Universidade de Vigo

CLT in nonlinear wavelet regression with left-truncated α-mixing observations

Han-Ying Liang and Jacobo de Uña Álvarez

Report 08/10

Discussion Papers in Statistics and Operation Research

Departamento de Estatística e Investigación Operativa

Facultade de Ciencias Económicas e Empresariales
Lagoas-Marcosende, s/n · 36310 Vigo
Tfno.: +34 986 812440 - Fax: +34 986 812401
http://eioweb.uvigo.es/
E-mail: depc05@uvigo.es

Universidade de Vigo

CLT in nonlinear wavelet regression with left-truncated α-mixing observations

Han-Ying Liang and Jacobo de Uña Álvarez

Report 08/10

Discussion Papers in Statistics and Operation Research

Imprime: GAMESAL

Edita:

Universidade de Vigo
Facultade de CC. Económicas e Empresariales
Departamento de Estatística e Investigación Operativa
As Lagoas Marcosende, s/n 36310 Vigo
Tfno.: +34 986 812440

I.S.S.N: 1888-5756

Depósito Legal: VG 1402-2007

CLT in nonlinear wavelet regression with left-truncated α-mixing observations

Han-Ying Liang[a,b] Jacobo de Uña-Álvarez[b]

[a]Department of Mathematics, Tongji University, Shanghai 200092, P. R. China

E-mail: hyliang83@yahoo.com

[b]Facultad de Ciencias Económicas y Empresariales, Universidad de Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Spain *E-mail: jacobo@uvigo.es*

Abstract. In this paper we define a new nonlinear wavelet-based estimator of the regression function under a left truncation model. Asymptotic normality of the estimator is established. It is assumed that the lifetime observations form a stationary α-mixing sequence. Also, the asymptotic normality of the nonlinear wavelet-based estimator of the covariate's density is proved.

Key words and phrases: Asymptotic normality, nonlinear wavelets; nonparametric regression; truncated data; α-mixing sequence

AMS 2000 subject classifications: Primary: 62G07; Secondary: 62G20.

1 Introduction

The importance of wavelets in curve estimation is well known since the initial works by Kerkyacharian and Picard (1992, 1993), Donoho and Johnstone (1994, 1995), and Donoho et al. (1995, 1996). In these papers, adaptation of wavelets (in the minimax sense) to the degree of smooth of the underlying function is analyzed, for a wide range of functional spaces and a number of loss functions. This is a remarkable property of the wavelet method when compared to other common estimation techniques (such as the kernel method) which may fail in unsmooth situations. Hall and Patil (1995) gave for the first time an asymptotic expression of the mean integrated squared error (MISE) of a nonlinear wavelet density estimator, and compared its performance to that corresponding to the kernel density estimator. These authors showed that the asymptotic MISE formula is the same in both the smooth and unsmooth density cases, a fact that is not true for the kernel method. Similar results are available for the problem of estimating a regression function, see for example Hall and Patil (1996).

In Reliability and Survival Analysis, incomplete data are often encountered, censoring and truncation being two important sources of incompleteness. Some authors have studied wavelet density estimation with censored data. For example, Antoniadis et al. (1999) considered linear wavelet density estimation under random censoring, and provided asymptotic MISE convergence rate under smoothness assumptions on the underlying density function. Li (2003) proposed a non-linear wavelet density estimator with censored data and derived a result similar to the main result, Theorem 2.1, of Hall and Patil (1995), about the MISE; see also Rodríguez-Casal and de Uña-Álvarez (2004) who considered the Koziol-Green model of random censorship. All of the above works are devoted to the independent setting. For the dependent case, Liang et al. (2005) discussed the global L_2 error of the nonlinear wavelet estimators of the density function in the Besov space under censoring and stationary α-mixing assumptions; for complete data, Li and Xiao (2006, 2007) provided an asymptotic expression of the MISE in nonlinear wavelet regression with complete long memory data; while Truoug and Patil (2001) gave the result corresponding to α-mixing data. However, little is known about the wavelet-based estimator of the covariate's distribution and the regression function under random truncation, even for independent data. In this paper we introduce nonlinear wavelet estimators for left-truncated data and we investigate asymptotic properties under mixing conditions.

Let Y be a lifetime variable with continuous distribution function (df) F and let X be a one-dimensional covariate taking its values in $\mathbb{R}$ with df V and corresponding density v. Introduce the regression function of Y given X:

$$\mathbb{E}(Y|X=x) := m(x) \ , \ x \in \mathbb{R}, \tag{1.1}$$

which can be written as $m(x) = \frac{h(x)}{v(x)}$, where $h(x) = \int_{\mathbb{R}} y f(x,y) dy$ with $f(\cdot,\cdot)$ being the joint density function of (X,Y). In practice, the response lifetime variable Y may be subject to right censoring and/or left truncation. In this paper we consider the left truncation model. Left-truncated data occur in astronomy, economics, epidemiology and biometry; see, e.g., Woodroofe (1985), Feigelson and Babu (1992), Wang et al. (1986), Tsai et al. (1987) and He and Yang (1994).

Under the assumption that the lifetime observations are mutually independent, the nonparametric kernel estimator of $m(\cdot)$ has first been considered for complete data; see, e.g., Devroye et al. (1996), Györfi et al. (1998) and the references therein. For censored data, the estimation of the regression function $m(\cdot)$ has also been studied by Fan and Gijbels (1996) and the recent work of Kohler et al. (2002), and many others. Recently, Ould-Saïd and Lemdani (2006) constructed a new nonparametric kernel estimator of the regression function $m(\cdot)$ for the left-truncation model and studied its asymptotic properties. In this paper we define a new nonlinear wavelet-based estimator of the regression function under the left-truncation model, and establish asymptotic normality for the estimator of the regression function when the data exhibit some kind of dependence. It is assumed that the lifetime observations form a stationary

α-mixing sequence. Also, the asymptotic normality of the nonlinear wavelet-based estimator of the covariable's density is considered.

In the sequel, $\{(X_k, Y_k, T_k) =: \xi_k,\ k \geq 1\}$ is assumed to be a stationary α-mixing sequence of random vectors distributed as (X, Y, T), where T is the truncation variable. For the components of (X, Y, T), in addition to the assumptions and notation for X and Y we made above, we assume throughout that T and (X, Y) are independent, and T has continuous df G. Let $F(\cdot, \cdot)$ be the joint df of (X, Y). Without loss of generality, we assume that both Y and T are nonnegative random variables. Recall that the sequence $\{\xi_k, k \geq 1\}$ is said to be α-mixing if the α-mixing coefficient

$$\alpha(n) := \sup_{k\geq 1} \sup\{|P(AB) - P(A)P(B)| : A \in \mathcal{F}_{n+k}^{\infty}, B \in \mathcal{F}_1^k\}$$

converges to zero as $n \to \infty$, where $\mathcal{F}_l^m$ denotes the σ-algebra generated by $\xi_l, \xi_{l+1}, \ldots, \xi_m$ with $l \leq m$. Among various mixing conditions used in the literature, α-mixing is reasonably weak and has many practical applications; see, e.g., Doukhan (1994), page 99, for more details; and see Cai and Kim (2003) for motivation in the scope of Survival Analysis. In particular, the stationary autoregressive-moving average (ARMA) processes, which are widely applied in time series analysis, are α-mixing with exponential mixing coefficient, i.e., $\alpha(k) = \rho^k$ for some $0 < \rho < 1$.

The rest of this paper is organized as follows. In the next section, we give some notation for the left-truncation model. Basic elements of the wavelet theory, and the definition of the nonlinear wavelet-based estimators of $m(\cdot)$, $v(\cdot)$ and $h(\cdot)$ are given too. Main results are described in Section 3, while their proofs are given in Section 4. In the Appendix, we collect some preliminary lemmas, which are used in Section 4.

2 Notation and wavelet-based estimators

In the random left-truncation model, the lifetime Y_i is interfered by the truncation random variable T_i in such a way that both Y_i and T_i are observable only when $Y_i \geq T_i$, whereas nothing is observed if $Y_i < T_i$ for $i = 1, \cdots, N$, where N is the potential sample size. Due to the occurrence of truncation, the N is unknown, and n (the size of the actually observed sample) is random with $n \leq N$. Let $\theta = \mathbb{P}(Y \geq T)$ be the probability that the random variable Y is observable. Since $\theta = 0$ implies that no data can be observed, we suppose throughout the paper that $\theta > 0$. Without confusion we still denote $(X_i, Y_i, T_i), i = 1, \cdots, n$ the observed sequence.

Following Stute (1993) the conditional dfs of Y and T given no occurrence of the truncation are

$$F^*(y) = \mathbb{P}(Y \le y | Y \ge T) = \theta^{-1} \int_0^y G(u) dF(u)$$

and $G^*(y) = \mathbb{P}(T \le y | Y \ge T) = \theta^{-1} \int_0^\infty G(y \wedge u) dF(u)$, which can be estimated by

$$F_n^*(y) = n^{-1} \sum_{i=1}^n I(Y_i \le y) \quad \text{and} \quad G_n^*(y) = n^{-1} \sum_{i=1}^n I(T_i \le y),$$

respectively, where $I(\cdot)$ is the indicator function. Since $C(y) = \mathbb{P}(T \le y \le Y | Y \ge T) = \theta^{-1} G(y)[1 - F(y-)] = G^*(y) - F^*(y-)$, the empirical estimator of $C(y)$ is defined by

$$C_n(y) = n^{-1} \sum_{i=1}^n I(T_i \le y \le Y_i) = G_n^*(y) - F_n^*(y-),$$

where for any df H, $H(y-)$ denotes the left-limit of H at y.

Since N is unknown and n is known (although random), our results would not be stated with respect to the probability measure $\mathbb{P}$ (related to the N-sample) but will involve the conditional probability P with respect to the actually observed n-sample. Also $\mathbb{E}$ and E will denote the expectation operators under $\mathbb{P}$ and P, respectively.

Following the idea of Lynden-Bell (1971), the nonparametric maximum likelihood estimators of F and G are given by

$$1 - F_n(y) = \prod_{i:Y_i \le y} \left(1 - \frac{1}{nC_n(Y_i)}\right) \quad \text{and} \quad G_n(y) = \prod_{i:T_i > y} \left(1 - \frac{1}{nC_n(T_i)}\right).$$

The estimator of θ is defined by $\theta_n = G_n(y)[1 - F_n(y-)] C_n^{-1}(y)$. He and Yang (1998) proved that θ_n does not depend on y and that it is well-defined whenever $C_n(y) \neq 0$. For any df H, let $a_H = \inf\{y : H(y) > 0\}$ and $b_H = \sup\{y : H(y) < 1\}$ be its two endpoints. Consequently, θ is identifiable only if $a_G \le a_F$ and $b_G \le b_F$.

Now we introduce some notation corresponding to wavelets. Let $\phi(x)$ and $\psi(x)$ be father and mother wavelets, having the properties: $\phi(\cdot)$ and $\psi(\cdot)$ are bounded and compactly supported; $\int \phi^2 = \int \psi^2 = 1$, $\mu_k = \int y^k \psi(y) dy = 0$ for $0 \le k \le r - 1$ and $\mu_r = r!\kappa \neq 0$, where $\kappa = (r!)^{-1} \int y^r \psi(y) dy$. Let

$$\phi_{mj}(x) = 2^{m/2} \phi(2^m x - j), \quad \psi_{ij}(x) = 2^{i/2} \psi(2^i x - j), \quad x \in R, \quad m, i \in \mathbb{Z}, \tag{2.1}$$

then, the collection $\{\phi_{mj}, \psi_{ij}, j \in \mathbb{Z}, i \ge m\}$ is an orthonormal basis (ONB) of $L_2(\mathbb{R})$. Furthermore, let V_m and W_i be linear subspaces of $L_2(\mathbb{R})$ with the ONB ϕ_{mj}, $j \in \mathbb{Z}$ and ψ_{ij}, $j \in \mathbb{Z}$ respectively, we have the following decomposition $L_2(\mathbb{R}) = V_m \oplus W_m \oplus W_{m+1} \oplus W_{m+2} \oplus \cdots$. For more on wavelets see Daubechies (1992) or Härdle et al. (1998).

