappus (2015) estimators for the densities and characteristic functions of X and ϵ. Let us first introduce the LV estimator. Define

$$\psi(u_1, u_2) = E[e^{\mathrm{i}(u_1Y_1+u_2Y_2)}] = \varphi_X(u_1 + u_2)\varphi_\epsilon(u_1)\varphi_\epsilon(u_2).$$

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 $\hat{\psi}(u_1, u_2) = \frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}(u_1Y_{j,1}+u_2Y_{j,2})}$ and $\frac{\partial\hat{\psi}(u_1,u_2)}{\partial u_1} = \frac{1}{n}\sum_{j=1}^n \mathrm{i}Y_{j,1}e^{\mathrm{i}(u_1Y_{j,1}+u_2Y_{j,2})}$. 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 $\hat{\varphi}_\epsilon$.

We next introduce a regularized version of the LV estimator developed by Comte and Kappus (2015). 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 Comte and Kappus (2015) 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 also 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, Comte and Kappus (2015) investigated the risk bounds and convergence rates for the L_2-loss function. In this paper, we study the uniform convergence rates of Comte and Kappus' (2015) estimators.

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

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, $\varphi_K(u) = 0$ for any $|u| > 1$.*

This assumption says that K is a p-th order kernel function. Below, we give some examples of kernel functions which satisfy our assumptions.

Remark 1 (Examples of kernel functions)**.** Construction of a kernel function satisfying Assumption K is typically done by specifying its Fourier transform φ_K. Let $\zeta : \mathbb{R} \to \mathbb{R}$ be a function that is even (i.e., $\zeta(u) = \zeta(-u)$), supported on $[-1, 1]$, $(p+2)$-times continuously differentiable, and such that $\zeta^{(\ell)}(0) = 1$ 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, $|K(x)| = o(|x|^{-p-2})$ as $|x| \to \infty$ (which follows from a change of variables), so that $(1 \vee |x|^p)K$ is integrable, and

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

Here, 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

$$\zeta(u) = \begin{cases} 1 & \text{if } |u| \leq c_0, \\ \exp\left\{ \frac{-b\exp(-b/(|u|-c_0)^2)}{(|u|-1)^2} \right\} & \text{if } c_0 < |u| < 1, \\ 0 & \text{if } 1 \leq |u|, \end{cases} \tag{6}$$

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 proceed, we consider the following two scenarios for 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*,* ρ_ϵ*,* $C_\epsilon \geq c_\epsilon$*,* ω_ϵ*, and some constants* $\beta_x, \beta_\epsilon \in \mathbb{R}$*, it holds*

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

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.

Our main results are presented as follows.

Theorem. *Suppose that Assumption M holds true, and* $E|Y_1|^{2+\eta} < \infty$ *for some* $\eta > 0$*.*

(i): *Under Assumption OS and*

$$n^{-1/2} T_n^{3\beta_x + 2\beta_\epsilon + 1} \log T_n \to 0 \textit{ as } n \to \infty,$$

it holds

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

Additionally, suppose that Assumption K holds for $p \geq \max\{\beta_x, \beta_\epsilon\}$*. Set* $T_n = h^{-1}$*. Then*

$$\sup_{|x| \leq h^{-1}} |\hat{f}_X(x) - f_X(x)| = O_p\left(n^{-1/2} h^{-2\beta_x - 2\beta_\epsilon - 2} \log h^{-1} + h^{\beta_x - 1}\right),$$
$$\sup_{|x| \leq h^{-1}} |\hat{f}_\epsilon(x) - f_\epsilon(x)| = O_p\left(n^{-1/2} h^{-3\beta_x - 2\beta_\epsilon - 2} \log h^{-1} + h^{\beta_\epsilon - 1}\right).$$

