Kernel entropy estimation for linear processes II

Yudan Xiong and Fangjun Xu*

March 29, 2024

Abstract

Let $X = \{X_n : n \in \mathbb{N}\}$ be a linear process with bounded probability density function $f(x)$. Under certain conditions, we use the kernel estimator

$$\frac{2}{n(n-1)h_n} \sum_{1\le i<j\le n} K\Big(\frac{X_i - X_j}{h_n}\Big)$$

to estimate the quadratic functional of $\int_{\mathbb{R}} f^2(x)dx$ of the linear process $X = \{X_n : n \in \mathbb{N}\}$ and improve the corresponding results in [4].

Keywords: linear process, kernel entropy estimation, quadratic functional, projection operator

2010 Mathematics Subject Classification: Primary 60F05, 62M10; Secondary 60G10, 62G05.

1 Introduction

Let $X = \{X_n : n \in \mathbb{N}\}$ be a linear process defined by

$$X_n = \sum_{i=0}^{\infty} a_i \varepsilon_{n-i}, \tag{1.1}$$

where the innovations ε_i are independent and identical distributed (i.i.d.) real-valued random variables in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a_i are real coefficients such that $\sum\limits_{i=0}^{n} a_i\varepsilon_{n-i}$ converges in distributions. Assume that the linear process $X = \{X_n : n \in \mathbb{N}\}$ has a bounded density function f. Clearly, the quadratic functional $\int_{\mathbb{R}} f^2(x)\, dx$ is well-defined and it plays an important role in the study of the entropies of the linear process $X = \{X_n : n \in \mathbb{N}\}$, say, quadratic Rényi entropy $R(f) = -\ln(\int_{\mathbb{R}} f^2(x)\, dx)$ and Shannon entropy $S(f) = -\int_{\mathbb{R}} f(x) \ln f(x)\, dx$, see [4] and references therein. To estimate the quadratic functional $\int_{\mathbb{R}} f^2(x)\, dx$ of the linear process $X = \{X_n : n \in \mathbb{N}\}$ defined in (1.1), the authors in [4] used the kernel estimator

$$T_n(h_n) = \frac{2}{n(n-1)h_n} \sum_{1\le j<i\le n} K\Big(\frac{X_i - X_j}{h_n}\Big), \tag{1.2}$$

where the kernel K is a symmetric and bounded function with $\int_{\mathbb{R}} K(u)\, du = 1$ and $\int_{\mathbb{R}} u^2|K(u)|\, du < \infty$. The bandwidth sequence h_n satisfies $0 < h_n \to 0$ as $n \to \infty$. Then, under certain conditions,

*F. Xu is partially supported by National Natural Science Foundation of China (Grant No.12371156).

asymptotic properties of the estimator $T_n(h_n)$ were established, see Theorems 2.1 and 2.2 in [4]. In this paper, we are going to improve these two theorems. Some ideas are borrowed from [3].

Let $\phi_\varepsilon(\lambda)$ be the characteristic function innovations. That is, for each $\lambda \in \mathbb{R}$, $\phi_\varepsilon(\lambda) = \mathbb{E}[e^{\iota\lambda\varepsilon_1}]$ where ι is the imaginary unit $\sqrt{-1}$. Moreover, for any integrable function $g(x)$, its Fourier transform is defined as $\widehat{g}(u) = \int_{\mathbb{R}} e^{\iota x u} g(x)\, dx$. The following is our main result.

Theorem 1.1 *For some $\gamma \in (0, 1]$, assume that*

(A1) *there are at least three coefficients $\{a_i : i \geq 0\}$ of the linear process X are non-zero and $\sum\limits_{i=0}^{\infty} |a_i|^\gamma < \infty$,*

(A2) *there exists a positive constant c_γ such that $\mathbb{E}\,|e^{\iota\lambda\varepsilon_1} - \phi_\varepsilon(\lambda)|^2 \leq c_\gamma \left(|\lambda|^{2\gamma} \wedge 1\right)$ for all $\lambda \in \mathbb{R}$,*

(A3) *(i) $\int_{\mathbb{R}} |\lambda|^{3\gamma} |\phi_\varepsilon(\lambda)|^2\, d\lambda < \infty$ or (ii) $\int_{\mathbb{R}} |\lambda|^{2\gamma} |\phi_\varepsilon(\lambda)|^2\, d\lambda < \infty$ and $\sup\limits_{\lambda \in \mathbb{R}} |\widehat{K}(\lambda)||\lambda|^\gamma < \infty$.*

Then there exist positive constants c_1 and c_2 such that

$$\left|\mathbb{E}\, T_n(h_n) - \int_{\mathbb{R}} f^2(x)\, dx\right| \leq c_1 \Big(\frac{1}{n} + h_n^{2\gamma}\Big), \tag{1.3}$$

$$\mathbb{E}\left|T_n(h_n) - \mathbb{E}\, T_n(h_n) - \frac{1}{n}\sum_{i=1}^n Y_i\right| \leq c_2 \Big(\frac{1}{\sqrt{n^3 h_n^2}} + \frac{1}{\sqrt{n^2 h_n}} + \frac{h_n^\gamma}{\sqrt{n}}\Big), \tag{1.4}$$

where $Y_i = 2\big(f(X_i) - \mathbb{E}\, f(X_i)\big)$ for $1 \leq i \leq n$, and, if additionally $nh_n \to \infty$ as $n \to \infty$,

$$\sqrt{n}\,\Big[T_n(h_n) - \mathbb{E}\, T_n(h_n)\Big] \xrightarrow{\mathcal{L}} N(0, 4\sigma^2) \tag{1.5}$$

for some $\sigma^2 \in (0, \infty)$.

Remark: (i) If at most two coefficients of the linear process X are non-zero, then X is an i.i.d or 2-dependent sequence. The corresponding results are available in [1, 2, 4].

(ii) By Assumptions **(A1)** and **(A3)**, we could easily obtain that $\mathbb{E}[e^{\iota u X_n}] = \prod\limits_{i=0}^{\infty} \phi_\varepsilon(u a_i) \in L^1(\mathbb{R})$. So the probability density function $f(x)$ of the underlying linear process X is bounded.

Next we compare our result with those in [4]. The assumption (3)

$$\mathbb{E}\,|e^{\iota\lambda\varepsilon_1} - \phi_\varepsilon(\lambda)|^2 \leq c_{\gamma,4} \left(|\lambda|^{4\gamma} \wedge 1\right)$$

in Theorem 2.1 of [4] is weaken to our assumption **(A2)**. According to Example 3.2 in [4], our Theorem 1.1 works for more linear processes than Theorem 2.1 in [4]. The result (9)

$$\mathbb{E}\left|T_n(h_n) - \mathbb{E}\, T_n(h_n) - \frac{1}{n}\sum_{i=1}^n Y_i\right| \leq c_4 \Big(\frac{1}{nh_n} + \frac{h_n^\gamma}{\sqrt{n}}\Big)$$

in Theorem 2.2 of [4] is improved to (1.4) in our Theorem 1.1. Moreover, to obtain the central limit theorem in (1.5), we only require the bandwidth to satisfy $\lim\limits_{n\to\infty} nh_n = \infty$ instead of $\lim\limits_{n\to\infty} \sqrt{n} h_n = \infty$ in Theorem 2.2 of [4].