Remark 2.1 *Set $K(t,x) = \sum_{j=-\infty}^{\infty} \phi(t-j)\phi(x-j)$. Then, $K(t,x)$ has the following properties: (i) $K(t,x)$ is uniformly bounded; (ii) $K(t,x) = 0$ for $|t-x| > 4L$ if support $\phi(x) \subset [-L, L]$; (iii) $K(t,x)$ satisfies the moment condition (cf. Härdle et al. (1998), Theorem 8.3, page 93), i.e., $\int (t-x)^k K(t,x)dt = \delta_{0k}$ for $k = 0, 1, \cdots, r-1$, where δ_{ij} denotes the Kronecker delta [i.e. $\delta_{ij} = 1$, if $i = j$; 0, otherwise].*

If the function $v(\cdot)$ belongs to $L_2(\mathbb{R})$, we have the following wavelet expansion:

$$v(x) = \sum_{j=-\infty}^{\infty} a_{mj}\phi_{mj}(x) + \sum_{i=m}^{\infty}\sum_{j=-\infty}^{\infty} a_{ij}\psi_{ij}(x), \tag{2.2}$$

where $a_{mj} = \int v(x)\phi_{mj}(x)dx = \int \phi_{mj}(x)V(dx)$ and $a_{ij} = \int v(x)\psi_{ij}(x)dx = \int \psi_{ij}(x)V(dx)$ are the wavelet coefficients of the function $v(\cdot)$ and the series in (2.2) converges in $L_2(\mathbb{R})$. These coefficients can be estimated by replacing V by a proper empirical df. In our random left-truncation model, to build estimator of $V(\cdot)$, we first consider the conditional joint distribution of (X, Y, T)

$$H^*(x,y,t) = \mathbb{P}(X \le x, Y \le y, T \le t | Y \ge T) = \frac{1}{\theta}\int_{u \le x}\int_{a_G \le w \le y} G(w \wedge t)F(du, dw).$$

Taking $t = +\infty$, we get the conditional joint df of (X, Y)

$$F^*(x,y) := \mathbb{P}(X \le x, Y \le y | Y \ge T) = \theta^{-1}\int_{u \le x}\int_{a_G \le w \le y} G(w)F(du, dw),$$

from which we obtain

$$F(dx, dy) = \theta G^{-1}(y)F^*(dx, dy) \quad \text{for} \quad y > a_G. \tag{2.3}$$

Integrating over y we get the df of X, $V(x) = \theta \int_{u \le x}\int_{y \ge a_G} \frac{1}{G(y)} F^*(du, dy)$. A natural estimator of V is then given by

$$V_n(x) = \frac{\theta_n}{n}\sum_{k=1}^{n} G_n^{-1}(Y_k)I(X_k \le x). \tag{2.4}$$

Note that in Eq. (2.4) and the forthcoming formulae, the sum is taken only for k such that $G_n(Y_k) \ne 0$. In view of (2.4), we define the non-linear wavelet estimator of $v(x)$ as

$$\hat{v}_n(x) = \sum_{j=-\infty}^{\infty} \hat{a}_{mj}\phi_{mj}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \hat{a}_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x), \tag{2.5}$$

where $\delta > 0$ is a "threshold" and $\pi \ge 1$ is another smoothing parameter, and the wavelet coefficients $\hat{a}_{mj}$ and $\hat{a}_{ij}$ are defined as follows:

$$\hat{a}_{mj} = \int \phi_{mj} dV_n = \frac{\theta_n}{n}\sum_{k=1}^{n}\frac{1}{G_n(Y_k)}\phi_{mj}(X_k), \ \hat{a}_{ij} = \int \psi_{ij} dV_n = \frac{\theta_n}{n}\sum_{k=1}^{n}\frac{1}{G_n(Y_k)}\psi_{ij}(X_k). \tag{2.6}$$

If the function h is square-integrable then its wavelet expansion is given by

$$h(x) = \sum_{j=-\infty}^{\infty} b_{mj}\phi_{mj}(x) + \sum_{i=m}^{\infty}\sum_{j=-\infty}^{\infty} b_{ij}\psi_{ij}(x), \tag{2.7}$$

where $b_{mj} = \int h(x)\phi_{mj}(x)dx$ and $b_{ij} = \int h(x)\psi_{ij}(x)dx$. Note that

$$H_n(x) = \theta_n n^{-1}\sum_{k=1}^{n} Y_k G_n^{-1}(Y_k) I(X_k \le x)$$

is an estimator of $H(x) = \int_{u\le x} h(u)du$ (see Ould-Saïd and Lemdani (2006)). So, the proposed non-linear wavelet estimator of $h(x)$ is

$$\hat{h}_n(x) = \sum_{j=-\infty}^{\infty} \hat{b}_{mj}\phi_{mj}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \hat{b}_{ij} I(|\hat{b}_{ij}| > \delta)\psi_{ij}(x), \tag{2.8}$$

where $\hat{b}_{mj} = \theta_n n^{-1}\sum_{k=1}^{n} Y_k G_n^{-1}(Y_k)\phi_{mj}(X_k)$, $\hat{b}_{ij} = \theta_n n^{-1}\sum_{k=1}^{n} Y_k G_n^{-1}(Y_k)\psi_{ij}(X_k)$. Further, from (2.5) and (2.8), a wavelet estimator of $m(x)$ is given by $\hat{m}(x) = \hat{h}_n(x)/\hat{v}_n(x)$.

3 Main results

In the sequel, let $C, C_0, C_1, \cdots$ and c denote generic finite positive constants, whose values are unimportant and may change from line to line, $A_n = O(B_n)$ stand for $A_n \le CB_n$, $a_n \asymp b_n$ mean $0 < \liminf a_n/b_n \le \limsup a_n/b_n < \infty$. Throughout this paper, we assume that

$$a_G < a_F, \quad b_G \le b_F < \infty. \tag{3.1}$$

Let $x_m = [2^m x]/2^m$ for $x \in \mathbb{R}$ and

$$\Sigma(x) = \begin{pmatrix} \Sigma_0(x) & \Sigma_1(x) \\ \Sigma_1(x) & \Sigma_2(x) \end{pmatrix}, \ \Sigma_i(x) = \int\int \frac{y^{2-i} f(x,y)}{G(y)}\Big[\sum_l \phi(u+l)\phi(l)\Big]^2 dudy \ \ (i = 0, 1, 2).$$

In order to formulate the main results, we need to impose the following assumptions.

(A1) For all integers $j \ge 1$, the joint conditional density $v_j^*(\cdot,\cdot)$ of X_1 and X_{j+1} exists on $R \times R$ and satisfies $v_j^*(t_1, t_2) \le C$ for all $t_1, t_2 \in \mathbb{R}$ with $|t_1 - t_2| \le \delta_0$ for some $\delta_0 > 0$.

(A2) (i) The density $v(\cdot)$ satisfies $0 < v(x) \le C_2$ for $x \in \mathbb{R}$;

(ii) $f(x, y)$ is bounded and continuous with respect to the first component.

(A3) The smoothing parameters π and δ are functions of n. Suppose that $\pi \to \infty$ as $n \to \infty$ in such a manner that $2^\pi\delta^2 = O(n^{-\epsilon})$ for some $0 < \epsilon < 1$, $\delta \ge C_3(n^{-1}\ln(n))^{1/2}$.

(A4) The sequence $\alpha(n)$ satisfies

(i) there exist positive integers $p := p(n)$ and $q := q(n)$ such that $p + q \leq n$, and as $n \to \infty$, $p/n \to 0$, $qp^{-1} \to 0$ and $(n/p)\alpha(q) \to 0$;

(ii) there exist $\gamma > 2$ and $\eta > 1 - 2/\gamma$ such that $\sum_{l=1}^{\infty} l^{\eta}[\alpha(l)]^{1-2/\gamma} < \infty$.

3.1 Asymptotic normality of covariate's density

Theorem 3.1 *In addition to the conditions on $\phi(\cdot)$ and $\psi(\cdot)$ stated in Section 2 and the assumptions (A1)-(A4) and equation (3.1). Assume that the r-th derivative $v^{(r)}(\cdot)$ of $v(\cdot)$ is continuous and bounded. Let $\alpha(k) = O(k^{-\lambda})$ for some*

$$\lambda > \max\{3, d(d+\mu)/(2\mu), 1 + 4r/[\epsilon(2r+1)], (\tau-1)(2\tau+1)(2-\epsilon)/(2\epsilon(\tau-2))\}, \tag{3.2}$$

where $\tau > 2$, $d > 2$, $\mu > 0$, and

$$\epsilon(\lambda + 1 + 2b) + 2b/(2r+1) \geq 2(b+1) \quad \text{for} \quad b > 1. \tag{3.3}$$

If $2^m \asymp n^{1/(2r+1)}$ and $(p2^m/n)^{d/2-1} 2^{m\mu/(d+\mu)} \to 0$, then

$$\sqrt{n2^{-m}}(\hat{v}_n(x_m) - v(x_m) - a(x_m)) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x)) \quad x \in \mathbb{R},$$

where $a(x) = (r!)^{-1} v^{(r)}(x) 2^{-rm} \int u^r [\sum_l \phi(u+l)\phi(l)] du$ and $\sigma^2(x) = \theta \Sigma_2(x)$. Further, if $n2^{-(2r+1)m} \to 0$, then $\sqrt{n2^{-m}}(\hat{v}_n(x_m) - v(x_m)) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x))$.

3.2 Asymptotic normality of regression function

Theorem 3.2 *Suppose that the assumptions in Theorem 3.1 are satisfied, and that the r-th derivative $h^{(r)}(\cdot)$ of $h(\cdot)$ is continuous and bounded. If*

$$2^m \asymp (n \ln(n))^{1/(2r+1)} \quad \text{and} \quad (p2^m/n)^{d/2-1} 2^{m\mu/(d+\mu)} \to 0,$$

then $\sqrt{n2^{-m}}(\hat{m}_n(x_m) - m(x_m)) \xrightarrow{\mathcal{D}} N(0, \Delta^2(x))$ $x \in \mathbb{R}$, where

$$\Delta^2(x) = \frac{\theta[\Sigma_0(x) v^2(x) + \Sigma_2(x) h^2(x) - 2v(x)h(x)\Sigma_1(x)]}{v^4(x)}.$$

Remark 3.1 *In Theorems 3.1 and 3.2, if we replace $\alpha(k) = O(k^{-\lambda})$ by $\alpha(k) = O(\rho^k)$ for some $0 < \rho < 1$, then (3.2) and (3.3) are automatically satisfied.*

4 Proofs of main results

We are now ready to prove our main results.

