



**HAL**  
open science

# Kernel density estimation for linear processes with quasi-associated innovations

Lahcen Douge

► **To cite this version:**

Lahcen Douge. Kernel density estimation for linear processes with quasi-associated innovations. *Annales de l'ISUP*, 2019, 63 (2-3), pp.21-32. hal-03604361

**HAL Id: hal-03604361**

**<https://hal.science/hal-03604361>**

Submitted on 10 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Kernel density estimation for linear processes with quasi-associated innovations

Lahcen DOUGE

*FSTG BP 549 Cadi Ayyad University, Marrakech, Morocco*

**e-mail: lahcen.douge@uca.ac.ma**

**Abstract:** In this paper, we study the kernel estimate of the density function of linear processes with quasi-associated innovations. We prove the asymptotic normality of the kernel density estimator under mild regularity conditions and some conditions on the coefficients of linear processes and on the decay of the covariances.

### 1. Introduction

Let  $X_1, X_2, \dots$  be an identically distributed sequence of random variables defined by  $X_i = \sum_{r=0}^{\infty} a_r Z_{i-r}$ ,  $i \geq 1$ , where  $\{Z_t\}_{t \in \mathbb{Z}}$  is a strictly stationary sequence of quasi-associated random variables with mean zero and finite variance, and  $\{a_r\}_{r \in \mathbb{N}}$  is a sequence of real numbers. The concept of quasi-association was introduced by [2] and provides a unified approach to studying both families of positively or negatively associated and Gaussian random variables. A sequence  $(Z_i)_{i \in \mathbb{Z}}$  of real-random variables is said to be quasi-associated if for any finite and disjoint subsets  $I, J \subset \mathbb{Z}$  and all bounded Lipschitz functions  $g : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ , one has

$$(1.1) \quad \left| \text{Cov}(g(Z_i, i \in I), h(Z_j, j \in J)) \right| \leq \text{Lip}(g)\text{Lip}(h) \sum_{i \in I} \sum_{j \in J} |\text{Cov}(Z_i, Z_j)|,$$

where  $|I|$  denotes the cardinality of  $I$  and

$$\text{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1}$$

the Lipschitz modulus of a function  $h$ .

Assume that  $X_1$  has a probability density  $f$  and, for any  $1 \leq i, j \leq n$ ,  $(X_i, X_j)$  has a probability density  $f_{i,j}$ . Further, let  $\{h_n\}_{n \geq 0}$  be a sequence of positive constants such that, as  $n \rightarrow \infty$ ,  $h_n \rightarrow 0$  and  $nh_n \rightarrow \infty$ . The classical kernel estimator of  $f$  is defined as

$$f_n(x) := \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad x \in \mathbb{R},$$



where the kernel  $K$  is a bounded function such that

$$\int_{\mathbb{R}} K(u) du = 1.$$

The asymptotic normality of the kernel density estimators was studied by [3], [6] and [12] for linear processes with independent innovations under some conditions on  $\{a_r\}$  and  $\{h_n\}$ . [12] improved Hallin and Tran's results by imposing  $\sum_{r=0}^{\infty} |a_r| < \infty$  and only the natural conditions on  $\{h_n\}$ . For linear processes with  $\alpha$ -mixing innovations, [9] showed that the kernel density estimators are asymptotically normal under general and mild conditions. In this paper, we establish the asymptotic normality of the kernel density estimate for linear processes with quasi-associated innovations. We note that quasi-association and mixing define two distinct but not disjoint classes of processes. We mention also that an important property of quasi-associated random variables is that non-correlation implies independence and the dependence in this case is evaluated only through the covariance structure which is much easier to compute than a mixing coefficient. Our result is obtained under the same general conditions on the density function  $f$  and the kernel  $K$ , and the coefficients  $\{a_n\}$  as in [6] and [9]. Some additional conditions on the bandwidth and the decay of the covariances are assumed.

For other works on density estimation for linear processes, we can refer to [7], [8], [10] and [13].