(ii): *Under Assumptions SS and*

$$n^{-1/2} T_n^{1 - 3\beta_x - 2\beta_\epsilon} (\log T_n) \exp\left(\frac{3T_n^{\rho_x}}{\mu_x} + \frac{2T_n^{\rho_\epsilon}}{\mu_\epsilon}\right) \to 0 \textit{ as } n \to \infty,$$

it holds

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

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

Additionally, suppose that Assumption K holds and that there exists $0<c\le 1$ *such that* $\varphi_K(x)=1$ *for* $|x|\le c$. *Set* $T_n=h^{-1}$. *Then*

$$\sup_{|x|\le h^{-1}} |\hat{f}_X(x)-f_X(x)|$$
$$= O_p\left(n^{-1/2}h^{2\beta_x+2\beta_\epsilon-2}(\log h^{-1})\exp\left(\frac{2h^{-\rho_x}}{\mu_x}+\frac{2h^{-\rho_\epsilon}}{\mu_\epsilon}\right)+h^{\rho_x/q-\beta_x-1}\exp\left(-\frac{c^{\rho_x}h^{-\rho_x}}{\mu_x}\right)\right),$$

$$\sup_{|x|\le h^{-1}} |\hat{f}_\epsilon(x)-f_\epsilon(x)|$$
$$= O_p\left(n^{-1/2}h^{3\beta_x+2\beta_\epsilon-2}(\log h^{-1})\exp\left(\frac{3h^{-\rho_x}}{\mu_x}+\frac{2h^{-\rho_\epsilon}}{\mu_\epsilon}\right)+h^{\rho_\epsilon/q-\beta_\epsilon-1}\exp\left(-\frac{c^{\rho_\epsilon}h^{-\rho_\epsilon}}{\mu_\epsilon}\right)\right),$$

where $q=1$ *when* $\beta_x,\beta_\epsilon>0$, *and* $q>1$ *when* $\beta_x,\beta_\epsilon\le 0$.

(iii): *The same uniform convergence results in (i) and (ii) hold true even if we replace the LV estimator* $(\hat{\varphi}_X,\hat{\varphi}_\epsilon,\hat{f}_X,\hat{f}_\epsilon)$ *with the Comte and Kappus (2015) estimator* $(\tilde{\varphi}_X,\tilde{\varphi}_\epsilon,\tilde{f}_X,\tilde{f}_\epsilon)$.

Remark 2. To obtain the uniform consistency of $\hat{\varphi}_X$ (or $\tilde{\varphi}_X$) on $[-T_n,T_n]$, it is sufficient to assume

$$n^{-1/2}T_n^{2\beta_x+2\beta_\epsilon+1}\log T_n\to 0 \qquad \text{under Assumption OS,}$$
$$n^{-1/2}T_n^{1-2\beta_x-2\beta_\epsilon}(\log T_n)\exp\left(\frac{2T_n^{\rho_x}}{\mu_x}+\frac{2T_n^{\rho_\epsilon}}{\mu_\epsilon}\right)\to 0 \qquad \text{under Assumption SS,}$$

as $n\to\infty$. Note that we can take $c=1$ in Theorem (ii) if we use the sinc kernel function. We can also derive uniform convergence rates when X is ordinary smooth and ϵ is super smooth and vise versa since the decay rates of characteristic functions are not essential in our proofs. See Appendix A for details.

Remark 3 (Comparison with LV)**.** We note that LV established the uniform convergence rates of their estimators under the assumption that both X and ϵ have bounded support. On the other hand, our theorem does not require such boundedness. Also, the convergence rates obtained in our theorem 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 (i) implies that

$$\sup_{u\in[-T_n,T_n]} |\hat{\varphi}_X(u)-\varphi_X(u)| = O_p\left(\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\alpha-\frac{\alpha}{2(1+\beta_x+\beta_\epsilon)}}\right),$$

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 4 (Comparison with Bonhomme and Robin, 2010)**.** The convergence rates in our theorem are also faster than those given in Bonhomme and Robin (2010). For example, under

Assumption OS, Bonhomme and Robin (2010, Theorem 1) implies that

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

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 multivariate version of Neumann and Reiss (2009, Theorem 4.1) (see Lemma 3 in Appendix C below for details).