Proof of Theorem 3.1. By using Lemma 5.7 it follows that

$$\begin{aligned}
& \hat{v}_n(x) - v(x) - a(x) \\
= & \sum_{j=-\infty}^{\infty} (\hat{a}_{mj} - a_{mj})\phi_{mj}(x) + \Big[\sum_{j=-\infty}^{\infty} a_{mj}\phi_{mj}(x) - v(x) - a(x)\Big] + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \hat{a}_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) \\
= & \sum_{j=-\infty}^{\infty} (\tilde{a}_{mj} - a_{mj})\phi_{mj}(x) + \Big[\sum_{j=-\infty}^{\infty} a_{mj}\phi_{mj}(x) - v(x) - a(x)\Big] \\
& + \sum_{j=-\infty}^{\infty} A_{mj}\phi_{mj}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} A_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \tilde{a}_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) \\
:= & I_1(x) + I_2(x) + I_3(x) + I_4(x) + I_5(x).
\end{aligned}$$

It suffices to show that

$$\sqrt{n2^{-m}} I_1(x_m) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x)),\ I_2(x_m) = o(2^{-rm}),$$
$$\sqrt{n2^{-m}} I_3(x) = o_p(1),\ \sqrt{n2^{-m}} I_4(x) = o_p(1),\ \sqrt{n2^{-m}} I_5(x) = o_p(1).$$

Step 1. We prove $\sqrt{n2^{-m}} I_1(x_m) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x))$. It is easy to see that

$$\sqrt{n2^{-m}} I_1(x_m) = \sum_{k=1}^{n} \Big\{ \sqrt{\frac{2^m}{n}} \Big(\frac{\theta K(2^m X_k, 2^m x_m)}{G(Y_k)} - \int v(t) K(2^m t, 2^m x_m) dt \Big) \Big\} := \sum_{k=1}^{n} Z_{nk},$$

where $K(t,x) = \sum_{j=-\infty}^{\infty} \phi(t-j)\phi(x-j)$, $Z_{nk} = \sqrt{\frac{2^m}{n}} \Big(\frac{\theta K(2^m X_k, 2^m x_m)}{G(Y_k)} - \int v(t) K(2^m t, 2^m x_m) dt \Big)$ and $EZ_{nk} = 0$. Let $w := w_n = [\frac{n}{p+q}]$ and $\sum_{k=1}^{n} Z_{nk} = S_n' + S_n'' + S_n'''$, where

$$\begin{aligned}
S_n' &= \sum_{l=1}^{w} y_{nl},\ S_n'' = \sum_{l=1}^{w} y_{nl}',\ S_n''' = y_{nw+1}', \\
y_{nl} &= \sum_{i=s_l}^{s_l+p-1} Z_{ni},\ y_{nl}' = \sum_{i=t_l}^{t_l+q-1} Z_{ni},\ y_{nw+1}' = \sum_{i=w(p+q)+1}^{n} Z_{ni}
\end{aligned}$$

and $s_l = (l-1)(p+q)+1,\ t_l = (l-1)(p+q)+p+1,\ l = 1, \ldots, w$. It is sufficient to prove that

$$E(S_n'')^2 \to 0, \quad E(S_n''')^2 \to 0, \tag{4.1}$$

$$\mathrm{Var}(S_n') \to \sigma^2(x), \tag{4.2}$$

$$\Big| E\exp\Big(it\sum_{l=1}^{w} y_{nl} \Big) - \prod_{l=1}^{w} E\exp(ity_{nl}) \Big| \to 0, \tag{4.3}$$

$$A_n(\epsilon) = \sum_{l=1}^{w} Ey_{nl}^2 I(|y_{nl}| > \epsilon\sigma(x)) \to 0 \ \ \forall \epsilon > 0. \tag{4.4}$$

Relation (4.1) implies that S''_n and S'''_n are asymptotically negligible, (4.3) shows that the summands y_{nl} in S'_n are asymptotically independent, and (4.2) and (4.4) are the standard Lindeberg-Feller conditions for asymptotic normality of S'_n under independence.

We first establish (4.1). We observe that

$$\begin{aligned} E(S''_n)^2 &= \sum_{l=1}^{w}\sum_{i=t_l}^{t_l+q-1} EZ_{ni}^2 + 2\sum_{l=1}^{w}\sum_{t_l\le i<j\le t_l+q-1} \mathrm{Cov}(Z_{ni},Z_{nj}) + 2\sum_{1\le i<j\le w}\mathrm{Cov}(y'_{ni},y'_{nj}) \\ &:= I_{11}(x_m)+I_{12}(x_m)+I_{13}(x_m). \end{aligned} \tag{4.5}$$

From $K(t,x)=K(t+k,x+k)$ for $k\in\mathbb{Z}$ and (2.3), it follows that

$$\begin{aligned} nEZ_{ni}^2 &= 2^m E\Big(\frac{\theta^2K^2(2^mX_1,2^mx_m)}{G^2(Y_1)}\Big) - 2^m\Big(\int v(t)K(2^mt,2^mx_m)dt\Big)^2 \\ &= 2^m\theta\int\int\frac{f(t,y)}{G(y)}K^2(2^mt,2^mx_m)dtdy - 2^m\Big(\int v(t)K(2^mt,2^mx_m)dt\Big)^2 \\ &= 2^m\theta\int\int\frac{f(t,y)}{G(y)}K^2(2^mt-2^mx_m,0)dtdy - 2^m\Big(\int v(t)K(2^mt-2^mx_m,0)dt\Big)^2 \\ &= \theta\int\int\frac{f(2^{-m}u+x_m,y)}{G(y)}K^2(u,0)dudy - \frac{1}{2^m}\Big(\int v(2^{-m}u+x_m)K(u,0)du\Big)^2 \\ &\to \theta\int\int\frac{f(x,y)}{G(y)}K^2(u,0)dudy = \theta\int\int\frac{f(x,y)}{G(y)}\Big[\sum_l\phi(u+l)\phi(l)\Big]^2dudy, \end{aligned} \tag{4.6}$$

which yields that $I_{11}(x_m)=O(wq/n)=O(q/p)=o(1)$. Note that

$$|I_{12}(x_m)|\le 2\sum_{1\le i<j\le n}|\mathrm{Cov}(Z_{ni},Z_{nj})| \quad\text{and}\quad |I_{13}(x_m)|\le 2\sum_{1\le i<j\le n}|\mathrm{Cov}(Z_{ni},Z_{nj})|.$$

Therefore, in order to prove $I_{12}(x_m)=o(1)$ and $I_{13}(x_m)=o(1)$, we need only to show that

$$\sum_{1\le i<j\le n}|\mathrm{Cov}(Z_{ni},Z_{nj})|\to 0. \tag{4.7}$$

Next, let c_n (specified below) be a sequence of integers such that $c_n\to\infty$ and $c_n2^{-m}\to 0$. Put

$$\begin{aligned} S_1 &= \{(i,j)|i,j\in\{1,2,\cdots,n\},1\le j-i\le c_n\}, \\ S_2 &= \{(i,j)|i,j\in\{1,2,\cdots,n\},c_n+1\le j-i\le n-1\}. \end{aligned}$$

We write

$$\sum_{1\le i<j\le n}|\mathrm{Cov}(Z_{ni},Z_{nj})| = \sum_{S_1}|\mathrm{Cov}(Z_{ni},Z_{nj})| + \sum_{S_2}|\mathrm{Cov}(Z_{ni},Z_{nj})|. \tag{4.8}$$

From (A1), similarly as in (4.6), for $i < j$ we have

$$
\begin{aligned}
&|\mathrm{Cov}(Z_{ni}, Z_{nj})| \\
&= \frac{2^m}{n}\Big\{E\Big(\frac{\theta K(2^m X_i, 2^m x_m)}{G(Y_i)} \cdot \frac{\theta K(2^m X_j, 2^m x_m)}{G(Y_j)}\Big) - \Big(\int v(t)K(2^m t, 2^m x_m)dt\Big)^2\Big\} \\
&\leq \frac{2^m}{n}\Big\{\frac{\theta^2}{G^2(a_F)} E|K(2^m X_i, 2^m x_m)K(2^m X_j, 2^m x_m)| + \Big(\int v(t)K(2^m t, 2^m x_m)dt\Big)^2\Big\} \\
&= \frac{2^m}{n}\Big\{\frac{\theta^2}{G^2(a_F)} \int\int |K(2^m t_1, 2^m x_m)K(2^m t_2, 2^m x_m)| v^*_{j-i}(t_1, t_2)dt_1 dt_2 \\
&\quad + \Big(\int v(t)K(2^m t, 2^m x_m)dt\Big)^2\Big\} \\
&\leq \frac{2^m}{n}\Big\{C\int\int |K(2^m t_1 - 2^m x_m, 0)K(2^m t_2 - 2^m x_m, 0)|dt_1 dt_2 \\
&\quad + \Big(\int v(t)K(2^m t - 2^m x_m, 0)dt\Big)^2\Big\} = O((n2^m)^{-1}). \qquad (4.9)
\end{aligned}
$$

Hence

$$\sum_{S_1} |\mathrm{Cov}(Z_{ni}, Z_{nj})| = O(c_n 2^{-m}) \to 0. \qquad (4.10)$$

On the other hand, it follows from Lemma 5.1 that $|\mathrm{Cov}(Z_{ni}, Z_{nj})| \leq C[\alpha(j-i)]^{1-2/\gamma}(E|Z_{ni}|^\gamma)^{2/\gamma}$ and

$$
\begin{aligned}
E|Z_{ni}|^\gamma &\leq 2^\gamma\Big(\frac{2^m}{n}\Big)^{\gamma/2} E\Big|\frac{\theta K(2^m X_k, 2^m x_m)}{G(Y_k)}\Big|^\gamma \leq \Big(\frac{2^m}{n}\Big)^{\gamma/2}\frac{2^\gamma\theta^\gamma}{G^\gamma(a_F)}\int |K(2^m t, 2^m x_m)|^\gamma v(t)dt \\
&= \Big(\frac{2^m}{n}\Big)^{\gamma/2}\frac{2^\gamma\theta^\gamma}{G^\gamma(a_F)} \cdot 2^{-m}\int |K(u, 0)|^\gamma v(2^{-m}u + x_m)du = O(n^{-1}(2^m/n)^{\gamma/2-1}). \quad (4.11)
\end{aligned}
$$

Then

$$
\begin{aligned}
\sum_{S_2} |\mathrm{Cov}(Z_{ni}, Z_{nj})| &\leq C\sum_{j=1}^{n}\sum_{j-i=c_n+1}^{n-1} [\alpha(j-i)]^{1-2/r} n^{-1} 2^{(1-2/\gamma)m} \\
&\leq C2^{(1-2/\gamma)m}\sum_{l=c_n}^{\infty}[\alpha(l)]^{1-2/\gamma} \leq Cc_n^{-\eta}2^{(1-2/\gamma)m}\sum_{l=c_n}^{\infty} l^\eta[\alpha(l)]^{1-2/\gamma}.
\end{aligned}
$$

Hence, by choosing $c_n = 2^{m(1-2/\gamma)/\eta}$, from (A4) it follows that

$$\sum_{S_2} |\mathrm{Cov}(Z_{ni}, Z_{nj})| \to 0. \qquad (4.12)$$

Therefore, (4.7) follows from (4.8), (4.10) and (4.12).