## 2. Notation, assumptions and main result

Divide the set  $\{1, \dots, n\}$  into  $k$  large  $p$ -blocks,  $I_j$ , and small  $q$ -blocks,  $J_j$ ,  $j = 1, \dots, k$ , as follows

$$\begin{aligned} I_j &= \{(j-1)(p+q)+1, \dots, (j-1)(p+q)+p\}, \\ J_j &= \{(j-1)(p+q)+p+1, \dots, j(p+q)\}, \end{aligned}$$

where  $p$  and  $q$  are positive integers tending to  $\infty$  as  $n \rightarrow \infty$  and  $k$  is defined by  $k = [n/(p+q)]$ , where  $[x]$  stands for the integral part of  $x$ . We suppose that

$$\frac{qk}{n} \rightarrow 0 \quad \text{and} \quad \frac{pk}{n} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Define

$$\tilde{X}_i := \sum_{r=0}^s a_r Z_{i-r}, \quad i \geq 1,$$

where  $s < q$  is a positive integer tending to  $\infty$  as  $n \rightarrow \infty$ .

Consider the kernel estimator  $g_n(x)$  defined by

$$g_n(x) := \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - \tilde{X}_i}{h_n}\right), \quad x \in \mathbb{R}.$$



For each  $x \in \mathbb{R}$ , set

$$Z_{ni}(x) := \frac{1}{\sqrt{nh_n}} \left( K \left( \frac{x - \tilde{X}_i}{h_n} \right) - \mathbb{E} K \left( \frac{x - \tilde{X}_i}{h_n} \right) \right)$$

and

$$S_n := \sqrt{nh_n} (g_n(x) - \mathbb{E} g_n(x)) = \sum_{i=1}^n Z_{ni}(x).$$

For  $j = 1, \dots, k$ , let  $\eta_j$ ,  $\xi_j$  and  $\zeta_k$  be defined as follows

$$\eta_j := \sum_{i \in I_j} Z_{ni}(x), \quad \xi_j := \sum_{i \in J_j} Z_{ni}(x), \quad \zeta_k := \sum_{i=k(p+q)+1}^n Z_{ni}(x),$$

so that

$$S_n = \sum_{j=1}^k \eta_j + \sum_{j=1}^k \xi_j + \zeta_k =: T_n + T'_n + T''_n.$$

In order to formulate our main result, we now list some assumptions. Denote by  $C$  (different) constants whose values are allowed to change. Through this paper we consider that all limits are taken as  $n \rightarrow \infty$ .

### Assumptions

(A1) For all  $x, y \in \mathbb{R}$ ,

$$\sup_{i \neq j} |f_{i,j}(x, y) - f(x)f(y)| < C.$$

(A2) The kernel  $K$  satisfies a Lipschitz condition of order 1, that is, for all  $x, y \in \mathbb{R}$ ,

$$|K(x) - K(y)| \leq C|x - y|.$$

(A3) The sequence  $\{a_r\}_{r \in \mathbb{N}}$  satisfies the condition  $|a_r| = \mathcal{O}(r^{-a})$ , for some  $a > 4$ .

(A4)  $\theta_k := |\text{Cov}(Z_1, Z_{k+1})| = \mathcal{O}(k^{-b})$ ,  $k \geq 0$  and  $b > 2 + \frac{a-2}{2(a-1)} \left( \frac{a+2}{a-3} + \frac{3}{2} + 3(a-1) \right)$ .

(A5)  $n h_n^{(a+2)/(a-3)} (\log n)^{-\delta} \rightarrow \infty$ , for some  $\delta > 0$ .

(A6)  $(\log n)^{\delta(a-3)/4} n h_n^{[2(b-2)/(a-2)-3](a-1)-3/2} \rightarrow 0$ .

Notice that if  $x$  is a Lebesgue point of  $f$  and  $(x, x)$  is a Lebesgue point of all functions  $f_{1,v}$ ,  $v \geq 2$ , by applying Theorem 3 in Chapter 2 of [4], the condition (A1) can be replaced by the condition  $\sup_{v \geq 2} |f_{1,v}(x, x)| < C$ . The condition (A2) is a mild condition on the kernel  $K$ . The conditions (A5) is necessary to ensure  $s < p$ , for a suitable choice of  $s$  and  $p$ , and allows, together with the condition (A6), to get Lemma 3.



**Theorem 1.** Assume that assumptions (A1)-(A6) hold. Then, for all Lebesgue points  $x$  of  $f$  such that  $f(x) > 0$ ,

$$\sqrt{nh_n} (f_n(x) - E f_n(x)) \xrightarrow{d} \mathcal{N}(0, \sigma^2(x)),$$

where

$$\sigma^2(x) = f(x) \int_{\mathbb{R}} K^2(u) du.$$

" $\xrightarrow{d}$ " denotes convergence in distribution.

