NONLINEAR WAVELET DENSITY ESTIMATION FOR TRUNCATED AND DEPENDENT OBSERVATIONS

JUAN-JUAN CAI* and HAN-YING LIANG†
Department of Mathematics, Tongji University
Shanghai 200092, P. R. China
*caicaiy@gmail.com
†hyliang83@yahoo.com

Received 17 December 2009
Revised 19 March 2010

In this paper, we provide an asymptotic expression for mean integrated squared error (MISE) of nonlinear wavelet density estimator for a truncation model. It is assumed that the lifetime observations form a stationary $\alpha$-mixing sequence. Unlike for kernel estimator, the MISE expression of the nonlinear wavelet estimator is not affected by the presence of discontinuities in the curves. Also, we establish asymptotic normality of the nonlinear wavelet estimator.

Keywords: Nonlinear wavelet density estimator; truncated data; $\alpha$-mixing; mean integrated squared error; asymptotic normality.

AMS Subject Classification: 62G07, 62E20

1. Introduction

In recent years, wavelet methods in nonparametric curve estimation have become a well-known and powerful technique. We refer to the monograph by Härdle et al.\textsuperscript{16} for a systematic discussion of wavelets and their applications in statistics. The major advantage of the wavelet method is its adaptability (in the minimax sense) to the degree of smoothness of the underlying unknown curve. These wavelet estimators typically achieve the optimal convergence rates over exceptionally large function spaces. For more information and related references, see the initial works by Kerkyacharian and Picard,\textsuperscript{18,19} Donoho and Johnstone,\textsuperscript{8,9} and Donoho et al.\textsuperscript{10,11} Hall and Patil\textsuperscript{15} 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

\textsuperscript{†}Corresponding author.
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.

In medical follow-up or in engineering life testing studies, one may not be able to observe the variable of interest, referred to hereafter as the lifetime. Among the different forms in which incomplete data appear, right censoring and left-truncation are two common ones. Some authors have studied wavelet density estimation with censored data. For example, Antoniadis et al. considered linear wavelet density estimation under random censoring, providing the MISE convergence rate under smoothness assumptions on the density function; Li proposed a nonlinear wavelet estimator of the density function with censored data and derived a result similar to the main result, Theorem 2.1 of Ref. 15. Also, Rodríguez-Casal and de Uña-Alvarez investigated the asymptotic expression of the MISE for the nonlinear wavelet estimator of the density function under the Koziol–Green model of random censorship. In addition, Li discussed hazard rate estimation for censored data by wavelet methods. All of the above works are devoted to the independent setting. For dependent case, Liang et al. discussed global $L^2$ error of the nonlinear wavelet estimators of the density function in the Besov space under censoring and stationary $\alpha$-mixing assumptions, Truong and Patil gave an asymptotic expression of the MISE in nonlinear wavelet regression for complete data with $\alpha$-mixing setting. In addition, Tian considered estimations and optimal designs for two-dimensional Haar-wavelet regression models, Chen et al. investigated efficient statistical modeling of wavelet coefficients for image denoising, Kittisuwan et al. discussed image and audio-speech denoising based on higher-order statistical modeling of wavelet coefficients and local variance estimation.

In the left-truncation model, if the lifetime observations in the sample are assumed to be mutually independent, the nonparametric product-limit (PL) estimator of survival function has been studied extensively by many authors during recent years, such as Woodroofe, Chao and Lo, Keiding and Gill, and Stute. Under dependent assumptions, Sun and Zhou provided strong presentation for the kernel estimators of the density and the hazard rate, and gave asymptotic normality and uniform consistency of the kernel density estimators. Motivated by the work of Sun and Zhou and the advantage of the wavelet method, we, in this paper, focus on the nonlinear wavelet estimator of the density function with left-truncated data when the data exhibit some kind of dependence.

Let $\{(X_k, T_k), k \geq 1\}$ be a sequence of random vectors from $(X, T)$. The random variable $X$ with a common unknown continuous distribution function (df) $F$ and density function $f$ can be regarded as the lifetimes of the items under biomedical studies, and $T$ is the truncation variable. We assume throughout that $X$ and $T$ are independent, and $T$ has continuous df $G$. Without loss of generality, we assume that both $X$ and $T$ are non-negative random variables, as usual in survival analysis. In the left-truncation model, $(X_i, T_i)$ are observed only when $X_i \geq T_i$. Let $\gamma = P(T \leq X)$ be the probability that the random variable $X$ is observable. Since $\gamma = 0$ implies that no data can be observed, we suppose throughout the paper that $\gamma > 0$. 