As to $E(S_n''')^2$, from (4.6), (4.7) and (A4)(i) we have

$$
\begin{aligned}
E(S_n''')^2 &= \sum_{i=w(p+q)+1}^{n} \mathrm{Var}(Z_{ni}) + 2\sum_{w(p+q)+1\leq i<j\leq n} \mathrm{Cov}(Z_{ni}, Z_{nj}) \\
&\leq C \cdot \frac{n - w(p+q)}{n} + 2\sum_{1\leq i<j\leq n} |\mathrm{Cov}(Z_{ni}, Z_{nj})| \to 0
\end{aligned}
$$

and (4.1) is proved. Besides, since $wp/n \to 1$, (4.2) follows from (4.6) and (4.7). As to (4.3), according to Lemma 5.2 we have

$$\Big|E\exp\Big(it\sum_{l=1}^{w} y_{nl}\Big) - \prod_{l=1}^{w} E\exp(ity_{nl})\Big| \le 16w\alpha(q+1) \le C(n/p)\alpha(q),$$

which tends to zero by (A4)(i).

Finally, we establish (4.4). Here, we use Lemma 5.3 to evaluate $A_n(\epsilon)$. Let $d > 2$ and $\mu > 0$ such that $\lambda > \frac{d(d+\mu)}{2\mu}$, by using Lemma 5.3 and (4.11), for sufficiently small $\beta > 0$ we have

$$\begin{aligned}
A_n(\epsilon) &\le (\epsilon\sigma(x))^{2-d}\sum_{l=1}^{w} E|y_{nl}|^d \\
&\le C(\epsilon\sigma(x))^{2-d}\sum_{l=1}^{w}\Big\{p^{\beta}\sum_{i=s_l}^{s_l+p-1} E|Z_{ni}|^d + \Big(\sum_{i=s_l}^{s_l+p-1}\|Z_{ni}\|_{d+\mu}^2\Big)^{d/2}\Big\} \\
&\le C\Big\{wp^{\beta+1}n^{-1}(2^m/n)^{d/2-1} + wp^{d/2}\Big(n^{-1}(2^m/n)^{(d+\mu)/2-1}\Big)^{d/(d+\mu)}\Big\} \\
&\le C(p2^m/n)^{d/2-1}2^{m\mu/(d+\mu)} \to 0.
\end{aligned}$$

Step 2. We verify $I_2(x_m) = o(2^{-rm})$. From Remark 2.1, by applying an argument similar as that in (4.6) and a Taylor expansion, it follows that

$$\begin{aligned}
I_2(x_m) &= \int [v(t) - v(x_m)]2^m K(2^m t, 2^m x_m)dt - a(x_m) \\
&= \int [v(x_m + 2^{-m}u) - v(x_m)]K(u,0)du - a(x_m) \\
&= \int \sum_{k=1}^{r}\frac{v^{(k)}(x_m)}{k!}2^{-km}u^k K(u,0)du + o(2^{-rm}) - a(x_m) \\
&= \frac{v^{(r)}(x_m)}{r!}2^{-rm}\int u^r K(u,0)du + o(2^{-rm}) - a(x_m) = o(2^{-rm}).
\end{aligned}$$

Step 3. We prove $\sqrt{n2^{-m}}I_3(x) = o_p(1)$ and $\sqrt{n2^{-m}}I_4(x) = o_p(1)$. From Lemma 5.7 we have

$$\sqrt{n2^{-m}}|I_3(x)| = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big)\sqrt{n2^{-m}}\cdot\frac{\theta}{n}\sum_{k=1}^{n}\sum_{j=-\infty}^{\infty} 2^m\frac{|\phi(2^m X_k - j)|}{G(Y_k)}|\phi(2^m x - j)| \quad a.s.$$

and, for each m, since the support of $\phi(\cdot)$ is compact, there exists a finite number of non-zero j terms of the form $\phi(2^m x - j)$. Hence, according to

$$\theta 2^m E\Big(\frac{|\phi(2^m X_k - j)|}{G(Y_k)}\Big) = 2^m\int|\phi(2^m x - j)|v(x)dx = \int|\phi(u)|v((u+j)/2^m)du = O(1)$$

we have $\sqrt{n2^{-m}}I_3(x) = O_p((2^{-m}\ln\ln(n))^{1/2}) = o_p(1)$.

Note that

$$\sqrt{n2^{-m}}|I_4(x)| = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big)\sqrt{n2^{-m}} \cdot \frac{\theta}{n}\sum_{i=m}^{\pi}\sum_{k=1}^{n}\sum_{j=-\infty}^{\infty}\frac{|\psi_{ij}(X_k)||\psi_{ij}(x)|}{G(Y_k)} \quad a.s.,$$

and the compactness of the support of $\psi(\cdot)$ implies that, for each i, there exists a finite number of non-zero j terms $\psi_{ij}(x)$. So, from

$$\begin{aligned}\theta E\Big(\frac{|\psi_{ij}(X_k)||\psi_{ij}(x)|}{G(Y_k)}\Big) &= 2^i|\phi(2^i x - j)|\int|\psi(2^i t - j)|v(t)dt \\ &= |\phi(2^i x - j)|\int|\psi(u)|v((u+j)/2^i)du = O(1)\end{aligned}$$

and $\pi = O(\ln(n))$, it follows that $\sqrt{n2^{-m}}I_4(x) = O_p(\pi(2^{-m}\ln\ln(n))^{1/2}) = o_p(1)$.

Step 4. We prove $\sqrt{n2^{-m}}I_5(x) = o_p(1)$. Note that

$$\begin{aligned}I_5(x) &= \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}(\tilde{a}_{ij} - a_{ij})I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}a_{ij}I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) \\ &:= I_{51}(x) + I_{52}(x).\end{aligned}$$

Let β_i $(i = 1, 2, 3, 4)$ be positive numbers such that $\beta_1 + \beta_2 = 1$ and $\beta_3 + \beta_4 = 1$, then

$$\begin{aligned}|I_{51}(x)| &\le \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|a_{ij}| > \beta_1\delta) \\ &\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|\tilde{a}_{ij} - a_{ij}| > \beta_2\beta_3\delta) \\ &\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|A_{ij}| > \beta_2\beta_4\delta) \\ &\le \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|a_{ij}| > \beta_1\delta) \\ &\quad + 2\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|\tilde{a}_{ij} - a_{ij}| > \beta_2\beta_3\delta) \\ &\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|\tilde{a}_{ij} - a_{ij}||\psi_{ij}(x)|I(|A_{ij}| > \beta_2\beta_4\delta, |\tilde{a}_{ij} - a_{ij}| \le \beta_2\beta_3\delta) \\ &:= I_{511}(x) + 2I_{512}(x) + I_{513}(x).\end{aligned}$$

Since the r-th derivative $v^{(r)}(\cdot)$ of $v(\cdot)$ is continuous and bounded, by a Taylor expansion, it follows that

$$\begin{aligned} a_{ij} &= 2^{-i/2}\int \psi(u)v\Big(\frac{u+j}{2^i}\Big)du = 2^{-i/2}\int \psi(u)\Big[\sum_{l=0}^{r-1}\frac{1}{l!}(u/2^i)^l v^{(l)}(j/2^i) \\ &\quad +\frac{1}{(r-1)!}(u/2^i)^r\int_0^1(1-t)^{r-1}v^{(r)}((j+tu)/2^i)dt\Big]du \\ &= 2^{-(r+1/2)i}\frac{1}{(r-1)!}\int u^r\psi(u)\Big[\int_0^1(1-t)^{r-1}v^{(r)}((j+tu)/2^i)dt\Big]du, \end{aligned} \tag{4.13}$$

which yields that $|a_{ij}| \le c2^{-(r+1/2)i}$, and $I(|a_{ij}| > \beta_1\delta) = 0$ for large n by $2^{(r+1/2)i}\delta \ge 2^{(r+1/2)m}\delta \to \infty$. Hence, $\sqrt{n2^{-m}}I_{511}(x) = o_p(1)$.

Let a be positive constant such that $a^{-1}+b^{-1} = 1$. Since there exists a finite number of non-zero j terms $\psi_{ij}(x)$, by the Hölder inequality and Lemma 5.9 we have

$$\begin{aligned} EI_{512}^2(x) &\le C\pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}E\{(\tilde a_{ij}-a_{ij})^2\psi_{ij}^2(x)I(|\tilde a_{ij}-a_{ij}| > \beta_2\beta_3\delta)\} \\ &\le C\pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}\psi_{ij}^2(x)E^{1/a}|\tilde a_{ij}-a_{ij}|^{2a}P^{1/b}(|\tilde a_{ij}-a_{ij}| > \beta_2\beta_3\delta) \\ &\le C\pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}\psi_{ij}^2(x)n^{-1}P^{1/b}(|\tilde a_{ij}-a_{ij}| > \beta_2\beta_3\delta). \end{aligned}$$

Next, we evaluate $P(|\tilde a_{ij}-a_{ij}| > \beta_2\beta_3\delta)$. Set $\xi_{ijk} = \frac{\theta}{G(Y_k)}\psi_{ij}(X_k)$. Then $E\xi_{ijk} = a_{ij}$ and $|\xi_{ijk}-E\xi_{ijk}| \le C2^{i/2} := S$, $E(\xi_{ijk}-E\xi_{ijk})^2 \le E\xi_{ijk}^2 \le C$, $|\mathrm{Cov}(\xi_{ijs},\xi_{ijt})| = O(2^{-i})$ for $s \ne t$. Hence, by Lemma 5.5, taking $m=\infty$, for $N \in \mathbb{N}, 0 < N \le n/2$ we have

$$D_N = \max_{1\le l\le 2N}\mathrm{Var}\Big(\sum_{k=1}^{l}\xi_{ijk}\Big) \le CN\Big((2^{i/2})^{2/r}(2^{-i})^{1-1/r}+C\Big) \le CN. \tag{4.14}$$

So, according to Lemma 5.4, taking $N = [(\delta^2 2^i)^{-1/2}]$, it follows that

$$\begin{aligned} &P(|\tilde a_{ij}-a_{ij}| > \beta_2\beta_3\delta) = P\Big(\Big|\sum_{k=1}^{n}(\xi_{ijk}-E\xi_{ijk})\Big| > n\beta_2\beta_3\delta\Big) \\ &\le 4\exp\Big\{-\frac{n^2\beta_2^2\beta_3^2\delta^2/16}{nN^{-1}D_N + Cn\beta_2\beta_3\delta SN}\Big\} + \frac{32S}{n\beta_2\beta_3\delta}n\alpha(N) \\ &\le 4\exp\{-C_4\delta^2 n\} + C(2^{i(\lambda+1)}\delta^{2(\lambda-1)})^{1/2}. \end{aligned} \tag{4.15}$$

Then

$$\begin{aligned} EI_{512}^2(x) &\le C\pi\sum_{i=m}^{\pi}2^i n^{-1}\Big\{\exp\{-C_5\delta^2 n\} + C(2^{i(\lambda+1)}\delta^{2(\lambda-1)})^{1/(2b)}\Big\} \\ &\le C\frac{\pi 2^{\pi}}{n}\exp\{-C_5\delta^2 n\} + C\frac{\pi}{n}2^{\pi[(\lambda+1)/(2b)+1]}\delta^{(\lambda-1)/b} \\ &= o(n^{-2r/(2r+1)}) = o(n^{-1}2^m), \end{aligned}$$

which follows by choosing $\delta \geq C_3(n^{-1}\ln(n))^{1/2}$ with C_3 such that $C_3C_5 = 2r/(2r+1)$, and by noticing that $2^\pi\delta^2 = O(n^{-\epsilon})$, $\delta \geq C_3(n^{-1}\ln(n))^{1/2}$ and $\epsilon(\lambda+1+2b)+2b/(2r+1) \geq 2(b+1)$ imply $n^{-1}\pi 2^{\pi[(\lambda+1)/(2b)+1]}\delta^{(\lambda-1)/b} = o(n^{-2r/(2r+1)})$. Therefore $\sqrt{n2^{-m}}I_{512}(x) = o_p(1)$.