We introduce the following lemmas which we need in the proof of the theorem. The first lemma shows that, for all  $x \in \mathbb{R}$ ,  $f_n(x)$  and  $g_n(x)$  have the same asymptotic distribution.

**Lemma 1.** Assume that assumptions (A2) and (A3) hold. Then, for any  $\varepsilon > 0$  and for all  $x \in \mathbb{R}$ ,

$$P\left[\sqrt{nh_n} |f_n(x) - g_n(x)| > \varepsilon\right] \rightarrow 0.$$

**Proof.** Using (A2), we get

$$\begin{aligned} \sqrt{nh_n} |f_n(x) - g_n(x)| &\leq (nh_n)^{-1/2} \sum_{i=1}^n \left| K\left(\frac{x - X_i}{h_n}\right) - K\left(\frac{x - \tilde{X}_i}{h_n}\right) \right| \\ &\leq C n^{-1/2} h_n^{-3/2} \sum_{i=1}^n |X_i - \tilde{X}_i|. \end{aligned}$$

Then

$$\begin{aligned} P\left[\sqrt{nh_n} |f_n(x) - g_n(x)| > \varepsilon\right] &\leq P\left[\sum_{i=1}^n |X_i - \tilde{X}_i| > \frac{n^{1/2} h_n^{3/2} \varepsilon}{C}\right] \\ &\leq \sum_{i=1}^n P\left[|X_i - \tilde{X}_i| > \frac{h_n^{3/2} \varepsilon}{C n^{1/2}}\right] \\ &\leq C n h_n^{-3} \varepsilon^{-2} \sum_{i=1}^n E |X_i - \tilde{X}_i|^2 \\ &\leq C n h_n^{-3} \varepsilon^{-2} \sum_{i=1}^n E \left( \sum_{r=s+1}^{\infty} a_r Z_{i-r} \right)^2 \\ &\leq C n^2 h_n^{-3} \varepsilon^{-2} E Z_1^2 \left( \sum_{r=s+1}^{\infty} |a_r| \right)^2. \end{aligned}$$

Choose  $s = [(\log n)^{\delta(a-3)/2} n^2 h_n^{-3}]^{1/2(a-1)}$ . Thus, by (A3),

$$P\left[\sqrt{nh_n} |f_n(x) - g_n(x)| > \varepsilon\right] \leq C n^2 h_n^{-3} \varepsilon^{-2} s^{-2(a-1)} = \mathcal{O}((\log n)^{-\delta(a-3)/2}). \quad \blacksquare$$



**Lemma 2.** Assume that assumptions (A1)-(A2) hold. Then, for all  $v \geq 2$  and for all  $x \in \mathbb{R}$ ,

$$(2.1) \quad \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \leq C \left[ h_n^2 + \frac{1}{h_n} \sum_{r=s+1}^{\infty} |a_r| \right].$$

**Proof.** For all  $v \geq 2$ , by (A1),

$$\begin{aligned} & \frac{1}{h_n^2} \left| \text{Cov} \left( K \left( \frac{x - X_1}{h_n} \right), K \left( \frac{x - X_v}{h_n} \right) \right) \right| \\ &= \left| \int_{\mathbb{R}^2} K(u)K(w) \left[ f_{1,v}(x - h_n u, x - h_n w) - f(x - h_n u)f(x - h_n w) \right] du dw \right| \\ &\leq C. \end{aligned}$$

By (A2), for all  $v \geq 2$ ,

$$\begin{aligned} & \left| \text{Cov} \left( K \left( \frac{x - X_1}{h_n} \right), K \left( \frac{x - X_v}{h_n} \right) \right) - \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \\ &\leq 2 \|K\|_{\infty} \left[ \mathbb{E} \left| K \left( \frac{x - X_1}{h_n} \right) - K \left( \frac{x - \tilde{X}_1}{h_n} \right) \right| + \mathbb{E} \left| K \left( \frac{x - X_v}{h_n} \right) - K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right| \right] \\ &\leq \frac{C}{h_n} \left[ \mathbb{E} |X_1 - \tilde{X}_1| + \mathbb{E} |X_v - \tilde{X}_v| \right] \\ &\leq \frac{C}{h_n} \sum_{r=s+1}^{\infty} |a_r|. \end{aligned}$$