As a consequence of Theorem 1.1, we can easily obtain the following result.

Corollary 1.2 *Under the assumptions* **(A1)**, **(A2)** *and* **(A3)** *in Theorem 1.1. Let h_n have the order* $\begin{cases} n^{-\frac{3}{4\gamma+2}} & \text{for} \quad \gamma \in (0,1/4], \\ n^{-\frac{2}{4\gamma+1}} & \text{for} \quad \gamma \in (1/4,1]. \end{cases}$ *If $0 < \gamma \le 1/4$, then*

$$T_n(h_n) - \int_{\mathbb{R}} f^2(x)\,dx = O_{\mathbb{P}}(n^{-\frac{3\gamma}{2\gamma+1}});$$

if $1/4 < \gamma \le 1$, then

$$\sqrt{n}\Big[T_n(h_n) - \int_{\mathbb{R}} f^2(x)\,dx\Big] \xrightarrow{\mathcal{L}} N(0, 4\sigma^2),$$

and if we further assume that the kernel function K is nonnegative, the estimator $-\ln(\frac{1}{n}+T_n(h_n))$ of the quadratic Rényi entropy $R(f) = -\ln(\int_{\mathbb{R}} f^2(x)dx)$ satisfies

$$\sqrt{n}\Big[-\ln(\frac{1}{n}+T_n(h_n)) - R(f)\Big] \xrightarrow{\mathcal{L}} N\Big(0, \frac{4\sigma^2}{(\int_{\mathbb{R}} f^2(x)dx)^2}\Big)$$

for some $\sigma^2 \in (0,\infty)$.

Throughout this paper, if not mentioned otherwise, the letter c with or without a subscript denotes a generic positive finite constant whose exact value is independent of n and may change from line to line. For a complex number z, we use $\overline{z}$ and $|z|$ to denote its conjugate and modulus, respectively. Moreover, we let $\phi(\lambda)$ be the characteristic function of linear process $X = \{X_n : n \in \mathbb{N}\}$. That is, $\phi(\lambda) = \mathbb{E}[e^{\iota\lambda X_n}]$ for each $\lambda \in \mathbb{R}$.

2 Proof of the main result

In this section, we will prove Theorem 1.1. Firstly, we give two important definitions. For each $i \in \mathbb{Z}$, let $\mathcal{F}_i$ be the σ-field generated by $\{\varepsilon_k : k \le i\}$. Given an integrable complex-valued random variable Z, we define the following projection operator $\mathscr{P}_i$ as

$$\mathscr{P}_i Z = \mathbb{E}[Z|\mathcal{F}_i] - \mathbb{E}[Z|\mathcal{F}_{i-1}]$$

for each $i \in \mathbb{Z}$. Then, for any two integrable complex-valued random variables Z and W, $\mathbb{E}[\mathscr{P}_i Z \mathscr{P}_j W] = 0$ if $i \ne j$.

For any $\delta \in (0,1)$, define

$$T_n^\delta(h_n) = \frac{2}{n(n-1)h_n} \sum_{1\le j<i\le n} K_\delta\Big(\frac{X_i - X_j}{h_n}\Big), \tag{2.1}$$

where $K_\delta(x) = \int_{\mathbb{R}} K(y)\frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(x-y)^2}{2\delta^2}}\,dy$.

It is easy to see that $K_\delta(x)$ is a symmetric and bounded function with $\int_{\mathbb{R}} K_\delta(u)\,du = 1$ and

$$\begin{aligned} &\int_{\mathbb{R}} u^2 |K_\delta(u)|\,du \\ &\le \int_{\mathbb{R}}\int_{\mathbb{R}} u^2|K(y)|\frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(u-y)^2}{2\delta^2}}\,dy\,du \\ &\le 2\int_{\mathbb{R}}\int_{\mathbb{R}} (u-y)^2|K(y)|\frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(u-y)^2}{2\delta^2}}\,du\,dy + 2\int_{\mathbb{R}}\int_{\mathbb{R}} y^2|K(y)|\frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(u-y)^2}{2\delta^2}}\,du\,dy \\ &\le 2\int_{\mathbb{R}} (1+|y|^2)|K(y)|dy < \infty. \end{aligned} \tag{2.2}$$

Moreover, for any $1 \le i < j \le n$, by the bounded convergence theorem,

$$\lim_{\delta\downarrow 0} \mathbb{E}\left| K\Big(\frac{X_i - X_j}{h_n}\Big) - K_\delta\Big(\frac{X_i - X_j}{h_n}\Big)\right|^2 = 0.$$

Therefore, $\limsup\limits_{\delta\downarrow 0} \mathbb{E}|T_n(h_n) - T_n^\delta(h_n)|^2 = 0$.

Proof of Theorem 1.1 The proof will be done in several steps.

Step 1. We give the estimation for $\left|\mathbb{E}\, T_n(h_n) - \int_{\mathbb{R}} f^2(x)\mathrm{d}x\right|$. Using the Fourier inverse transform, Plancherel formula, symmetry of $K_\delta(x)$ and $\int_{\mathbb{R}} K_\delta(u)du = 1$,

$$\begin{aligned}
\mathbb{E}\, T_n^\delta(h_n) - \int_{\mathbb{R}} f^2(x)\, dx &= \frac{1}{\pi n(n-1)} \sum_{1\le i<j\le n} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\, \mathbb{E}\left[e^{-\iota\lambda(X_i - X_j)}\right] d\lambda - \frac{1}{2\pi}\int_{\mathbb{R}} |\phi(\lambda)|^2\, d\lambda \\
&= \frac{1}{\pi n(n-1)} \sum_{1\le i<j\le n} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\, \mathbb{E}\left[(e^{-\iota\lambda X_i} - \phi(-\lambda))(e^{\iota\lambda X_j} - \phi(\lambda))\right] d\lambda \\
&\qquad + \frac{1}{2\pi}\int_{\mathbb{R}} \big(\widehat{K_\delta}(\lambda h_n) - \widehat{K_\delta}(0)\big)|\phi(\lambda)|^2\, d\lambda \\
&=: \mathrm{I}_1 + \mathrm{I}_2.
\end{aligned}$$

By Lemma 7.1 in [4], $\int_{\mathbb{R}} |\lambda|^{2\gamma}|\phi_\varepsilon(\lambda)|^2 < \infty$ and

$$|\widehat{K_\delta}(\lambda)| = |\widehat{K}(\lambda)e^{-\frac{1}{2}\delta^2\lambda^2}| \le |\widehat{K}(\lambda)| \le \int_{\mathbb{R}} |K(x)|dx < \infty$$

for all $\lambda \in \mathbb{R}$, we can obtain that I_1 is less than a constant multiple of $\frac{1}{n}$.