Similarly to the arguments in (4.14), it follows that $\mathrm{Var}(\theta\sum_{k=1}^n |\psi_{ij}(X_k)|G^{-1}(Y_k)) \leq Cn$. Hence

$$
\begin{aligned}
EA_{ij}^2 &= O\Big(\frac{\ln\ln(n)}{n}\Big)E\Big(\frac{\theta}{n}\sum_{k=1}^n \frac{|\psi_{ij}(X_k)|}{G(Y_k)}\Big)^2 \\
&= O\Big(\frac{\ln\ln(n)}{n}\Big)\Big\{\mathrm{Var}\Big(\frac{\theta}{n}\sum_{k=1}^n \frac{|\psi_{ij}(X_k)|}{G(Y_k)}\Big) + \Big(\theta E|\psi_{ij}(X_1)|G^{-1}(Y_1)\Big)^2\Big\} \\
&= O\Big(\frac{\ln\ln(n)}{n}\Big)\Big\{\frac{1}{n}+\frac{1}{2^i}\Big\}. \qquad (4.16)
\end{aligned}
$$

Then

$$
\begin{aligned}
EI_{513}^2(x) &\leq C\pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}\psi_{ij}^2(x)EA_{ij}^2 = O\Big(\frac{\ln\ln(n)}{n}\Big)\pi\sum_{i=m}^{\pi}\Big\{\frac{2^i}{n}+1\Big\} \\
&= O\Big(\frac{\ln\ln(n)}{n}\Big)\Big\{\frac{\pi 2^\pi}{n}+\pi^2\Big\} = o(n^{-1}2^m),
\end{aligned}
$$

which yields $\sqrt{n2^{-m}}I_{513}(x) = o_p(1)$. As to $I_{52}(x)$, we have

$$
\begin{aligned}
|I_{52}(x)| &\leq \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|a_{ij}\psi_{ij}(x)|I(|a_{ij}|>\beta_1\delta) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|a_{ij}\psi_{ij}(x)|I(|\tilde{a}_{ij}-a_{ij}|>\beta_2\beta_3\delta) \\
&\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}|a_{ij}\psi_{ij}(x)|I(|A_{ij}|>\beta_2\beta_4\delta) \\
&:= I_{521}(x)+I_{522}(x)+I_{523}(x).
\end{aligned}
$$

By $I(|a_{ij}|>\beta_1\delta) = 0$ for large n, we have $\sqrt{n2^{-m}}I_{521}(x) = o_p(1)$. Since $2^\pi\delta^2 = O(n^{-\epsilon})$, $\delta \geq C_3(n^{-1}\ln(n))^{1/2}$ and $\lambda > 1+4r/[\epsilon(2r+1)]$ imply $\pi 2^{\pi[(\lambda+1)/2-2r]}\delta^{\lambda-1} = o(n^{-2r/(2r+1)})$, from (4.15) and (4.16), it follows that

$$
\begin{aligned}
EI_{522}^2(x) &\leq \pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}a_{ij}^2\psi_{ij}^2(x)P(|\tilde{a}_{ij}-a_{ij}|>\beta_2\beta_3\delta) \\
&\leq C\pi\sum_{i=m}^{\pi}\Big\{2^{-2ri}\exp(-C_4\delta^2 n)+2^{i[(\lambda+1)/2-2r]}\delta^{\lambda-1}\Big\} \\
&\leq C\pi\Big\{2^{-2rm}\exp(-C_4\delta^2 n)+2^{\pi[(\lambda+1)/2-2r]}\delta^{\lambda-1}\Big\} \\
&= o(n^{-2r/(2r+1)}) = o(n^{-1}2^m), \\
EI_{523}^2(x) &\leq \pi\sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty}a_{ij}^2\psi_{ij}^2(x)\delta^{-2}EA_{ij}^2 = O\Big(\frac{\ln\ln(n)}{n}\Big)\delta^{-2}\pi\sum_{i=m}^{\pi}2^{-2ri}\Big\{\frac{1}{n}+\frac{1}{2^i}\Big\} \\
&= O\Big(\frac{\ln\ln(n)}{n}\Big)\delta^{-2}\pi\Big\{n^{-1}2^{-2rm}+2^{-(2r+1)m}\Big\} = o(n^{-1}2^m).
\end{aligned}
$$

Therefore, $\sqrt{n2^{-m}}I_{522}(x) = o_p(1)$ and $\sqrt{n2^{-m}}I_{523}(x) = o_p(1)$. ■

Proof of Theorem 3.2. We write

$$\begin{aligned}
\hat{h}_n(x) - h(x) &= \sum_{j=-\infty}^{\infty} (\tilde{b}_{mj} - b_{mj})\phi_{mj}(x) + \Big[\sum_{j=-\infty}^{\infty} b_{mj}\phi_{mj}(x) - h(x)\Big] + \sum_{j=-\infty}^{\infty} B_{mj}\phi_{mj}(x) \\
&\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} B_{ij} I(|\hat{b}_{ij}| > \delta)\psi_{ij}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \tilde{b}_{ij} I(|\hat{b}_{ij}| > \delta)\psi_{ij}(x) \\
&:= \Lambda_1(x) + \Lambda_2(x) + \Lambda_3(x) + \Lambda_4(x) + \Lambda_5(x),
\end{aligned}$$

$$\begin{aligned}
\hat{v}_n(x) - v(x) &= \sum_{j=-\infty}^{\infty} (\tilde{a}_{mj} - a_{mj})\phi_{mj}(x) + \Big[\sum_{j=-\infty}^{\infty} a_{mj}\phi_{mj}(x) - v(x)\Big] + \sum_{j=-\infty}^{\infty} A_{mj}\phi_{mj}(x) \\
&\quad + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} A_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) + \sum_{i=m}^{\pi}\sum_{j=-\infty}^{\infty} \tilde{a}_{ij} I(|\hat{a}_{ij}| > \delta)\psi_{ij}(x) \\
&:= I_1(x) + I_2'(x) + I_3(x) + I_4(x) + I_5(x).
\end{aligned}$$

From the proof of Theorem 3.1 it follows that

$$\sqrt{n2^{-m}}I_2'(x_m) = o(1),\ \sqrt{n2^{-m}}I_i(x) = o_p(1)\ (i = 3, 4, 5). \tag{4.17}$$

Since (3.1) implies that Y is bounded, by using arguments similar to those in the proof of Theorem 3.1, one can verify that

$$\sqrt{n2^{-m}}\Lambda_2(x_m) = o(1),\ \sqrt{n2^{-m}}\Lambda_i(x) = o_p(1)\ (i = 3, 4, 5). \tag{4.18}$$

Next we prove

$$\sqrt{n2^{-m}}(\Lambda_1(x_m), I_1(x_m))^{\tau} \xrightarrow{\mathcal{D}} N(0, \theta\Sigma(x)). \tag{4.19}$$

Consider the mapping $L(\cdot)$ from $\mathbb{R}^2$ to $\mathbb{R}$ defined by $L(x, y) = x/y$ for $y \neq 0$. Since $\hat{m}_n(x)$ and $m(x)$ are the respective images of $(\hat{h}_n(x), \hat{v}_n(x))$ and $(h(x), v(x))$ by $L(\cdot)$, we deduce from Lemmas 6.7 to 6.9 and from Mann-Wald's Theorem (see Rao 1965, page 321) that $\sqrt{n2^{-m}}(\hat{m}_n(x) - m(x))$ converges in distribution to $N(0, \theta\nabla L^{\tau}\Sigma(x)\nabla L)$ by (4.17)-(4.19), where the gradient ∇L is evaluated at $(h(x), v(x))$, i.e., $\nabla L = (1/v(x), -h(x)/v^2(x))^{\tau}$. Therefore

$$\begin{aligned}
\Delta^2(x) &= \theta\nabla L^{\tau}\Sigma(x)\nabla L = \theta(1/v(x), -h(x)/v^2(x)) \begin{pmatrix} \Sigma_0(\mathbf{x}) & \Sigma_1(\mathbf{x}) \\ \Sigma_1(\mathbf{x}) & \Sigma_2(\mathbf{x}) \end{pmatrix} \begin{pmatrix} 1/v(x) \\ -h(x)/v^2(x) \end{pmatrix} \\
&= \frac{\theta[\Sigma_0(x)v^2(x) + \Sigma_2(x)h^2(x) - 2v(x)h(x)\Sigma_1(x)]}{v^4(x)}.
\end{aligned}$$

To prove (4.19), it suffices to show that, for any given pair of real numbers $A = (a_1, a_2)^\tau \neq 0$

$$\sqrt{n2^{-m}}(a_1\Lambda_1(x_m) + a_2 I_1(x_m)) \xrightarrow{\mathcal{D}} N(0, \tau^2(\mathbf{x})) \tag{4.20}$$

with $\tau^2(x) = \theta A^\tau \Sigma(x) A = \theta[a_1^2\Sigma_0(x) + a_2^2\Sigma_2(x) + 2a_1a_2\Sigma_1(x)]$. Put

$$\begin{aligned} W_{ni} &= a_1\sqrt{\frac{2^m}{n}}\Big(\frac{\theta Y_i K(2^m X_i, 2^m x_m)}{G(Y_i)} - \int h(t)K(2^m t, 2^m x_m)dt\Big) \\ &\quad + a_2\sqrt{\frac{2^m}{n}}\Big(\frac{\theta K(2^m X_i, 2^m x_m)}{G(Y_i)} - \int v(t)K(2^m t, 2^m x_m)dt\Big). \end{aligned}$$

Then $EW_{ni} = 0$. Let the definitions of w, s_l and t_l be the same as in *Step 1* in the proof of Theorem 3.1. Define $u_{nl} = \sum_{i=s_l}^{s_l+p-1} W_{ni}$, $u'_{nl} = \sum_{i=t_l}^{t_l+q-1} W_{ni}$, $u''_{nw+1} = \sum_{i=w(p+q)+1}^{n} W_{ni}$. Then

$$\sqrt{n2^{-m}}(a_1\Lambda_1(x_m) + a_2 I_1(x_m)) = \sum_{i=1}^{n} W_{ni} = \sum_{l=1}^{w} u_{nl} + \sum_{l=1}^{w} u'_{nl} + u''_{nw+1} := \Omega'_n + \Omega''_n + \Omega'''_n.$$

Hence, for (4.20), we only need to prove that

$$E(\Omega''_n)^2 \to 0, \quad E(\Omega'''_n)^2 \to 0, \tag{4.21}$$

$$\mathrm{Var}(\Omega'_n) \to \tau^2(x), \tag{4.22}$$

$$\Big|E\exp\Big(it\sum_{l=1}^{w} u_{nl}\Big) - \prod_{l=1}^{w} E\exp(itu_{nl})\Big| \to 0, \tag{4.23}$$

$$g_n(\epsilon) = \sum_{l=1}^{w} E\xi_{nl}^2 I(|u_{nl}| > \epsilon\tau(x)) \to 0 \ \ \forall\epsilon > 0. \tag{4.24}$$

We first prove (4.21). We observe that

$$\begin{aligned} E(\Omega''_n)^2 &= \sum_{l=1}^{w}\sum_{i=t_l}^{t_l+q-1} EW_{ni}^2 + 2\sum_{l=1}^{w}\sum_{t_l\le i<j\le t_l+q-1} \mathrm{Cov}(W_{ni}, W_{nj}) + 2\sum_{1\le i<j\le w} \mathrm{Cov}(u'_{ni}, u'_{nj}) \\ &:= J_1(x_m) + J_2(x_m) + J_3(x_m). \end{aligned}$$