Then

$$\left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \leq C \left[ h_n^2 + \frac{1}{h_n} \sum_{r=s+1}^{\infty} |a_r| \right]. \quad \blacksquare$$

**Lemma 3.** Assume that assumptions (A1)-(A6) hold. Then, for all  $x \in \mathbb{R}$ ,

$$(2.2) \quad \frac{1}{h_n} \sum_{v=2}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \rightarrow 0.$$

**Proof.** For  $2 \leq v \leq 2s$ , define

$$X_v^* := \sum_{r=0}^{[v/2]} a_r Z_{v-r}.$$



Observe that

$$\begin{aligned}
 & \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \\
 &= \mathbb{E} \left[ \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right) - \mathbb{E} K \left( \frac{x - \tilde{X}_1}{h_n} \right) \right) \left( K \left( \frac{x - \tilde{X}_v}{h_n} \right) - K \left( \frac{x - X_v^*}{h_n} \right) \right) \right] \\
 (2.3) \quad &+ \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - X_v^*}{h_n} \right) \right).
 \end{aligned}$$

Using (A2) and the boundedness of  $K$ , we obtain

$$\begin{aligned}
 & \left| \mathbb{E} \left[ \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right) - \mathbb{E} K \left( \frac{x - \tilde{X}_1}{h_n} \right) \right) \left( K \left( \frac{x - \tilde{X}_v}{h_n} \right) - K \left( \frac{x - X_v^*}{h_n} \right) \right) \right] \right| \\
 & \leq \frac{C}{h_n} \mathbb{E} |\tilde{X}_v - X_v^*| \\
 (2.4) \quad & \leq \frac{C}{h_n} \sum_{r=[v/2]+1}^{\infty} |a_r|.
 \end{aligned}$$

For  $n$  large enough such that  $h_n \leq 1$ , by (1.1) we get

$$(2.5) \quad \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - X_v^*}{h_n} \right) \right) \right| \leq C \frac{sv}{h_n^2} \theta_{v-[v/2]-1}.$$

By using again (1.1) we obtain, for  $v \geq 2s$ ,

$$(2.6) \quad \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \leq C \frac{s^2}{h_n^2} \theta_{v-s-1}.$$

Let  $m = m(n) = [h_n^{\theta-1}]$  with  $\theta$  is an arbitrary number such that  $0 < \theta < 1$  and  $\frac{2}{a-2} < 1 - \theta < \frac{1}{a-1} \left( \frac{a+2}{a-3} + \frac{3}{2} \right)$ . By (A5), we have, for  $n$  large enough,  $m < s$ . Then

$$\begin{aligned}
 & \sum_{v=2}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \\
 &= \sum_{v=2}^m \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| + \sum_{v=m+1}^{2s} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \\
 &+ \sum_{v=2s+1}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right|.
 \end{aligned}$$



Using (A3) and (A4), by (2.1), (2.3), (2.4), (2.5) and (2.6), we get, for  $n$  large enough,

$$\begin{aligned}
 & \frac{1}{h_n} \sum_{v=2}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \\
 & \leq C \left[ h_n^\theta + h_n^{-3+\theta} \sum_{r=s+1}^{\infty} |a_r| \right] + C h_n^{-2} \left[ \sum_{v=m+1}^{2s} \sum_{r=[v/2]+1}^{\infty} |a_r| + s h_n^{-1} \sum_{v=m+1}^{2s} v \theta_{v-[v/2]-1} \right] \\
 & \quad + C h_n^{-3} s^2 \sum_{v=2s+1}^{\infty} \theta_{v-s-1} \\
 & \leq C \left( h_n^\theta + h_n^{-3+\theta} s^{-(a-1)} + h_n^{-2} m^{-(a-2)} + h_n^{-3} s m^{-b+2} + h_n^{-3} s^{-b+3} \right) \\
 & \leq C \left[ h_n^\theta + \left( n h_n^{3/2-\theta} \right)^{-1} + h_n^{-2+(1-\theta)(a-2)} \right. \\
 & \quad \left. + \left( (\log n)^{\delta(a-3)/4} n h_n^{[(1-\theta)(b-2)-3](a-1)-3/2} \right)^{1/(a-1)} \right].
 \end{aligned}$$

Then, by (A5) and (A6)

$$\frac{1}{h_n} \sum_{v=2}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_v}{h_n} \right) \right) \right| \rightarrow 0.$$

This completes the proof of (2.2).