Moreover, by symmetry of $K_\delta(x)$ and $\int_{\mathbb{R}} K_\delta(u)du = 1$,

$$\left|\widehat{K_\delta}(\lambda h_n) - \widehat{K}_\delta(0)\right| = \left|\int_{\mathbb{R}} (e^{\iota\lambda h_n u} - 1)K_\delta(u)\, du\right| \le |\lambda|^{2\gamma} h_n^{2\gamma} \int_{\mathbb{R}} u^2 |K_\delta(u)|\, du. \tag{2.3}$$

Then, by $\int_{\mathbb{R}} |\lambda|^{2\gamma}|\phi_\varepsilon(\lambda)|^2 < \infty$ and (2.2), $|\mathrm{I}_2|$ is less than a constant multiple of $h_n^{2\gamma}$.

Hence,

$$\begin{aligned}
\left|\mathbb{E}\, T_n(h_n) - \int_{\mathbb{R}} f^2(x)\, dx\right| &\le \limsup_{\delta\downarrow 0}\left|\mathbb{E}\, T_n(h_n) - \mathbb{E}\, T_n^\delta(h_n)\right| + \limsup_{\delta\downarrow 0}\left|\mathbb{E}\, T_n^\delta(h_n) - \int_{\mathbb{R}} f^2(x)\, dx\right| \\
&\le c_1\Big(\frac{1}{n} + h_n^{2\gamma}\Big).
\end{aligned}$$

Step 2. We give the decomposition for $T_n^\delta(h_n) - \mathbb{E}\, T_n^\delta(h_n)$.

Recall the definition of $T_n^\delta(h_n)$ in (2.1). Then, by the Fourier inverse transform,

$$T_n^\delta(h_n) - \mathbb{E}\, T_n^\delta(h_n) = A_{n,\delta} + D_{n,\delta} - \mathbb{E}\, D_{n,\delta} + N_{n,\delta}, \tag{2.4}$$

where

$$\begin{aligned}
A_{n,\delta} &= \frac{2}{n(n-1)h_n} \sum_{1\le i<j\le n,\, j-i\le p_2} \Big[K_\delta\big(\frac{X_i - X_j}{h_n}\big) - \mathbb{E}\, K_\delta\big(\frac{X_i - X_j}{h_n}\big)\Big], \\
D_{n,\delta} &= \frac{1}{\pi n(n-1)} \sum_{1\le i<j\le n,\, j-i>p_2} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\big(e^{-\iota\lambda X_i} - \phi(-\lambda)\big)\big(e^{\iota\lambda X_j} - \phi(\lambda)\big)\, d\lambda, \\
N_{n,\delta} &= \frac{1}{\pi n(n-1)} \sum_{1\le i<j\le n,\, j-i>p_2} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\Big[\big(e^{-\iota\lambda X_i} - \phi(-\lambda)\big)\phi(\lambda) + \big(e^{\iota\lambda X_j} - \phi(\lambda)\big)\phi(-\lambda)\Big]\, d\lambda
\end{aligned}$$

and $p_2 = \min\{i > p_1 : a_i \neq 0\}$ with $p_1 = \min\{i \geq 1 : a_i \neq 0\}$. The finiteness of p_1 and p_2 follow from the assumption **(A1)**.

Step 3. We estimate $\mathbb{E}\,|A_{n,\delta}|^2$. For any $1 \leq i < j \leq n$, define

$$A_{n,\delta}^{i,j} = \frac{1}{h_n}\Big[K_\delta\big(\frac{X_i - X_j}{h_n}\big) - \mathbb{E}\,K_\delta\big(\frac{X_i - X_j}{h_n}\big)\Big].$$

Then

$$\begin{aligned}
\mathbb{E}\,|A_{n,\delta}|^2 &= \frac{4}{n^2(n-1)^2}\mathbb{E}\Big|\sum_{1\leq i<j\leq n,\, j-i\leq p_2} A_{n,\delta}^{i,j}\Big|^2 \\
&= \frac{4}{n^2(n-1)^2}\sum_{\substack{1\leq i_\theta<j_\theta\leq n,\, j_\theta-i_\theta\leq p_2\\ \theta=1,2}} \mathbb{E}\left[A_{n,\delta}^{i_1,j_1}A_{n,\delta}^{i_2,j_2}\right] \\
&\leq \frac{c_2}{n^3h_n^2} + \frac{c_2}{n^4}\sum_{\substack{1\leq i_\theta<j_\theta\leq n,\, j_\theta-i_\theta\leq p_2\\ \theta=1,2,\,|j_1-j_2|>4p_2}} \Big|\mathbb{E}\left[A_{n,\delta}^{i_1,j_1}A_{n,\delta}^{i_2,j_2}\right]\Big| \\
&\leq \frac{c_2}{n^3h_n^2} + \frac{c_2}{n^4}\sum_{\substack{1\leq i_\theta<j_\theta\leq n,\, j_\theta-i_\theta\leq p_2\\ \theta=1,2,\,|j_1-j_2|>4p_2}} \sum_{k=-\infty}^{\min\{j_1,j_2\}}\Big|\mathbb{E}\left[\mathscr{P}_k(A_{n,\delta}^{i_1,j_1})\mathscr{P}_k(A_{n,\delta}^{i_2,j_2})\right]\Big|. \qquad (2.5)
\end{aligned}$$

For any $1 \leq i < j \leq n$, it is easy to see that $X_i - X_j = \sum_{\ell=0}^{\infty} a_\ell^{i,j}\varepsilon_{j-\ell}$ where

$$a_\ell^{i,j} = \begin{cases} -a_\ell & \text{if } 0 \leq \ell < j-i, \\ a_{\ell-j+i} - a_\ell & \text{if } \ell \geq j-i. \end{cases} \qquad (2.6)$$