Similarly to the arguments in (4.6), we have

$$\begin{aligned}
nEW_{ni}^2 &= a_1^2\Big[2^m E\Big(\frac{\theta^2 Y_i^2 K^2(2^m X_i, 2^m x_m)}{G^2(Y_i)}\Big) - 2^m\Big(\int h(t)K(2^m t, 2^m x_m)dt\Big)^2\Big] \\
&\quad + a_2^2\Big[2^m E\Big(\frac{\theta^2 K^2(2^m X_i, 2^m x_m)}{G^2(Y_i)}\Big) - 2^m\Big(\int v(t)K(2^m t, 2^m x_m)dt\Big)^2\Big] \\
&\quad + 2a_1a_2\Big[2^m E\Big(\frac{\theta^2 Y_i K^2(2^m X_i, 2^m x_m)}{G^2(Y_i)}\Big) \\
&\quad - 2^m\Big(\int h(t)K(2^m t, 2^m x_m)dt\Big)\Big(\int v(t)K(2^m t, 2^m x_m)dt\Big)\Big] \\
&= a_1^2\Big[\theta\int\int\frac{y^2 f(2^{-m}u + x_m, y)}{G(y)}K^2(u,0)dudy - \frac{1}{2^m}\Big(\int h(2^{-m}u + x_m)K(u,0)du\Big)^2\Big] \\
&\quad + a_2^2\Big[\theta\int\int\frac{f(2^{-m}u + x_m, y)}{G(y)}K^2(u,0)dudy - \frac{1}{2^m}\Big(\int v(2^{-m}u + x_m)K(u,0)du\Big)^2\Big] \\
&\quad + 2a_1a_2\Big[\theta\int\int\frac{yf(2^{-m}u + x_m, y)}{G(y)}K^2(u,0)dudy \\
&\quad - \frac{1}{2^m}\Big(\int h(2^{-m}u + x_m)K(u,0)du\Big)\Big(\int v(2^{-m}u + x_m)K(u,0)du\Big)\Big] \\
&\to \theta\Big[a_1^2\int\int\frac{y^2 f(x,y)}{G(y)}K^2(u,0)dudy + a_2^2\int\int\frac{f(x,y)}{G(y)}K^2(u,0)dudy \\
&\quad + 2a_1a_2\int\int\frac{yf(x,y)}{G(y)}K^2(u,0)dudy\Big] = \tau^2(x),
\end{aligned} \tag{4.25}$$

which yields that $J_1(x_m) = O(wq/n) \to 0$.

Similarly to the arguments in *Step 1* in the proof of Theorem 3.1, to prove $J_2(x_m) \to 0$, $J_3(x_m) \to 0$ and $E(\Omega_n''')^2 \to 0$, we only need to show that

$$\sum_{1\le i<j\le n} |\mathrm{Cov}(W_{ni}, W_{nj})| \to 0. \tag{4.26}$$

Note that

$$\sum_{1\le i<j\le n} |\mathrm{Cov}(W_{ni}, W_{nj})| = \sum_{S_1} |\mathrm{Cov}(W_{ni}, W_{nj})| + \sum_{S_2} |\mathrm{Cov}(W_{ni}, W_{nj})|, \tag{4.27}$$

where the definition of S_1 and S_2 is the same as in Theorem 3.1.

From (A1), (2.3) and (3.1), similarly to the arguments in (4.6) or (4.9), for $i < j$ we have

$$\begin{aligned}
& |\mathrm{Cov}(W_{ni}, W_{nj})| \\
= & \Big| E\Big\{ a_1^2 \frac{2^m}{n} \cdot \frac{\theta^2 Y_i Y_j K(2^m X_i, 2^m x_m) K(2^m X_j, 2^m x_m)}{G(Y_i)G(Y_j)} \\
& + a_1 a_2 \frac{2^m}{n} \cdot \frac{\theta^2 Y_i K(2^m X_i, 2^m x_m) K(2^m X_j, 2^m x_m)}{G(Y_i)G(Y_j)} \\
& + a_1 a_2 \frac{2^m}{n} \cdot \frac{\theta^2 Y_j K(2^m X_i, 2^m x_m) K(2^m X_j, 2^m x_m)}{G(Y_i)G(Y_j)} \\
& + a_2^2 \frac{2^m}{n} \cdot \frac{\theta^2 K(2^m X_i, 2^m x_m) K(2^m X_j, 2^m x_m)}{G(Y_i)G(Y_j)} \Big\} \\
& - \Big[E\Big(a_1 \sqrt{\frac{2^m}{n}} \cdot \frac{\theta Y_j K(2^m X_i, 2^m x_m)}{G(Y_i)} + a_2 \sqrt{\frac{2^m}{n}} \cdot \frac{\theta K(2^m X_i, 2^m x_m)}{G(Y_i)} \Big) \Big]^2 \Big| \\
\leq & C \frac{2^m}{n} E|K(2^m X_i, 2^m x_m) K(2^m X_j, 2^m x_m)| \\
& + \Big(a_1 \sqrt{\frac{2^m}{n}} \int\int y f(t,y) K(2^m t, 2^m x_m) dt dy + a_2 \sqrt{\frac{2^m}{n}} \int v(t) K(2^m t, 2^m x_m) dt \Big)^2 \Big\} \\
\leq & \frac{C}{n2^m} \int\int |K(t_1, 0) K(t_2, 0)| dt_1 dt_2 + \Big[\frac{a_1}{\sqrt{n2^m}} \int\int y f(2^{-m} u + x_m, y) K(u, 0) du dy \\
& + \frac{a_2}{\sqrt{n2^m}} \int v(2^{-m} u + x_m) K(u, 0) du \Big]^2 = O((n2^m)^{-1}).
\end{aligned} \tag{4.28}$$

Hence

$$\sum_{S_1} |\mathrm{Cov}(W_{ni}, W_{nj})| = O(c_n 2^{-m}) \to 0. \tag{4.29}$$

On the other hand, it follows from Lemma 5.1 that $|\mathrm{Cov}(W_{ni}, W_{nj})| \leq C[\alpha(j-i)]^{1-2/\gamma}(E|W_{ni}|^\gamma)^{2/\gamma}$ and

$$\begin{aligned}
E|W_{ni}|^\gamma & \leq C\Big(\frac{2^m}{n}\Big)^{\gamma/2} E\Big| \frac{\theta Y_i K(2^m X_i, 2^m x_m)}{G(Y_i)} \Big|^\gamma \leq C\Big(\frac{2^m}{n}\Big)^{\gamma/2} 2^{-m} \int |K(u,0)|^\gamma v(2^{-m} u + x_m) du \\
& = O(n^{-1}(2^m/n)^{\gamma/2-1}).
\end{aligned} \tag{4.30}$$

Similarly to the arguments in *Step 1* of the proof of Theorem 3.1, (4.30) and (A4) imply $\sum_{S_2} |\mathrm{Cov}(W_{ni}, W_{nj})| \to 0$, further (4.26) holds by (4.27) and (4.29).

We now verify (4.22). Note that

$$\mathrm{Var}(\Omega_n') = \sum_{l=1}^{w} \sum_{i=s_l}^{s_l+p-1} EW_{ni}^2 + 2\sum_{l=1}^{w} \sum_{s_l \leq i < j \leq s_l+p-1} \mathrm{Cov}(W_{ni}, W_{nj}) + 2 \sum_{1 \leq i < j \leq w} \sum_{l_1=s_i}^{s_i+p-1} \sum_{l_2=s_j}^{s_j+p-1} \mathrm{Cov}(W_{nl_1}, W_{nl_2}).$$

Then, from $wp/n \to 1$, and (4.25)-(4.26), it follows that $\mathrm{Var}(\Omega_n') \to \tau^2(x)$.

Following the lines in the proof of Theorem 3.1, one can verify (4.23).

By using (4.30), from the proof of $A_n(\epsilon) \to 0$ in Theorem 3.1, $g_n(\epsilon) \to 0$ can be proved similarly. ■

Acknowledgments. This work was supported by the Grants MTM2005-01274 (FEDER support included) of the Spanish Ministerio de Educación y Ciencia, and PGIDIT07PXIB300191PR of the Xunta de Galicia, and also by the National Natural Science Foundation of China (10871146).

References

[1] Antoniadis, A., Grëgoire, G. and Nason, G. (1999). Density and hazard rate estimation for right-censored data by using wavelet methods, *J. R. Statist. Soc.* **B 61**, 63-84.

[2] Bradley, R. C. (1983). Approximation theorems for strongly mixing random variables. *Michigan. Math. J.* **30**, 69-81.

[3] Cai, J. and Kim, J. (2003). Nonparametric quantile estimation with correlated failure time data. *Lifetime Data Analysis* **9**, 357-371.

[4] Cai, Z. W., Roussas, G. G. (1992). Uniform strong estimation under α-mixing, with rate. *Statist. Probab. Lett.* **15**, 47-55.

[5] Daubechies, I. (1992). *Ten Lectures on Wavelets.* SIAM, Philadelphia.

[6] Devroye, L., Györfi, L. and Lugosi, G. (1996). *A probability theory of pattern recognition.* Berlin Heidelberg, New York: Springer.

[7] Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425-455.

[8] Donoho, D. L. and Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *JASA* **90**, 1200-1224.

[9] Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage-asymptopia ? (with discussion) *J. R. Statist. Soc.* **B 57**, 301-369.

[10] Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24**, 508-539.

[11] Doukhan. P. (1994). *Mixing: Properties and Examples.* Lecture Notes in Statistics, vol. 85. Springer, Berlin.

[12] Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and applications.* London: Chapman Hall.

[13] Feigelson, E. D. and Babu, G. J. (1992). *Statistical challenges in modern astronomy.* Berlin Heidelberg, New York: Springer.

[14] Györfi, L., Kohler, M. and Walk, H. (1998). Weak and strong universal consistency semi-recursive partitioning and kernel regression estimates. *Statist. Decisions* **16**, 1-18.

[15] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application.* Academic Press, New York.

[16] Hall, P. and Patil, P. (1995). Formulae for mean integrated squared error of non-linear wavelet-based density estimators. *Ann. Statist.* **23**, 905-928.

[17] Hall, P. and Patil, P. (1996). On the choice of smoothing parameter, threshold and truncation in nonparametric regression by non linear wavelet methods. *J. Roy. Statist. Soc. Ser. B* **58(2)**, 361-377.

[18] Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation and Statistical Applications.* Lecture Notes in Statistics, **129**, Springer, New York.

[19] He, S. and Yang, G. (1994). Estimating a lifetime distribution under different sampling plan. In S.S. Gupta J. O. Berger (Eds.) *Statistical decision theory and related topics* (Vol. 5, 73-85) Berlin Heidelberg, New York: Springer.

[20] He, S., Yang, G. (1998). Estimation of the truncation probability in the random truncation model. *Ann. Statist.* **26**, 1011-1027.

[21] Kerkyacharian, G. and Picard, D. (1992). Density estimation in Besov spaces. *Statist. Probab. Lett.* **13**, 15-24.

[22] Kerkyacharian, G. and Picard, D. (1993). Density estimation by kernel and wavelets methods: optimality of Besov spaces. *Statist. Probab. Lett.* **18(4)**, 327-336.