**Lemma 4.** Assume that assumptions (A1)-(A6) hold. Then, for all Lebesgue points  $x$  of  $f$ ,

$$n h_n \text{Var } g_n(x) \rightarrow \sigma^2(x).$$

**Proof.** We will show that

$$(2.7) \quad \text{E} (T_n)^2 \rightarrow \sigma^2(x)$$

and

$$(2.8) \quad \text{E} (T_n')^2 + \text{E} (T_n'')^2 \rightarrow 0.$$

First, consider (2.7). By using the strict stationarity of the sequence  $\{Z_t\}_{t \in \mathbb{Z}}$ , we have

$$\text{E} (T_n)^2 = k \text{Var} (\eta_1) + 2 \sum_{1 \leq i < j \leq k} \text{Cov} (\eta_i, \eta_j).$$

Let us establish that

$$(2.9) \quad k \text{Var} (\eta_1) \rightarrow \sigma^2(x)$$



and

$$(2.10) \quad \sum_{1 \leq i < j \leq k} \text{Cov}(\eta_i, \eta_j) \rightarrow 0.$$

By the strict stationarity of the sequence  $\{Z_t\}_{t \in \mathbb{Z}}$ , we get

$$k \text{Var}(\eta_1) = \frac{pk}{nh_n} \text{Var} K\left(\frac{x - \tilde{X}_1}{h_n}\right) + \frac{2k}{nh_n} \sum_{1 \leq i < j \leq p} \text{Cov}\left(K\left(\frac{x - \tilde{X}_i}{h_n}\right), K\left(\frac{x - \tilde{X}_j}{h_n}\right)\right)$$

By Theorem 3 in Chapter 2 of [4], for all Lebesgue points  $x$  of  $f$ ,

$$(2.11) \quad \frac{1}{h_n} \text{Var} K\left(\frac{x - X_1}{h_n}\right) \rightarrow f(x) \int_{\mathbb{R}} K^2(u) du.$$

By stationarity and (2.2), we obtain

$$(2.12) \quad \begin{aligned} & \frac{k}{nh_n} \left| \sum_{1 \leq i < j \leq p} \text{Cov}\left(K\left(\frac{x - \tilde{X}_i}{h_n}\right), K\left(\frac{x - \tilde{X}_j}{h_n}\right)\right) \right| \\ & \leq \frac{kp}{nh_n} \sum_{v=2}^{\infty} \left| \text{Cov}\left(K\left(\frac{x - \tilde{X}_1}{h_n}\right), K\left(\frac{x - \tilde{X}_v}{h_n}\right)\right) \right| \rightarrow 0. \end{aligned}$$

Now, by (A2) and (A5),

$$(2.13) \quad \begin{aligned} \frac{1}{h_n} \left| \text{Var} K\left(\frac{x - X_1}{h_n}\right) - \text{Var} K\left(\frac{x - \tilde{X}_1}{h_n}\right) \right| & \leq \frac{4\|K\|_{\infty}}{h_n} \mathbb{E} \left| K\left(\frac{x - X_1}{h_n}\right) - K\left(\frac{x - \tilde{X}_1}{h_n}\right) \right| \\ & \leq \frac{C}{h_n^2} \mathbb{E} |X_1 - \tilde{X}_1| \\ & \leq \frac{C}{h_n^2} \left[ \mathbb{E} \left( \sum_{r=s+1}^{\infty} a_r Z_{1-r} \right)^2 \right]^{1/2} \\ & \leq \frac{C}{h_n^2} \sum_{r=s+1}^{\infty} |a_r| \rightarrow 0, \end{aligned}$$

which, when combined with (2.11) and (2.12) completes the proof of (2.9).