By the Fourier inverse transform,

$$\begin{aligned}
\mathscr{P}_k(A_{n,\delta}^{i,j}) &= \mathbb{E}\Big[K_\delta\big(\frac{X_i-X_j}{h_n}\big)|\mathcal{F}_k\Big] - \mathbb{E}\Big[K_\delta\big(\frac{X_i-X_j}{h_n}\big)|\mathcal{F}_{k-1}\Big] \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\Big(\prod_{\ell=0}^{j-k-1}\phi_\varepsilon(-\lambda a_\ell^{i,j})\Big)\big(e^{-\iota\lambda a_{j-k}^{i,j}\varepsilon_k} - \phi_\varepsilon(-\lambda a_{j-k}^{i,j})\big)e^{-\iota\lambda\sum_{\ell=j-k+1}^{\infty}a_\ell^{i,j}\varepsilon_{j-\ell}}\,d\lambda.
\end{aligned}$$

For the case $k < j_1 < j_2$, by Cauchy-Schwarz inequality and assumption **(A2)**,

$$\begin{aligned}
&\Big|\mathbb{E}\left[\mathscr{P}_k(A_{n,\delta}^{i_1,j_1})\mathscr{P}_k(A_{n,\delta}^{i_2,j_2})\right]\Big| \\
&\leq \int_{\mathbb{R}^2}|\widehat{K}(\lambda_1h_n)||\widehat{K}(\lambda_2h_n)||\phi_\varepsilon(a_0\lambda_1)||\phi_\varepsilon(a_0\lambda_2)| \\
&\qquad\times\Big|\mathbb{E}\left[(e^{-\iota\lambda_1a_{j_1-k}^{i_1,j_1}\varepsilon_k} - \phi_\varepsilon(-\lambda_1a_{j_1-k}^{i_1,j_1}))(e^{-\iota\lambda_1a_{j_2-k}^{i_2,j_2}\varepsilon_k} - \phi_\varepsilon(-\lambda_2a_{j_2-k}^{i_2,j_2}))\right]\Big|\,d\lambda_1\,d\lambda_2 \\
&\leq c_3\,|a_{j_1-k}^{i_1,j_1}|^\gamma|a_{j_2-k}^{i_2,j_2}|^\gamma\int_{\mathbb{R}^2}|\widehat{K}(\lambda_1h_n)||\widehat{K}(\lambda_2h_n)||\phi_\varepsilon(a_0\lambda_1)||\phi_\varepsilon(a_0\lambda_2)||\lambda_1|^\gamma|\lambda_2|^\gamma d\lambda_1\,d\lambda_2 \\
&= c_3\,|a_{j_1-k}^{i_1,j_1}|^\gamma|a_{j_2-k}^{i_2,j_2}|^\gamma\left(\int_{\mathbb{R}}|\widehat{K}(\lambda h_n)||\phi_\varepsilon(a_0\lambda)||\lambda|^\gamma\,d\lambda\right)^2 \\
&\leq c_3\,|a_{j_1-k}^{i_1,j_1}|^\gamma|a_{j_2-k}^{i_2,j_2}|^\gamma\int_{\mathbb{R}}|\widehat{K}(\lambda h_n)|^2\,d\lambda\int_{\mathbb{R}}|\phi_\varepsilon(a_0\lambda)|^2|\lambda|^{2\gamma}\,d\lambda \\
&\leq c_4\,h_n^{-1}\,|a_{j_1-k}^{i_1,j_1}|^\gamma|a_{j_2-k}^{i_2,j_2}|^\gamma.
\end{aligned}$$

Recall the definition of $a_\ell^{i,j}$ in (2.6) and $\sum_{i=0}^{\infty}|a_i|^\gamma<\infty$. If $j-i=p_1$ and $a_0=a_{p_1}$, then there exists $\ell_0\geq p_1$ such that $a_{\ell_0}^{i,j}\neq 0$. Now, for the case $k=j_1<j_2$, by Cauchy-Schwarz inequality and assumption **(A2)**,

$$\begin{aligned}
&\left|\mathbb{E}\left[\mathscr{P}_{j_1}(A_{n,\delta}^{i_1,j_1})\mathscr{P}_{j_1}(A_{n,\delta}^{i_2,j_2})\right]\right|\\
&\leq \int_{\mathbb{R}^2}|\widehat{K}(\lambda_1 h_n)||\widehat{K}(\lambda_2 h_n)||\phi_\varepsilon(a_0\lambda_2)||\phi_\varepsilon(a_p^{i_1,j_1}\lambda_1-a_{j_2-j_1+p}^{i_2,j_2}\lambda_2)|\\
&\qquad\times\left|\mathbb{E}\left[\left(e^{-\iota\lambda_1 a_0^{i_1,j_1}\varepsilon_{j_1}}-\phi_\varepsilon(-\lambda_1 a_0^{i_1,j_1})\right)\left(e^{-\iota\lambda_2 a_{j_2-j_1}^{i_2,j_2}\varepsilon_{j_1}}-\phi_\varepsilon(-\lambda_2 a_{j_2-j_1}^{i_2,j_2})\right)\right]\right|d\lambda_1\,d\lambda_2\\
&\leq c_5\,|a_{j_2-j_1}^{i_2,j_2}|^\gamma\int_{\mathbb{R}}|\widehat{K}(\lambda_2 h_n)||\phi_\varepsilon(a_0\lambda_2)||\lambda_2|^\gamma\Big(\int_{\mathbb{R}}|\widehat{K}(\lambda_1 h_n)||\phi_\varepsilon(a_p^{i_1,j_1}\lambda_1-a_{j_2-j_1+p}^{i_2,j_2}\lambda_2)|\,d\lambda_1\Big)\,d\lambda_2\\
&\leq c_6\,h_n^{-1}\,|a_{j_2-j_1}^{i_2,j_2}|^\gamma,
\end{aligned}$$

where

$$a_p^{i_1,j_1}=\begin{cases}-a_{p_1} & \text{if } p_1<j_1-i_1\\ a_0 & \text{if } p_1>j_1-i_1\\ a_0-a_{p_1} & \text{if } j_1-i_1=p_1 \text{ and } a_0\neq a_{p_1}\\ a_{\ell_0}^{i_1,j_1} & \text{if } j_1-i_1=p_1 \text{ and } a_0=a_{p_1}\end{cases}$$

with

$$p=\begin{cases}j_1-p_1 & \text{if } p_1<j_1-i_1\\ j_1-i_1 & \text{if } p_1>j_1-i_1\\ p_1 & \text{if } j_1-i_1=p_1 \text{ and } a_0\neq a_{p_1}\\ \ell_0 & \text{if } j_1-i_1=p_1 \text{ and } a_0=a_{p_1}\end{cases}.$$