[23] Kohler, M., Máthé, K. and Pintér, M. (2002). Prediction from randomly right censored data. *J. Multivariate Anal.* **80**, 73-100.

[24] Li, L. (2003). Non-linear wavelet-based density estimators under random censorship. *J. Statist. Plann. Inference* **117(1)** 35-58.

[25] Li, L. and Xiao, Y. (2006). Wavelet-based estimators of mean regression function with long memory data. *Appl. Math. Mech. (English Ed.)* **27(7)**, 901-910.

[26] Li, L. and Xiao, Y. (2007). Mean integrated squared error of non-linear wavelet-based estimators with long memory data. *Ann. Inst. Statist. Math.* **59**, 299-324.

[27] Liang, H. Y., Mammitzsch, V. and Steinebach, J. (2005). Nonlinear wavelet density and hazard rate estimation for censored data under dependent observations. *Statist. Decisions.* **23**, 161-180.

[28] Liang, H. Y., Li, D. L. and Qi, Y. C. (2008). Strong convergence in nonparametric regression with truncated dependent data. *J. Multivariate Anal.* In press.

[29] Liebscher, E. (1996). Strong convergence of sums of α-mixing random variables with applications to density estimation. *Stochastic Processes Appl.*, **65(1)**, 69-80.

[30] Liebscher, E. (2001). Estimation of the density and the regression function under mixing conditions. *Statist. Decisions*, **19(1)**, 9-26.

[31] Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices of the Royal Astronomical Society* **155**, 95-118.

[32] Ould-Saïd, E. and Lemdani, M. (2006) Asymptotic properties of a nonparametric regression function estimator with randomly truncated data. *Ann. Inst. Statist. Math.* **58**, 357-378.

[33] Petrov, V. V. (1995). *Limit Theorems of Probability Theory*. Oxford Univ. Press Inc., New York.

[34] Rao, C. R. (1965). *A linear statistical inference and its applications*. New-York, Wiley.

[35] Rodríguez-Casal, A. and de Uña-Álvarez, J. (2004). Nonlinear wavelet density estimation under the Koziol-Green model. The International Conference on Recent Trends and Directions in Nonparametric Statistics. *J. Nonparametr. Statist.* **16(1-2)**, 91-109.

[36] Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21**, 146-156.

[37] Truoug, Y. and Patil, P. (2001). Asymptotics for wavelet based estimates of piecewise smooth regression for stationary time series. *Ann. Inst. Statist. Math.* **53**, 159-178.

[38] Tsai, W. Y., Jewell, N. P. and Wang, M. C. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika* **74**, 883-886.

[39] Volkonskii, V. A. and Rozanov, Y. A. (1959). Some limit theorems for random functions. *Theory Probab. Appl.* **4**, 178-197.

[40] Wang, M. C., Jemell, N. P. and Tsai, W. Y. (1986). Asymptotic properties of the product-limit estimate under random truncation. *Ann. Statist.* **14**, 1597-1605.

[41] Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13**, 163-177.

[42] Yang, S. C. (2007). Maximal moment inequality for partial sums of strong mixing sequences and application. *Acta. Math. Sinica, English Series* **23**, 1013-1024.

5 Appendix

In this section, we give some preliminary Lemmas, which have been used in Section 4. Let $\{Z_i, i \geq 1\}$ be a sequence of α-mixing real random variables with the mixing coefficients $\{\alpha(k)\}$.

Lemma 5.1 (Hall and Heyde (1980), Corollary A.2, p. 278) *Suppose that X and Y are random variables such that $E|X|^p < \infty, E|X|^q < \infty$, where $p, q > 1, p^{-1} + q^{-1} < 1$. Then*

$$|EXY - EXEY| \leq 8\|X\|_p \|Y\|_q \{ \sup_{A \in \sigma(X), B \in \sigma(Y)} |P(AB) - P(A)P(B)| \}^{1-p^{-1}-q^{-1}}.$$

Lemma 5.2 (Volkonskii and Rozanov, 1959) *Let $V_1, \cdots, V_n$ be α-mixing random variables measurable with respect to the σ-algebra $\mathcal{F}_{i_1}^{j_1}, \cdots, \mathcal{F}_{i_m}^{j_m}$, respectively, with $1 \leq i_1 < j_1 < \cdots < j_m \leq n$, $i_{l+1} - j_l \geq w \geq 1$ and $|V_j| \leq 1$ for $l, j = 1, 2, \cdots, m$. Then*

$$\Big| E\Big(\prod_{j=1}^m V_j \Big) - \prod_{j=1}^m EV_j \Big| \leq 16(m-1)\alpha_w,$$

where $\mathcal{F}_a^b = \sigma\{V_i, a < i \leq b\}$ denotes σ-field generated by $V_{a+1}, V_{a+2}, \cdots, V_b$, α_n is the mixing coefficient.

Lemma 5.3 (Yang (2007), Theorem 2.2) *Let $\lambda > 2$, $\mu > 0$, $EZ_i = 0$ and $E|Z_i|^{\lambda+\mu} < \infty$. Suppose that $\alpha(n) = O(n^{-r})$ for $r > \lambda(\lambda+\mu)/(2\mu)$. Then, for any given $\epsilon > 0$, there exists constant $C = C(r, \mu, \epsilon, \lambda)$ such that $E \max_{1 \leq k \leq n} |\sum_{i=1}^k Z_i|^\lambda \leq C\{n^\epsilon \sum_{i=1}^n E|Z_i|^\lambda + (\sum_{i=1}^n \|Z_i\|_{\lambda+\mu}^2)^{\lambda/2}\}$.*

Lemma 5.4 (Liebscher (2001), Proposition 5.1) *Assume that* $EZ_i = 0$ *and* $|Z_i| \leq S < \infty$ *a.s.* $(i = 1, 2, \cdots, n)$. *Then, for* $n, N \in \mathbb{N}, 0 < N \leq n/2$, $\epsilon > 0$,

$$P\Big(\Big|\sum_{i=1}^{n} Z_i\Big| > \epsilon\Big) \leq 4\exp\Big\{-\frac{\epsilon^2}{16}\Big(nN^{-1}D_N + \frac{1}{3}\epsilon SN\Big)^{-1}\Big\} + 32\frac{S}{\epsilon}n\alpha(N),$$

where $D_N = \max_{1\leq j\leq 2N} \mathrm{Var}(\sum_{i=1}^{j} Z_i)$.

Lemma 5.5 (Liebscher (1996), Lemma 2.3) *Assume* $\alpha(k) \leq C_1 k^{-r}$, *for some* $r > 1, C_1 > 0$. *Let* $\sup_{1\leq i,j\leq n, i\neq j} |\mathrm{Cov}(Z_i, Z_j)| := R^*(n) < \infty$ *be satisfied. Moreover, let* $R_m(n) < \infty$ *for some* m, $2r/(r-1) < m \leq \infty$, *where* $R_m(n) = \sup_{1\leq i\leq n}(E|Z_i|^m)^{1/m}$, *for* $1 \leq m < \infty$, *and* $R_\infty(n) = \sup_{1\leq i\leq n} \operatorname{ess\,sup}_{w\in\Omega}|Z_i|$. *Then*

$$Var\Big(\sum_{i=1}^{n} Z_i\Big) \leq n\Big\{C_2(r, m)(R_m(n))^{2m/(r(m-2))}(R^*(n))^{1-m/(r(m-2))} + R_2^2(n)\Big\}$$

holds with $C_2(r, m) := \frac{20r-40r/m}{r-1-2r/m}C_1^{1/r}$.

Lemma 5.6 (Liang et al. (2008)) *Suppose that* $\alpha(k) = O(k^{-r})$ *for some* $r > 3$. *Then*

$$\sup_y |G_n(y) - G(y)| = O((\ln\ln(n)/n)^{1/2}) \ \ a.s., \quad |\theta_n - \theta| = O((\ln\ln(n)/n)^{1/2}) \ \ a.s..$$

Lemma 5.7 *Let* $\hat{a}_{mj}$, $\hat{a}_{ij}$, $\hat{b}_{mj}$, $\hat{b}_{ij}$ *be as defined in Section 2. Set*

$$\tilde{a}_{mj} = \frac{\theta}{n}\sum_{k=1}^{n}\frac{1}{G(Y_k)}\phi_{mj}(X_k), \quad \tilde{a}_{ij} = \frac{\theta}{n}\sum_{k=1}^{n}\frac{1}{G(Y_k)}\psi_{ij}(X_k),$$
$$\hat{b}_{mj} = \frac{\theta}{n}\sum_{k=1}^{n}\frac{Y_k}{G(Y_k)}\phi_{mj}(X_k), \quad \hat{b}_{ij} = \frac{\theta}{n}\sum_{k=1}^{n}\frac{Y_k}{G(Y_k)}\psi_{ij}(X_k).$$

Then, under the assumption $\alpha(n) = O(n^{-r})$ *for some* $r > 3$, *we have*

$$\hat{a}_{mj} = \tilde{a}_{mj} + A_{mj}, \quad \hat{a}_{ij} = \tilde{a}_{ij} + A_{ij}, \quad \hat{b}_{mj} = \tilde{b}_{mj} + B_{mj}, \quad \hat{b}_{ij} = \tilde{b}_{ij} + B_{ij},$$

where

$$A_{mj} = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big) \cdot \frac{\theta}{n}\sum_{k=1}^{n}\frac{|\phi_{mj}(X_k)|}{G(Y_k)} \quad a.s.,$$
$$A_{ij} = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big) \cdot \frac{\theta}{n}\sum_{k=1}^{n}\frac{|\psi_{ij}(X_k)|}{G(Y_k)} \quad a.s.,$$
$$B_{mj} = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big) \cdot \frac{\theta}{n}\sum_{k=1}^{n}\frac{|Y_k\phi_{mj}(X_k)|}{G(Y_k)} \quad a.s.,$$
$$B_{ij} = O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big) \cdot \frac{\theta}{n}\sum_{k=1}^{n}\frac{|Y_k\psi_{ij}(X_k)|}{G(Y_k)} \quad a.s..$$

Proof. We observe that

$$\begin{aligned}\hat{a}_{mj} &= \frac{\theta}{n}\sum_{k=1}^{n}\frac{1}{G(Y_k)}\phi_{mj}(X_k)+\Big[\frac{\theta_n-\theta}{n}\sum_{k=1}^{n}\frac{1}{G_n(Y_k)}\phi_{mj}(X_k)+\frac{\theta}{n}\sum_{k=1}^{n}\Big(\frac{1}{G_n(Y_k)}-\frac{1}{G(Y_k)}\Big)\phi_{mj}(X_k)\Big]\\ &:= \tilde{a}_{mj}+A_{mj}.\end{aligned}$$

According to (3.1) and Lemma 5.6 we have

$$\begin{aligned}|A_{mj}| &\le \frac{|\theta_n-\theta|}{nG_n(a_F)}\sum_{k=1}^{n}|\phi_{mj}(X_k)|+\frac{\theta\sup_y|G_n(y)-G(y)|}{nG_n(a_F)}\sum_{k=1}^{n}\frac{|\phi_{mj}(X_k)|}{G(Y_k)}\\ &\le \frac{|\theta_n-\theta|}{n[G(a_F)-\sup_y|G_n(y)-G(y)|]}\sum_{k=1}^{n}|\phi_{mj}(X_k)|\\ &\quad+\frac{\theta\sup_y|G_n(y)-G(y)|}{n[G(a_F)-\sup_y|G_n(y)-G(y)|]}\sum_{k=1}^{n}\frac{|\phi_{mj}(X_k)|}{G(Y_k)}\\ &= O\Big(\Big(\frac{\ln\ln(n)}{n}\Big)^{1/2}\Big)\cdot\frac{\theta}{n}\sum_{k=1}^{n}\frac{|\phi_{mj}(X_k)|}{G(Y_k)}\quad a.s..\end{aligned}$$

The others can be verified by the same manner. ■

Lemma 5.8 (Bradley (1983), Theorem 3) *Let η and ξ be real-valued random variables. Suppose U is a uniform-$[0,1]$ random variable, independent of (η,ξ). Then there exists a real-valued random variable ξ^*, measurable w.r.t. (η,ξ,U), such that*

(1) ξ^* is independent of η,

(2) the probability distributions of ξ^* and ξ are identical, and

(3) $P(|\xi^*-\xi|\ge\epsilon)\le 18(\|\xi\|_r/\epsilon)^{r/(2r+1)}\{\sup_{A\in\sigma(\xi),B\in\sigma(\eta)}|P(AB)-P(A)P(B)|\}^{2r/(2r+1)}$, where $0<\epsilon\le\|\xi\|_r$, when $\|\xi\|_r>0$, and $\epsilon>0$, when $\|\xi\|_r=0$ and $\|\xi\|_\infty=\operatorname{ess\,sup}|\xi|$.