Remark 5. In this theorem, we obtained the same uniform convergence rates for the LV and Comte and Kappus (2015) estimators. On the other hand, it is open whether the LV estimator can achieve the L_2 convergence rate in Comte and Kappus (2015). 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 Comte and Kappus (2015).

Remark 6 (Multivariate version of Neumann and Reiss, 2009, Theorem 4.1)**.** Finally, we emphasize that Lemma 3 in the present paper 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 (2017) and Kurisu (2018), which study nonparametric inference on univariate Lévy (driven stochastic) processes under high- and low-frequency observations, to multivariate cases.

Appendix A. Proof of theorem

Here we only present the proof of Theorem (iii) for Comte and Kappus' (2015) estimator since the proofs of Theorem (i) and (ii) are its specialization. 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 $\tilde{\varphi}_X$.** Observe that

$$
\begin{aligned}
|\tilde{\varphi}_X(u) - \varphi_X(u)| &\le |\tilde{\varphi}_X(u) - \varphi_X(u)|\mathbb{I}\{|\Delta(u)| \le 1\} + |\tilde{\varphi}_X(u) - \varphi_X(u)|\mathbb{I}\{|\Delta(u)| > 1\} \\
&\le |\tilde{\varphi}_X^{mod}(u) - \varphi_X(u)|\mathbb{I}\{|\Delta(u)| \le 1\} + 2|\Delta(u)|\mathbb{I}\{|\Delta(u)| > 1\} \\
&= |\varphi_X(u)||1 - e^{\Delta(u)}|\mathbb{I}\{|\Delta(u)| \le 1\} + 2|\Delta(u)|\mathbb{I}\{|\Delta(u)| > 1\} \\
&\le 2|\varphi_X(u)||\Delta(u)|\mathbb{I}\{|\Delta(u)| \le 1\} + 2|\Delta(u)|\mathbb{I}\{|\Delta(u)| > 1\} \\
&\le 2(1 + |\varphi_X(u)|)|\Delta(u)|, \qquad (7)
\end{aligned}
$$

where the second inequality follows from the facts that $|\tilde{\varphi}_X(u) - \varphi_X(u)| \le |\tilde{\varphi}_X^{mod}(u) - \varphi_X(u)|$ and $|\tilde{\varphi}_X(u) - \varphi_X(u)| \le 2$, 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

$$
\begin{aligned}
\sup_{u\in[-T_n,T_n]} |\Delta_1(u)| &\le \sup_{u\in[-T_n,T_n]} |R_2(u)| \left(\int_0^{T_n} \frac{1}{|\psi(0,u_2)|} du_2\right), \\
\sup_{u\in[-T_n,T_n]} |\Delta_2(u)| &\le \sup_{u\in[-T_n,T_n]} |R_1(u)| \left(\int_0^{T_n} \left|\frac{\partial\psi(0,u_2)}{\partial u_1}\right| du_2\right) \\
&= \sup_{u\in[-T_n,T_n]} |R_1(u)| \left(\int_0^{T_n} |E[Y_1 e^{iu_2 Y_2}]| du_2\right) \le T_n E[|Y_1|] \sup_{u\in[-T_n,T_n]} |R_1(u)|, \\
\sup_{u\in[-T_n,T_n]} |\Delta_3(u)| &\le T_n \sup_{u\in[-T_n,T_n]} |R_1(u)| \sup_{u\in[-T_n,T_n]} |R_2(u)|.
\end{aligned}
$$

Therefore, the conclusion follows from Lemmas 1 and 2.

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

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

and this implies that

$$
\sup_{u\in[-T_n,T_n]}|\bar{\varphi}_X(u)-\varphi_X(u)| = O\left(\sup_{u\in[-T_n,T_n]}|\tilde{\varphi}_X(u)-\varphi_X(u)|\right). \tag{8}
$$

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 $|u|\le T_n$,

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

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

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

which yields

$$
\sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\psi(0,u))\right| = O\left(\sup_{u\in[-T_n,T_n]}|\zeta(u)|\right) = O_p\left(n^{-1/2}T_n^{\beta_x+\beta_\epsilon}\log T_n\right). \tag{9}
$$