Therefore, by the definition of $a_\ell^{i,j}$ in (2.6) and assumption **(A1)**,

$$\begin{aligned}
&\sum_{\substack{1\leq i_\theta<j_\theta\leq n,\,j_\theta-i_\theta\leq p_2\\ \theta=1,2,\,|j_1-j_2|>4p_2,\,j_1<j_2}}\ \sum_{k=-\infty}^{\min\{j_1,j_2\}}\left|\mathbb{E}\left[\mathscr{P}_k(A_{n,\delta}^{i_1,j_1})\mathscr{P}_k(A_{n,\delta}^{i_2,j_2})\right]\right|\\
&\leq c_7\,h_n^{-1}\sum_{\substack{1\leq i_\theta<j_\theta\leq n,\,j_\theta-i_\theta\leq p_2\\ \theta=1,2,\,|j_1-j_2|>4p_2,\,j_1<j_2}}\Big(|a_{j_2-j_1}^{i_2,j_2}|^\gamma+\sum_{k=-\infty}^{j_1-1}|a_{j_1-k}^{i_1,j_1}|^\gamma|a_{j_2-k}^{i_2,j_2}|^\gamma\Big)\\
&\leq c_8\,nh_n^{-1}.
\end{aligned}\tag{2.7}$$

Similarly,

$$\sum_{\substack{1\leq i_\theta<j_\theta\leq n,j_\theta-i_\theta\leq p_2\\ \theta=1,2,\,|j_1-j_2|>4p_2,\,j_1>j_2}}\ \sum_{k=-\infty}^{\min\{j_1,j_2\}}\left|\mathbb{E}\left[\mathscr{P}_k(A_{n,\delta}^{i_1,j_1})\mathscr{P}_k(A_{n,\delta}^{i_2,j_2})\right]\right|\leq c_9\,nh_n^{-1}.\tag{2.8}$$

Combining (2.5), (2.7) and (2.8) gives

$$\mathbb{E}\,|A_{n,\delta}|^2\leq\frac{c_{10}}{n^3h_n^2}.\tag{2.9}$$

Step 4. We estimate $\mathbb{E}\,|D_{n,\delta}|$. Note that

$$D_{n,\delta}=D_{n,\delta}^1+D_{n,\delta}^2,\tag{2.10}$$

where

$$D^1_{n,\delta}=\frac{1}{\pi n(n-1)}\sum_{1\le i<j\le n,\, j-i>p_2}\sum_{k=-\infty}^{i}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_k\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_k\Big(H(X_j)(\lambda)\Big)\,d\lambda$$

and

$$D^2_{n,\delta}=\frac{1}{\pi n(n-1)}\sum_{\substack{1\le i<j\le n,\, j-i>p_2\\ k\le i,\, \ell\le j,\, k\neq \ell}}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_k\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_\ell\Big(H(X_j)(\lambda)\Big)\,d\lambda$$

with $H(X_i)(\lambda)=e^{\iota\lambda X_i}-\mathbb{E}\,e^{\iota\lambda X_i}$ for $1\le i\le n$ and $\lambda\in\mathbb{R}$.

Using similar arguments as in **Step 3.**,

$$\begin{aligned}
\mathbb{E}\,|D^1_{n,\delta}| &= \frac{2}{\pi n(n-1)}\mathbb{E}\Bigg|\sum_{1\le i<j\le n,\, j-i>p_2}\sum_{k=-\infty}^{i}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_k\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_k\Big(H(X_j)(\lambda)\Big)\,d\lambda\Bigg|\\
&\le \frac{c_{11}}{n^2}\sum_{1\le i<j\le n,\, j-i>p_2}\sum_{k=-\infty}^{i}|a_{i-k}|^\gamma|a_{j-k}|^\gamma\int_{\mathbb{R}}|\phi_\varepsilon(a_0\lambda)||\phi_\varepsilon(a_{p_1}\lambda)||\lambda|^{2\gamma}\,d\lambda\\
&\le \frac{c_{12}}{n}. \qquad (2.11)
\end{aligned}$$

Moreover,

$$D^2_{n,\delta}=\mathrm{II}_1+\mathrm{II}_2+\mathrm{II}_3+\mathrm{II}_4+\mathrm{II}_5, \qquad (2.12)$$

where

$$\begin{aligned}
\mathrm{II}_1&=\frac{1}{\pi n(n-1)}\sum_{1\le i<j\le n,\, j-i>p_2}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_i\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_j\Big(H(X_j)(\lambda)\Big)\,d\lambda,\\
\mathrm{II}_2&=\frac{1}{\pi n(n-1)}\sum_{1\le i<j\le n,\, j-i>p_2}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_i\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_{j-p_1}\mathscr{P}_j\Big(H(X_j)(\lambda)\Big)\,d\lambda,\\
\mathrm{II}_3&=\frac{1}{\pi n(n-1)}\sum_{1\le i<j\le n,\, j-i>p_2}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_{i-p_1}\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_j\Big(H(X_j)(\lambda)\Big)\,d\lambda,\\
\mathrm{II}_4&=\frac{1}{\pi n(n-1)}\sum_{1\le i<j\le n,\, j-i>p_2}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_{i-p_1}\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_{j-p_1}\Big(H(X_j)(\lambda)\Big)\,d\lambda
\end{aligned}$$

and

$$\mathrm{II}_5=\frac{1}{\pi n(n-1)}\sum_{\substack{1\le i<j\le n,\, j-i>p_2,\, k\neq\ell\\ i-k>p_1 \text{ or } j-\ell>p_1}}\int_{\mathbb{R}}\widehat{K_\delta}(\lambda h_n)\mathscr{P}_k\Big(H(X_i)(-\lambda)\Big)\mathscr{P}_\ell\Big(H(X_j)(\lambda)\Big)\,d\lambda.$$