Now, we will establish (2.10). Using stationarity and (2.6), we get

$$\begin{aligned}
 \sum_{1 \leq i < j \leq k} |\text{Cov}(\eta_i, \eta_j)| &\leq k \sum_{l=1}^{k-1} |\text{Cov}(\eta_l, \eta_{l+1})| \\
 &\leq \frac{pk}{nh_n} \sum_{l=1}^{k-1} \sum_{r=l(p+q)-p+1}^{l(p+q)+p-1} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_l}{h_n} \right), K \left( \frac{x - \tilde{X}_r}{h_n} \right) \right) \right| \\
 &\leq \frac{pk}{nh_n} \sum_{r=q+1}^{\infty} \left| \text{Cov} \left( K \left( \frac{x - \tilde{X}_1}{h_n} \right), K \left( \frac{x - \tilde{X}_r}{h_n} \right) \right) \right| \\
 (2.14) \quad &\leq C \frac{pk s^2}{nh_n^3} \sum_{r=q+1}^{\infty} \theta_{r-s}.
 \end{aligned}$$

Then, by (A4) and (A6), and by choosing  $q = 2s$ , we obtain

$$(2.15) \quad \sum_{1 \leq i < j \leq k} |\text{Cov}(\eta_i, \eta_j)| \rightarrow 0.$$

Next, we turn to (2.8). Observe that

$$E(T'_n)^2 \leq k \text{Var}(\xi_1) + 2 \sum_{1 \leq i < j \leq k} |\text{Cov}(\xi_i, \xi_j)|.$$

Once again, by stationarity,

$$k \text{Var}(\xi_1) = \frac{qk}{nh_n} \text{Var} K \left( \frac{x - \tilde{X}_1}{h_n} \right) + \frac{2k}{nh_n} \sum_{1 \leq i < j \leq q} \text{Cov} \left( K \left( \frac{x - \tilde{X}_i}{h_n} \right), K \left( \frac{x - \tilde{X}_j}{h_n} \right) \right).$$

By (2.2), (2.11) and (2.13), it follows that

$$k \text{Var}(\xi_1) \rightarrow 0.$$

Then, by the same arguments in (2.14) and (2.15), we obtain, for  $n$  large enough,

$$\begin{aligned}
 \sum_{1 \leq i < j \leq k} |\text{Cov}(\xi_i, \xi_j)| &\leq C \frac{qk s^2}{nh_n^3} \sum_{r=p}^{\infty} \theta_{r-s} \\
 &\leq C \frac{qk s^2}{nh_n^3} \sum_{r=q}^{\infty} \theta_{r-s} \rightarrow 0.
 \end{aligned}$$

Therefore

$$E(T'_n)^2 \rightarrow 0.$$



Now, by stationnarity,

$$E\left(T_n''\right)^2 \leq \frac{p+q}{nh_n} \text{Var} K\left(\frac{x-\tilde{X}_1}{h_n}\right) + \frac{2}{nh_n} \sum_{1 \leq i < j \leq p+q} \left| \text{Cov}\left(K\left(\frac{x-\tilde{X}_i}{h_n}\right), K\left(\frac{x-\tilde{X}_j}{h_n}\right)\right) \right|$$

By (2.2), (2.11) and (2.13), we get  $E\left(T_n''\right)^2 \rightarrow 0$ , which completes the proof of (2.8).

**Proof of Theorem 1.** By Lemma 1, it suffices to establish the asymptotic normality of  $g_n(x)$ . For this aim, by (2.8), we will show that

$$(2.16) \quad T_n \xrightarrow{d} \mathcal{N}(0, \sigma^2(x)).$$

The proof of convergence in (2.16) consists in using Theorem 7.2 of [1], (2.9) and showing the following two results

$$(2.17) \quad \left| E\left(e^{it \sum_{j=1}^k \eta_j}\right) - \prod_{j=1}^k E\left(e^{it \eta_j}\right) \right| \rightarrow 0, \quad \forall t \in \mathbb{R},$$

and

$$(2.18) \quad k E\left(\eta_1^2 \mathbb{1}_{\{\eta_1 > \varepsilon \sqrt{k \text{Var}(\eta_1)}\}}\right) \rightarrow 0, \quad \forall \varepsilon > 0.$$

Let us begin by establishing (2.18). Since  $|\eta_1| \leq Cp/\sqrt{nh_n}$ , it follows that

$$\begin{aligned} k E\left(\eta_1^2 \mathbb{1}_{\{\eta_1 > \varepsilon \sqrt{k \text{Var}(\eta_1)}\}}\right) &\leq C \frac{k p^2}{nh_n} P\left(\eta_1 > \varepsilon \sqrt{k \text{Var}(\eta_1)}\right) \\ &\leq \frac{C p^2}{\varepsilon^2 nh_n}. \end{aligned}$$