Likewise, we can show that

$$
\sup_{u\in[-T_n,T_n]}|\log(\tilde{\varphi}_X(u)/\varphi_X(u))| = O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\tilde{\varphi}_X(u)-\varphi_X(u)}{\varphi_X(u)}\right|\right) = O_p\left(n^{-1/2}T_n^{3\beta_x+2\beta_\epsilon+1}\log T_n\right). \tag{10}
$$

Together with (8), (9), and (10), we have that

$$
\begin{aligned}
& \sup_{u\in[-T_n,T_n]}|\log\tilde{\varphi}_\epsilon(u)-\log\varphi_\epsilon(u)| \\
=\ & \sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\bar{\varphi}_X(u))-\log(\psi(0,u)/\varphi_X(u))\right| \\
\le\ & \sup_{u\in[-T_n,T_n]}\left|\log(\hat{\psi}(0,u)/\psi(0,u))\right|+\sup_{u\in[-T_n,T_n]}|\log(\bar{\varphi}_X(u)/\varphi_X(u))| \\
=\ & O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\hat{\psi}(0,u)-\psi(0,u)}{\psi(0,u)}\right|\right)+O\left(\sup_{u\in[-T_n,T_n]}\left|\frac{\tilde{\varphi}_X(u)-\varphi_X(u)}{\varphi_X(u)}\right|\right) \\
=\ & O_p\left(n^{-1/2}T_n^{3\beta_x+2\beta_\epsilon+1}\log T_n\right) = o_p(1).
\end{aligned} \tag{11}
$$

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

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

provided $\sup_{u\in[-T_n,T_n]} |\log\tilde{\varphi}_\epsilon(u) - \log\varphi_\epsilon(u)| < 1$. Therefore, (11) yields the desired result.

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

$$\begin{aligned}
& |\tilde{f}_X(x) - f_X(x)| \\
= & \left|\frac{1}{2\pi}\int_{-h^{-1}}^{h^{-1}} e^{-\mathrm{i}ux}\{\tilde{\varphi}_X(u) - \varphi_X(u)\}\varphi_K(hu)du + \frac{1}{2\pi}\int e^{-\mathrm{i}ux}\varphi_X(u)\{\varphi_K(hu) - 1\}du\right| \\
\leq & C_K\frac{h^{-1}}{\pi}\sup_{u\in[-h^{-1},h^{-1}]} |\tilde{\varphi}_X(u) - \varphi_X(u)| + \frac{1}{2\pi}\int |\varphi_X(u)||\varphi_K(hu) - 1|du, \qquad (12)
\end{aligned}$$

where the inequality follows from $|e^{-\mathrm{i}ux}| = 1$ and $\sup_{x\in\mathbb{R}} |\varphi_K(x)| \leq C_K$ for some positive constant $C_K < \infty$. By the first part of this theorem, the first term of (12) 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, there exists a function m such that $\varphi_K(x) = 1 + m(x)x^p$ for all $x \in [-1, 1]$ and $\varphi_K(x) = 0$ for $|x| > 1$, where m is continuous on $[-1, 1]$. Therefore, the second term of (12) satisfies

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

where the first inequality follows from Assumption OS, and the second inequality follows from the assumption on K. Moreover, under Assumptions SS and K with $\varphi_K(x) = 1$, $|x| \leq c$ for some $0 < c \leq 1$, there exists a positive constant C_0 such that

$$\begin{aligned}
\int |\varphi_X(u)||\varphi_K(hu) - 1|du &\leq C_0C_x\int_{ch^{-1}}^{\infty} |u|^{\beta_x}e^{-|u|^{\rho_x}/\mu_x}du \\
&= \begin{cases} O\left(h^{-\beta_x-1+\rho_x/q}\exp\left(-\frac{c^{\rho_x}h^{-\rho_x}}{\mu_x}\right)\right) & \text{if } \beta_x \leq 0, \\ O\left(h^{-\beta_x-1+\rho_x}\exp\left(-\frac{c^{\rho_x}h^{-\rho_x}}{\mu_x}\right)\right) & \text{if } \beta_x > 0, \end{cases}
\end{aligned}$$

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

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

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.