By Cauchy-Schwarz inequality and assumptions **(A1)**, **(A2)** and **(A3)**,

$$
\begin{aligned}
\mathbb{E}|\mathrm{II}_1|^2 &= \frac{1}{\pi^2 n^2 (n-1)^2} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \widehat{K_\delta}(\lambda_1 h_n)\widehat{K_\delta}(\lambda_2 h_n) \\
&\qquad \times \mathbb{E}\Big[\mathscr{P}_{i_1}\big(H(X_{i_1})(-\lambda_1)\big)\mathscr{P}_j\big(H(X_j)(\lambda_1)\big)\mathscr{P}_{i_2}\big(H(X_{i_2})(-\lambda_2)\big)\mathscr{P}_j\big(H(X_j)(\lambda_2)\big)\Big]\, d\lambda_1 d\lambda_2 \\
&\le \frac{c_{13}}{n^4} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \Big|\widehat{K}(\lambda_1 h_n)\widehat{K}(\lambda_2 h_n)\phi_\varepsilon\big(a_{p_1}(\lambda_1+\lambda_2)\big)\phi_\varepsilon\big(a_{p_2}(\lambda_1+\lambda_2)\big)\Big| \\
&\qquad \times \Big\{\Big|\mathbb{E}\Big[\big(e^{\iota a_{j-i_1}(\lambda_1+\lambda_2)} - \phi_\varepsilon(a_{j-i_1}(\lambda_1+\lambda_2))\big)\big(e^{-\iota a_0\lambda_1} - \phi_\varepsilon(-a_0\lambda_1)\big)\Big]\Big| 1_{\{i_1>i_2\}} \\
&\qquad\qquad + \Big|\mathbb{E}\Big[\big(e^{\iota a_{j-i_2}(\lambda_1+\lambda_2)} - \phi_\varepsilon(a_{j-i_2}(\lambda_1+\lambda_2))\big)(e^{-\iota a_0\lambda_1} - \phi_\varepsilon(-a_0\lambda_1)\Big]\Big| 1_{\{i_2>i_1\}} \\
&\qquad\qquad\quad + 1_{\{i_2=i_1\}}\Big\}\, d\lambda_1 d\lambda_2 \\
&\le \frac{c_{14}}{n^4} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \Big|\widehat{K}(\lambda_1 h_n)\widehat{K}(\lambda_2 h_n)\phi_\varepsilon\big(a_{p_1}(\lambda_1+\lambda_2)\big)\phi_\varepsilon\big(a_{p_2}(\lambda_1+\lambda_2)\big)\Big| \\
&\qquad \times \Big\{|\lambda_1+\lambda_2|^\gamma\Big(|a_{j-i_1}|^\gamma 1_{\{i_1>i_2\}} + |a_{j-i_2}|^\gamma 1_{\{i_2>i_1\}}\Big) + 1_{\{i_2=i_1\}}\Big\}\, d\lambda_1\, d\lambda_2 \\
&\le \frac{c_{15}}{n^2 h_n}
\end{aligned} \tag{2.13}
$$

and

$$
\begin{aligned}
\mathbb{E}|\mathrm{II}_2|^2 &= \frac{1}{\pi^2 n^2 (n-1)^2} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \widehat{K_\delta}(\lambda_1 h_n)\widehat{K_\delta}(\lambda_2 h_n) \\
&\qquad \times \mathbb{E}\Big[\mathscr{P}_{i_1}\big(H(X_{i_1})(-\lambda_1)\big)\mathscr{P}_{j-p_1}\big(H(X_j)(\lambda_1)\big)\mathscr{P}_{i_2}\big(H(X_{i_2})(-\lambda_2)\big)\mathscr{P}_{j-p_1}\big(H(X_j)(\lambda_2)\big)\Big]\, d\lambda_1\, d\lambda_2 \\
&\le \frac{c_{16}}{n^4} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \Big|\widehat{K}(\lambda_1 h_n)\widehat{K}(\lambda_2 h_n)\phi_\varepsilon(a_0\lambda_1)\phi_\varepsilon(a_0\lambda_2)\phi_\varepsilon\big(a_{p_2}(\lambda_1+\lambda_2)\big)\Big| \\
&\qquad \times \Big\{\Big|\mathbb{E}\Big[\big(e^{\iota a_{j-i_1}(\lambda_1+\lambda_2)} - \phi_\varepsilon(a_{j-i_1}(\lambda_1+\lambda_2))\big)\big(e^{-\iota a_0\lambda_1} - \phi_\varepsilon(-a_0\lambda_1)\big)\Big]\Big| 1_{\{i_1>i_2\}} \\
&\qquad\qquad + \Big|\mathbb{E}\Big[\big(e^{\iota a_{j-i_2}(\lambda_1+\lambda_2)} - \phi_\varepsilon(a_{j-i_2}(\lambda_1+\lambda_2))\big)(e^{-\iota a_0\lambda_1} - \phi_\varepsilon(-a_0\lambda_1)\Big]\Big| 1_{\{i_2>i_1\}} \\
&\qquad\qquad\quad + 1_{\{i_2=i_1\}}\Big\}\, d\lambda_1\, d\lambda_2 \\
&\le \frac{c_{17}}{n^4} \sum_{\substack{1\le i_\theta<j\le n,\, j-i_\theta>p_2\\ \theta=1,2}} \int_{\mathbb{R}^2} \Big|\widehat{K}(\lambda_1 h_n)\widehat{K}(\lambda_2 h_n)\phi_\varepsilon(a_0\lambda_1)\phi_\varepsilon(a_0\lambda_2)\phi_\varepsilon\big(a_{p_2}(\lambda_1+\lambda_2)\big)\Big| \\
&\qquad \times \Big\{|\lambda_1+\lambda_2|^\gamma\Big(|a_{j-i_1}|^\gamma 1_{\{i_1>i_2\}} + |a_{j-i_2}|^\gamma 1_{\{i_2>i_1\}}\Big) + 1_{\{i_2=i_1\}}\Big\}\, d\lambda_1\, d\lambda_2 \\
&\le \frac{c_{18}}{n^2 h_n}.
\end{aligned} \tag{2.14}
$$

Similarly,

$$
\mathbb{E}|\mathrm{II}_3|^2 \le \frac{c_{19}}{n^2 h_n} \quad \text{and} \quad \mathbb{E}|\mathrm{II}_4|^2 \le \frac{c_{19}}{n^2 h_n}. \tag{2.15}
$$

Note that

$$\mathrm{II}_4 = \mathrm{II}_{41} + \mathrm{II}_{42},$$

where

$$\mathrm{II}_{41} = \frac{1}{\pi n(n-1)} \sum_{\substack{1\le i<j\le n,\, j-i>p_2,\, k>\ell \\ i-k>p_1 \text{ or } j-\ell>p_1}} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n) \mathscr{P}_i\Big(H(X_i)(-\lambda)\Big) \mathscr{P}_\ell\Big(H(X_j)(\lambda)\Big)\, d\lambda,$$

and

$$\mathrm{II}_{42} = \frac{1}{\pi n(n-1)} \sum_{\substack{1\le i<j\le n,\, j-i>p_2,\, k<\ell \\ i-k>p_1 \text{ or } j-\ell>p_1}} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n) \mathscr{P}_i\Big(H(X_i)(-\lambda)\Big) \mathscr{P}_\ell\Big(H(X_j)(\lambda)\Big)\, d\lambda.$$