Remark 5.1 *Bradley (1983), Theorem 3, considers only the case $\|\xi\|_r>0$. Actually, if $\|\xi\|_r=0$, then $\xi=0$ a.s.; hence, on choosing $\xi^*=\xi=0$ a.s., then, for any $\epsilon>0$,* (1), (2) *and* (3) *in Lemma 5.8 are still true.*

Lemma 5.9 *Let $\tau>2$. Under the assumptions of Theorem 3.1, if $\lambda\ge(\tau-1)(2\tau+1)(2-\epsilon)/(2\epsilon(\tau-2))$, then $E|\tilde{a}_{ij}-a_{ij}|^\tau=O(n^{-\tau/2})$, $E|\tilde{b}_{ij}-b_{ij}|^\tau=O(n^{-\tau/2})$.*

Proof. Following the lines of Lemma 4.5 in Liang et al. (2005), one can verify Lemma 5.9. For the sake of completeness, here we give the proof of the second equation, the proof of the first equation is analogous. Choose $r(n)=[(n2^{-\pi})^{(\tau-2)/(2(\tau-1))}]$, and positive integers $k(n)$ and $\gamma(n)$ such that $n=r(n)k(n)+\gamma(n)$, with $0\le\gamma(n)<r(n)$. Set $W_k=\frac{1}{n}(\frac{\theta Y_k\psi_{ij}(X_k)}{G(Y_k)}-b_{ij})$. Then

$$\tilde{b}_{ij}-b_{ij} \;=\; \sum_{l=1}^{k(n)}\sum_{j=(l-1)r(n)+1}^{lr(n)}W_j+\sum_{j=r(n)k(n)+1}^{n}W_j.$$

The contribution of the remainder term $\sum_{i=r(n)k(n)+1}^{n}W_i$ is negligible (and is subsequently ignored) since it consists of at most $r(n)$ terms. So, without loss of generality, we assume $\gamma(n)=0$, and further $k(n)=2s(n)$. Then

$$\tilde{b}_{ij}-b_{ij}=\sum_{l=1}^{2s(n)}\sum_{j=(l-1)r(n)+1}^{lr(n)}W_j:=\sum_{l=1}^{2s(n)}\xi_n(l)=\sum_{l=1}^{s(n)}\xi_n(2l)+\sum_{l=1}^{s(n)}\xi_n(2l-1):=S(n)+T(n), \quad (5.1)$$

where $\xi_n(l)=\sum_{j=(l-1)r(n)+1}^{lr(n)}W_j$. Hence $E|\tilde{\beta}_{ij}-b_{ij}|^\tau\le C\{E|S(n)|^\tau+E|T(n)|^\tau\}$. Next, we evaluate only $E|T(n)|^\tau$, since the evaluation of $E|S(n)|^\tau$ is similar. In view of Lemma 5.4, there exist i.i.d. random variables $\xi_n^*(2l-1)$, $l=1,2,\dots,s(n)$ such that $\xi_n^*(2l-1)$ has the same distribution as $\xi_n(2l-1)$ for each l, and satisfies

$$P(|\xi_n^*(2l-1)-\xi_n(2l-1)|\ge\epsilon_l)\le 18\Big(\frac{\|\xi_n(2l-1)\|_\infty}{\epsilon_l}\Big)^{1/2}\alpha(r(n)), \quad (5.2)$$

where $0<\epsilon_l\le\|\xi_n(2l-1)\|_\infty$, if $\|\xi_n(2l-1)\|_\infty>0$, and $\epsilon_l>0$, if $\|\xi_n(2l-1)\|_\infty=0$. Then,

$$E|T(n)|^\tau\le C\Big\{E\Big|\sum_{l=1}^{s(n)}\xi_n^*(2l-1)\Big|^\tau+E\Big|\sum_{l=1}^{s(n)}(\xi_n^*(2l-1)-\xi_n(2l-1))\Big|^\tau\Big\}:=C\{T_1(n)+T_2(n)\}.$$

Let us take $M_n>0$ such that $s(n)M_n\asymp n^{-1/2}$, where $a_n\asymp b_n$ means $0<\liminf a_n/b_n\le\limsup a_n/b_n<\infty$, and assume $\|\xi_n(2l-1)\|_\infty\ge M_n$, for $l=1,2,\dots,s(n)$. Otherwise, by rearranging the terms appropriately, we may assume, without loss of generality, that $\|\xi_n(2l-1)\|_\infty\ge M_n$, for $l=1,2,\dots,s_1(n)$, and $\|\xi_n(2l-1)\|_\infty<M_n$, for $l=s_1(n)+1,\dots,s(n)$, where $s_1(n)$ is a positive integer with $s_1(n)\le s(n)$, in this case we have

$$T_2(n)\le C\Big\{(M_ns(n))^\tau+E\Big(\sum_{l=1}^{s_1(n)}|\xi_n^*(2l-1)-\xi_n(2l-1)|\Big)^\tau\Big\}.$$

Therefore,

$$T_2(n)\le C\Big\{(M_ns(n))^\tau+E\Big(\sum_{=1}^{s(n)}|\xi_n^*(2l-1)-\xi_n(2l-1)|I(|\xi_n^*(2l-1)-\xi_n(2l-1))|\ge M_n)\Big)^\tau\Big\},$$

where $\|\xi_n(2l-1)\|_\infty\ge M_n$. Observe that

$$|\xi_n^*(2l-1)-\xi_n(2l-1)|\le 2r(n)\Big(\frac{\theta b_F2^{i/2}\|\psi\|_\infty}{G(a_F)}+|b_{ij}|\Big)\frac{1}{n}\le\frac{C}{n}r(n)2^{\pi/2}.$$

Note that $2^\pi\delta^2 = O(n^{-\epsilon})$, $\delta \geq C_3(n^{-1}\log n)^{1/2}$ and $\lambda \geq (\tau-1)(2\tau+1)(2-\epsilon)/(2\epsilon(\tau-2))$ imply $n^{-(\frac{\lambda(\tau-2)}{2(\tau-1)}-\frac{1}{4})}2^{(\frac{\lambda(\tau-2)}{2(\tau-1)}+\frac{\tau}{2}+\frac{1}{4})\pi} = o(n^{-\tau/2})$. Then, according to (5.2) and $M_n s(n) = O(n^{-1/2})$, it follows that

$$\begin{aligned} T_2(n) &\leq C\Big\{\Big(\frac{1}{n}r(n)2^{\pi/2}\Big)^\tau (s(n))^{\tau-1}\sum_{l=1}^{s(n)} P(|\xi_n^*(2l-1)-\xi_n(2l-1))| \geq M_n)\Big\} + O(n^{-\tau/2}) \\ &\leq C\Big(\frac{1}{n}r(n)2^{\pi/2}\Big)^\tau (s(n))^\tau \Big(\frac{r(n)2^{\pi/2}}{nM_n}\Big)^{1/2}(r(n))^{-\lambda} + O(n^{-\tau/2}) \\ &\leq Cn^{-(\frac{\lambda(\tau-2)}{2(\tau-1)}-\frac{1}{4})}2^{(\frac{\lambda(\tau-2)}{2(\tau-1)}+\frac{\tau}{2}+\frac{1}{4})\pi} + O(n^{-\tau/2}) = O(n^{-\tau/2}). \end{aligned}$$

Next, we estimate $T_1(n)$. Applying the Rosenthal inequality for sums of independent random variables (cf. Petrov (1995), Theorem 2.9, page 59), we get

$$\begin{aligned} T_1(n) &\leq C\Big\{\sum_{l=1}^{s(n)} E|\xi_n^*(2l-1)|^\tau + \Big(\sum_{l=1}^{s(n)} E(\xi_n^*(2l-1))^2\Big)^{\tau/2}\Big\} \\ &\leq C\Big\{s(n)E|\xi_n(1)|^\tau + [s(n)E(\xi_n(1))^2]^{\tau/2}\Big\}. \end{aligned} \tag{5.3}$$

From (3.1) we have

$$\begin{aligned} E|\xi_n(1)|^\tau &= E\Big|\sum_{k=1}^{r(n)} W_k\Big|^\tau \leq (r(n))^\tau E|W_1|^\tau \leq (r(n))^\tau\Big(\frac{\theta b_F}{nG(a_F)}\Big)^\tau E|\psi_{ij}(X_1)|^\tau \\ &\leq C(r(n))^\tau n^{-\tau}\cdot 2^{(\tau/2-1)i}\int |\psi(u)|^\tau v\Big(\frac{u+j}{2^i}\Big)du \leq C(r(n))^\tau n^{-\tau}2^{(\tau/2-1)\pi}. \end{aligned}$$

Then

$$s(n)E|\xi_n(1)|^\tau = O(n^{-\tau/2}). \tag{5.4}$$

As to $E(\xi_n(1))^2$, by using Lemma 5.3, it follows that

$$E(\xi_n(1))^2 = E\Big|\sum_{k=1}^{r(n)} W_k\Big|^2 \leq r(n)\Big\{C(R_\infty(r(n))^{2/\lambda}(R^*(r(n)))^{1-1/\lambda} + R_2^2(r(n))\Big\},$$

where

$$\begin{aligned} &R_\infty(r(n)) := \sup_{1\leq k\leq r(n)} esssup_{w\in\Omega}|W_k| \leq C\Big(\frac{\theta b_F 2^{i/2}\|\psi\|_\infty}{G(a_F)} + |b_{ij}|\Big)\frac{1}{n} = O(2^{\pi/2}n^{-1}); \\ &R_2^2(r(n)) := E|W_1|^2 \leq \frac{C}{n^2}\int \psi^2(u)v\Big(\frac{u+j}{2^i}\Big)du = O(n^{-2}); \\ &R^*(r(n)) := \sup_{1\leq s,t\leq r(n), s\neq t} |\mathrm{Cov}(W_s,W_t)| \leq C(n^2 2^\pi)^{-1}. \end{aligned}$$

Therefore, $E(\xi_n(1))^2 \leq Cr(n)n^{-2}$, and $s(n)E(\xi_n(1))^2 \leq Cs(n)r(n)n^{-2} = O(n^{-1})$, which, together with (5.3) and (5.4), yields $T_1(n) = O(n^{-\tau/2})$. ■