A.4. **Proof for $\tilde{f}_\epsilon$.** The proof is similar to the one in Section A.3.

Appendix B. Lemmas

Lemma 1. *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 3 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 3 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 2. *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

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

Together with this and Lemma 1 yields that

$$\sup_{u\in[-T_n,T_n]} |\tilde{\psi}(0,u) - \psi(0,u)| = O_p(n^{-1/2}\log T_n). \tag{13}$$

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

$$\inf_{u\in[-T_n,T_n]} |\hat{\psi}(0,u)| \ge \inf_{u\in[-T_n,T_n]} |\psi(0,u)| - \sup_{u\in[-T_n,T_n]} |\hat{\psi}(0,u) - \psi(0,u)| = O_p(T_n^{-\beta_x-\beta_\epsilon}). \tag{14}$$

Combining (13) and (14), we finally obtain that

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

□

Appendix C. Multivariate version of Neumann and Reiss (2009, Theorem 4.1)

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

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

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

Lemma 3. *Assume* $E\left[\left(\prod_{j=1}^d |Y_{1,j}|^{(k_j\vee 1/2)}\right)^{2+\eta}\right]$ *for some* $\eta > 0$. *Then*

$$\sup_{n\ge1} E[\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)}] < \infty.$$

Remark 7. Since

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

Lemma 3 implies that

$$E\left[\sup_{\|\boldsymbol{u}\|\in[-T_n,T_n]}\left|\frac{\partial^{|\boldsymbol{k}|}}{\partial u_1^{k_1}\cdots\partial u_d^{k_d}}(\hat{\psi}(\boldsymbol{u})-\psi(\boldsymbol{u}))\right|\right] \le \frac{\sup_{n\ge1}E[\|C_n^{(\boldsymbol{k})}\|_{L_\infty(w)}]}{\sqrt{n}\inf_{\|\boldsymbol{u}\|\le T_n}w(\|\boldsymbol{u}\|)} = O\left(n^{-1/2}\log T_n\right).$$

Then we can show

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

from Markov's inequality.

Proof of Lemma 3. We prove the case when $\boldsymbol{k} = (k,0,\dots,0)'$. 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 $\ell \le f \le u$. 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 $E[(u_j(\boldsymbol{Y}) - \ell_j(\boldsymbol{Y}))^2] \le \epsilon^2$, 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

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

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

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

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

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

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

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

and also have that

$$\left| w(\|\boldsymbol{t}\|) \frac{\partial^k}{\partial t_1} \cos(\boldsymbol{t} \cdot \boldsymbol{y}) - w(\|\boldsymbol{t}_j\|) \frac{\partial^k}{\partial t_{j,1}} \cos(\boldsymbol{t}_j \cdot \boldsymbol{y}) \right| \leq |y_1|^k (w(\|\boldsymbol{t}\|) + w(\|\boldsymbol{t}_j\|)). \tag{15}$$

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

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

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 $\|\boldsymbol{z}_j\| \leq J(\epsilon)$, 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

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

Together with this and (15) yield that $N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(c)}) \leq (2J(\epsilon) + 1)^d$ (we can show the same bound for $N_{[\cdot]}(\epsilon, \mathbb{G}_{1,k}^{(s)})$). Then we have that

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

If follows from the Markov inequality that

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

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

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

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

[2] 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.

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

[4] 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.

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

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

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

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

[9] 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.

[10] Kato, K. and D. Kurisu (2017) Bootstrap confidence bands for spectral estimation of Lévy densities under high-frequency observations, forthcoming in *Stochastic Processes and Their Applications*.

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

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

[13] Kurisu, D. (2018) Nonparametric inference on Lévy measures of Lévy-driven Ornstein-Uhlenbeck processes under discrete observations, arXiv:1803.08671.

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

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

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

[19] Otsu, T. and L. N. Taylor (2018) Specification testing for errors-in-variables models, Working paper.

[20] 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