Clearly,

$$\mathbb{E}|\mathrm{II}_{41}|^2 = \frac{1}{\pi^2 n^2 (n-1)^2} \sum_{\substack{1\le i_\theta<j_\theta\le n,\, j_\theta-i_\theta>p_2 \\ i_\theta-k_\theta>p_1 \text{ or } j_\theta-\ell_\theta>p_1 \\ k_\theta>\ell_\theta,\, \theta=1,2}} \int_{\mathbb{R}^2} \widehat{K_\delta}(\lambda_1 h_n) \widehat{K_\delta}(\lambda_2 h_n)\, \mathrm{I}^{i_1,i_2,j_1,j_2}_{k_1,k_2,\ell_1,\ell_2}(\lambda_1,\lambda_2)\, d\lambda_1 d\lambda_2,$$

where

$$\begin{aligned} &\mathrm{I}^{i_1,i_2,j_1,j_2}_{k_1,k_2,\ell_1,\ell_2}(\lambda_1,\lambda_2) \\ &\quad = \mathbb{E}\Big[\mathscr{P}_{k_1}\Big(H(X_{i_1})(-\lambda_1)\Big)\mathscr{P}_{\ell_1}\Big(H(X_{j_1})(\lambda_1)\Big)\mathscr{P}_{k_2}\Big(H(X_{i_2})(-\lambda_2)\Big)\mathscr{P}_{\ell_2}\Big(H(X_{j_2})(\lambda_2)\Big)\Big] \end{aligned} \tag{2.16}$$

There are four possibilities for the orderings of k_1, ℓ_1, k_2, ℓ_2:

$$(1)\ k_1 \ge k_2 \ge \ell_1 \ge \ell_2,\ (2)\ k_1 \ge k_2 \ge \ell_2 \ge \ell_1,\ (3)\ k_2 \ge k_1 \ge \ell_1 \ge \ell_2,\ (4)\ k_2 \ge k_1 \ge \ell_2 \ge \ell_1.$$

Note that the expectation in (2.16) is equal to zero if $k_1 \ne k_2$. Hence, by Cauchy-Schwarz inequality and assumption **(A2)**,

$$\begin{aligned} &\Big|\mathrm{I}^{i_1,i_2,j_1,j_2}_{k_1,k_2,\ell_1,\ell_2}(\lambda_1,\lambda_2)\Big| \\ &\le 1_{\{k_1=k_2\}}\Big|\phi_\varepsilon(a_0\lambda_1)\phi_\varepsilon(a_{p_1}\lambda_1)\phi_\varepsilon(a_0\lambda_2)\phi_\varepsilon(a_{p_1}\lambda_2)\Big| \\ &\qquad \times \Big|\mathbb{E}\Big[\big(e^{\iota a_{i_1-k_1}\lambda_1\varepsilon_{k_1}} - \phi_\varepsilon(a_{i_1-k_1}\lambda_1)\big)\big(e^{\iota a_{i_2-k_2}\lambda_2\varepsilon_{k_2}} - \phi_\varepsilon(a_{i_2-k_2}\lambda_2)\big)\Big]\Big| \\ &\quad \times \bigg\{\Big|\mathbb{E}\Big[(e^{-\iota(a_{i_1-\ell_1}\lambda_1 + a_{i_2-\ell_1}\lambda_2)\varepsilon_{\ell_1}} - \phi_\varepsilon(-a_{i_1-\ell_1}\lambda_1 + a_{i_2-\ell_1}\lambda_2))(e^{\iota a_{j_1-\ell_1}\lambda_1\varepsilon_{\ell_1}} - \phi_\varepsilon(a_{j_1-\ell_1}\lambda_1))\Big]\Big| 1_{\{\ell_1>\ell_2\}} \\ &\qquad + \Big|\mathbb{E}\Big[\big(e^{-\iota(a_{i_1-\ell_2}\lambda_1 + a_{i_2-\ell_2}\lambda_2)\varepsilon_{\ell_2}} - \phi_\varepsilon(-a_{i_1-\ell_2}\lambda_1 + a_{i_2-\ell_2}\lambda_2)\big)\big(e^{\iota a_{j_2-\ell_2}\lambda_2\varepsilon_{\ell_2}} - \phi_\varepsilon(a_{j_2-\ell_2}\lambda_2)\big)\Big]\Big| 1_{\{\ell_2>\ell_1\}} \\ &\qquad\quad + \Big|\mathbb{E}\Big[\big(e^{\iota a_{j_1-\ell_1}\lambda_1\varepsilon_{\ell_1}} - \phi_\varepsilon(a_{j_1-\ell_1}\lambda_1)\big)\big(e^{\iota a_{j_2-\ell_2}\lambda_2\varepsilon_{\ell_2}} - \phi_\varepsilon(a_{j_2-\ell_2}\lambda_2)\big)\Big]\Big| 1_{\{\ell_2=\ell_1\}}\bigg\} \\ &\le c_{20} 1_{\{k_1=k_2\}}\Big|\phi_\varepsilon(a_0\lambda_1)\phi_\varepsilon(a_{p_1}\lambda_1)\phi_\varepsilon(a_0\lambda_2)\phi_\varepsilon(a_{p_1}\lambda_2)\Big| |a_{i_1-k_1}\lambda_1|^\gamma |a_{i_2-k_2}\lambda_2|^\gamma \\ &\qquad \times \Big\{|a_{i_1-\ell_1}\lambda_1 + a_{i_2-\ell_1}\lambda_2|^\gamma |a_{j_1-\ell_1}\lambda_1|^\gamma ||a_{j_2-\ell_2}\lambda_2|^\gamma| 1_{\{\ell_1>\ell_2\}} \\ &\qquad\qquad + |a_{i_1-\ell_2}\lambda_1 + a_{i_2-\ell_2}\lambda_2|^\gamma |a_{j_2-\ell_2}\lambda_2|^\gamma |a_{j_1-\ell_1}\lambda_1|^\gamma | 1_{\{\ell_2>\ell_1\}} \\ &\qquad\qquad\quad + |a_{j_1-\ell_1}\lambda_1|^\gamma |a_{j_2-\ell_2}\lambda_2|^\gamma 1_{\{\ell_2=\ell_1\}}\Big\}. \end{aligned}$$

Thus, by assumption (A3),

$$\mathbb{E}|\mathrm{II}_{41}|^2 \le \frac{c_{21}}{n^4} \sum_{\substack{1\le i_\theta<j_\theta\le n,\, j_\theta-i_\theta>p_2 \\ i_\theta-k_\theta>p_1 \text{ or } j_\theta-\ell_\theta>p_1 \\ k_\theta>\ell_\theta,\, \theta=1,2}} \int_{\mathbb{R}^2} \left|\widehat{K}(\lambda_1 h_n)\widehat{K}(\lambda_2 h_n)\right| \left|\mathrm{I}^{i_1,i_2,j_1,j_2}_{k_1,k_2,\ell_1,\ell_2}(\lambda_1,\lambda_2)\right| d\lambda_1\, d\lambda_2 \le \frac{c_{22}}{n^2}.$$

Similarly,

$$\mathbb{E}|\mathrm{II}_{42}|^2 \le \frac{c_{23}}{n^2}.$$

Therefore,

$$\mathbb{E}\,|\mathrm{II}_4|^2 = \mathbb{E}\,|\mathrm{II}_{41}+\mathrm{II}_{42}|^2 \le \frac{c_{24}}{n^2}. \tag{2.17}$$