Choose  $p = [(\log n)^{-\delta(a-3)/4(a-1)}(nh_n)^{1/2}]$ . Then, we get

$$k E\left(\eta_1^2 \mathbb{1}_{\{\eta_1 > \varepsilon \sqrt{k \text{Var}(\eta_1)}\}}\right) \rightarrow 0.$$

Now, let us establish (2.17).

$$\begin{aligned} &\left| E\left(e^{it \sum_{j=1}^k \eta_j}\right) - \prod_{j=1}^k E\left(e^{it \eta_j}\right) \right| \\ &\leq \left| E\left(e^{it \sum_{j=1}^k \eta_j}\right) - E\left(e^{it \sum_{j=1}^{k-1} \eta_j}\right) E\left(e^{it \eta_k}\right) \right| + \left| E\left(e^{it \sum_{j=1}^{k-1} \eta_j}\right) - \prod_{j=1}^{k-1} E\left(e^{it \eta_j}\right) \right| \\ &= \left| \text{Cov}\left(e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it \eta_k}\right) \right| + \left| E\left(e^{it \sum_{j=1}^{k-1} \eta_j}\right) - \prod_{j=1}^{k-1} E\left(e^{it \eta_j}\right) \right|. \end{aligned}$$

(2.19)



By a repetition of this argument, inequality (2.19) becomes

$$\left| E \left( e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k E \left( e^{it \eta_j} \right) \right| \leq \left| \text{Cov} \left( e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it \eta_k} \right) \right| \\ + \dots + \left| \text{Cov} \left( e^{it \eta_2}, e^{it \eta_1} \right) \right|.$$

Thus, by the same arguments as in (2.14) and (2.15), we obtain

$$\left| E \left( e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k E \left( e^{it \eta_j} \right) \right| \leq C \frac{p k s^2 t^2}{n h_n^3} \sum_{r \geq q+1} \theta_{r-s} \rightarrow 0. \quad \blacksquare$$

### 3. Conclusion

In this paper, we established the asymptotic normality of the kernel density estimator for linear processes with quasi-associated innovations. The almost sure convergence will be studied in the further works under both arithmetic and exponential decay of the covariances. Another interesting perspective of this work would be the investigation of asymptotic properties of the kernel estimate of the regression function for linear processes generated by a large class of weakly dependent innovations. We note that the linear processes with weakly dependent errors are of great importance in econometrics and discussed in [11] and [5].

### References

- [1] BILLINGSLEY, P. (1968). *Convergence of probability measures*. Wiley, New York.
- [2] BULINSKI, A. AND SUQUET, C. (2001). Normal approximation for quasi-associated random fields. *Statist. Probab. Lett.*, **54**(2), 215-226.
- [3] CHANDA, K.C. (1983). Density estimation for linear processes. *Ann. Inst. Statist. Math.*, **35**(3), 439-446.
- [4] DEVROYE, L. AND GYORFI, L. (1985). *Nonparametric density estimation: The L1 view*. Wiley, New York.
- [5] ENGLE, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
- [6] HALLIN, M. AND TRAN, L.T. (1996). Kernel density estimation for linear processes: asymptotic normality and optimal bandwidth derivation. *Ann. Inst. Statist. Math.*, **48**(3), 429-449.
- [7] HONDA, T. (2000). Nonparametric density estimation for a long-range dependent linear process. *Annals of the Institute of Statistical Mathematics*, **52**(4), 599-611.
- [8] HONDA, T. (2009). Nonparametric density estimation for linear processes with infinite variance. *Annals of the Institute of Statistical Mathematics*, **61**(2), 413-439.



- [9] LU, Z. (2001). Asymptotic normality of kernel density estimators under dependence. *Annals of the Institute of Statistical Mathematics*, **53**(3), 447-468.
- [10] SCHICK, A. AND WEFELMEYER, W. (2006). Pointwise convergence rates and central limit theorems for kernel density estimators in linear processes. *Statistics and probability letters*, **76**(16), 1756-1760.
- [11] WEISS, A. A. (1984). Arma models with arch errors. *Journal of time series analysis*, **5**(2), 129-143.
- [12] WU, W. B. AND MIELNICZUK, J. (2002). Kernel density estimation for linear processes. *The Annals of Statistics*, **30**(5), 1441-1459.
- [13] WU, W. B., HUANG, Y., AND HUANG, Y. (2010). Kernel estimation for time series: An asymptotic theory. *Stochastic Processes and their Applications*, **120**(12), 2412-2431.