Combining (2.10), (2.11), (2.12), (2.13), (2.14), (2.15) and (2.17) gives

$$\mathbb{E}\,|D_{n,\delta}| \le \frac{c_{25}}{\sqrt{n^2 h_n}}. \tag{2.18}$$

Step 5. We estimate $\mathbb{E}\left|N_{n,\delta}-\overline{N}_n\right|^2$ where

$$\overline{N}_n = \frac{1}{\pi n}\sum_{i=1}^n \int_{\mathbb{R}} \widehat{K}(0)\big(e^{-\iota\lambda X_i}-\phi(-\lambda)\big)\phi(\lambda)\, d\lambda.$$

Recall that

$$\begin{aligned} N_{n,\delta} &= \frac{1}{\pi n(n-1)} \sum_{1\le i<j\le n,\, j-i>p_2} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\Big[\big(e^{-\iota\lambda X_i}-\phi(-\lambda)\big)\phi(\lambda) + \big(e^{\iota\lambda X_j}-\phi(\lambda)\big)\phi(-\lambda)\Big]\, d\lambda \\ &= \frac{1}{2\pi n(n-1)} \sum_{1\le i\ne j\le n,\, |j-i|>p_2} \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\Big[\big(e^{-\iota\lambda X_i}-\phi(-\lambda)\big)\phi(\lambda) + \big(e^{\iota\lambda X_j}-\phi(\lambda)\big)\phi(-\lambda)\Big]\, d\lambda, \end{aligned}$$

where we use the symmetry of $K_\delta(x)$ in the last equality.

Let

$$\widetilde{N}_{n,\delta} = \frac{1}{\pi n}\sum_{i=1}^n \int_{\mathbb{R}} \widehat{K_\delta}(\lambda h_n)\big(e^{-\iota\lambda X_i}-\phi(-\lambda)\big)\phi(\lambda)\, d\lambda.$$

It is easy to see that $|N_{n,\delta}-\widetilde{N}_{n,\delta}|$ is less than a constant multiple of $\frac{1}{n}$.

Moreover, by Cauchy-Schwarz inequality, Fubini theorem, (2.3), assumption **(A3)** and Lemma 7.2 in [4],

$$\begin{aligned} \mathbb{E}\left|\widetilde{N}_{n,\delta}-\overline{N}_n\right|^2 &= \frac{1}{\pi^2}\mathbb{E}\left[\Big|\int_{\mathbb{R}} \big(\widehat{K_\delta}(\lambda h_n)-\widehat{K_\delta}(0)\big)\Big(\frac{1}{n}\sum_{i=1}^n e^{-\iota\lambda X_i}-\phi(-\lambda)\Big)\phi(\lambda)\, d\lambda\Big|^2\right] \\ &\le \Big(\int_{\mathbb{R}} \big|\widehat{K_\delta}(\lambda h_n)-\widehat{K_\delta}(0)\big|^2 |\phi(\lambda)|\, d\lambda\Big)\Big(\int_{\mathbb{R}} \mathbb{E}\Big|\frac{1}{n}\sum_{i=1}^n e^{-\iota\lambda X_i}-\phi(\lambda)\Big|^2 |\phi(\lambda)|\, d\lambda\Big) \\ &\le c_{26}\frac{h_n^{2\gamma}}{n}. \end{aligned}$$

Hence

$$\mathbb{E}\left|N_n-\overline{N}_n\right|^2 \le c_{27}\Big(\frac{1}{n^2}+\frac{h_n^{2\gamma}}{n}\Big). \tag{2.19}$$

Note that $\widehat{K}(0)=1$. By the Fourier inverse transform and Plancherel formula,

$$\overline{N}_n=\frac{2}{n}\sum_{i=1}^{n}\Big(f(X_i)-\int_{\mathbb{R}}f^2(x)\,dx\Big).$$

Combining (2.4), (2.9), (2.18) and (2.19) gives

$$\begin{aligned}
&\mathbb{E}\,\Big|T_n(h_n)-\mathbb{E}\,T_n(h_n)-\frac{1}{n}\sum_{i=1}^{n}Y_i\Big|\\
&\le \limsup_{\delta\downarrow 0}\mathbb{E}\,\Big|T_n(h_n)-\mathbb{E}\,T_n(h_n)-\big(T_n^{\delta}(h_n)-\mathbb{E}\,T_n^{\delta}(h_n)\big)\Big|+\limsup_{\delta\downarrow 0}\mathbb{E}\,\Big|T_n^{\delta}(h_n)-\mathbb{E}\,T_n^{\delta}(h_n)-\frac{1}{n}\sum_{i=1}^{n}Y_i\Big|\\
&\le c_{28}\Big(\frac{1}{\sqrt{n^3h_n^2}}+\frac{1}{\sqrt{n^2h_n}}+\frac{h_n^{\gamma}}{\sqrt{n}}\Big).
\end{aligned}$$

Step 6. We show the central limit theorem for

$$\frac{2}{n}\sum_{i=1}^{n}\Big(f(X_i)-\int_{\mathbb{R}}f^2(x)\,dx\Big).$$

Using Lemma 1 in [5] and similar arguments as in the proof of Theorem 2.1 in [4],

$$\frac{1}{n}\sum_{i=1}^{n}\Big(f(X_i)-\int_{\mathbb{R}}f^2(x)\,dx\Big)\xrightarrow{\mathcal{L}}N(0,\sigma^2)$$

for some $\sigma^2\in(0,\infty)$.

This finishes the proof of Theorem 1.1. □

References

[1] Giné, E., Nickl, R. (2008). A simple adaptive estimator of the integrated square of a density. *Bernoulli*, **14**: 47–61.

[2] Källberg, D., Leonenko, N. and Seleznjev, O. (2014). Statistical estimation of quadratic Rényi entropy for a stationary m-dependent sequence. *Journal of Nonparametric Statistics*, **26**: 385–411.

[3] Liu, H. and Xu, F. (2022). Kernel entropy estimation for long memory linear processes with infinite variance. arXiv:2210.03644v2, 2022.

[4] Sang, H., Sang, Y. and Xu, F. (2018). Kernel entropy estimation for linear processes. *Journal of Time Series Analysis*, **39**: 563–591.

[5] Wu, W.B.(2006). Unit root testing for functionals of linear processes. *Econometric Theory*, **22**: 1–14.

Yudan Xiong
School of Statistics, East China Normal University, Shanghai 200262, China
xyd_980107@163.com

Fangjun Xu
KLATASDS-MOE, School of Statistics, East China Normal University, Shanghai, 200062, China
NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, 200062, China
fjxu@finance.ecnu.edu.cn