In the sequel, the observed sample \( \{(X_k, T_k), 1 \leq k \leq n\} \) is assumed to be a stationary \( \alpha \)-mixing sequence. Recall that a sequence \( \{\xi_k, k \geq 1\} \) is said to be \( \alpha \)-mixing if the \( \alpha \)-mixing coefficient

\[
\alpha(n) \overset{\text{def}}{=} \sup_{k \geq 1} \sup\{ |P(AB) - P(A)P(B)| : A \in \mathcal{F}^m_{n+k}, B \in \mathcal{F}^n_k \}
\]

converges to zero as \( n \to \infty \), where \( \mathcal{F}^m_l = \sigma\{\xi_i, l \leq i \leq m\} \) denotes the \( \sigma \)-algebra generated by \( \xi_l, \xi_{l+1}, \ldots, \xi_m \) with \( l \leq m \). Among various mixing conditions used in the literature, \( \alpha \)-mixing is reasonably weak and is known to be fulfilled for many stochastic processes including many time series models. Gorodetskii and Withers derived the conditions under which a linear process is \( \alpha \)-mixing. In fact, under very mild assumptions linear autoregressive and more generally bilinear time series models are strongly mixing with mixing coefficients decaying exponentially, i.e. \( \alpha(k) = O(k^p) \) for some \( 0 < p < 1 \). Auestad and Tjøtheim provided illuminating discussions on the role of \( \alpha \)-mixing (including geometric ergodicity) for model identification in nonlinear time series analysis. Chen and Tsay showed that the functional autoregressive process is geometrically ergodic under certain conditions. Also see, e.g., Ref. 12, p. 99, for more details; and see Cai and Kim for motivation in the scope of survival analysis.

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

### 2. Notation and Wavelet-Based Estimator

For any df \( L(y) = P(\eta \leq y) \), define \( L^*(y) = P(\eta \leq y | T \leq X) \). The dfs of \( X \) and \( T \) given no occurrence of the truncation are

\[
F^*(x) = P(X \leq x | T \leq X) = \gamma^{-1} \int_0^x G(u)dF(u) 
\]

and

\[
G^*(x) = P(T \leq x | T \leq X) = \gamma^{-1} \int_0^x (1 - F(u))dG(u),
\]

which can be estimated by

\[
F_n^*(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x) \quad \text{and} \quad G_n^*(x) = n^{-1} \sum_{i=1}^n I(T_i \leq x),
\]

where \( I(\cdot) \) is the indicator function. Since

\[
C(x) = P(T \leq x | T \leq X) = \gamma^{-1} G(x)[1 - F(x)] = G^*(x) - F^*(x),
\]

the empirical estimator of \( C(x) \) is defined by

\[
C_n(x) = n^{-1} \sum_{i=1}^n I(T_i \leq x \leq X_i) = G_n^*(x) - F_n^*(x -),
\]

where \( F_n^*(x -) \) denotes the left-limit of \( F_n^* \) at \( x \). Following the idea of Lynden-Bell the PL estimator \( F_n \) of \( F \) is given by

\[
F_n(x) = 1 - \prod_{X_i \leq x} \left( 1 - \frac{1}{nC_n(X_i)} \right). 
\]

Now we introduce some notations corresponding to wavelets. Let \( \phi(x) \) and \( \psi(x) \) be father and mother wavelets, having the properties: \( \phi(x) \) and \( \psi(x) \) are bounded
590  J.-J. Cai & H.-Y. Liang

and compactly supported, \( \int \phi^2 = \int \psi^2 = 1; \) \( \int y^k \psi(y) dy = 0\) for \( 0 \leq k \leq r - 1, \)
\( r \) is a positive integer and \( \kappa = (r!)^{-1} \int y^r \psi(y) dy \neq 0. \) Let \( \phi_l(x) = p^{l/2} \phi(px - l), \) \( \psi_{kl}(x) = p^{l/2} \psi(pkx - l), \) \( x \in \mathbb{R}, \) for arbitrary \( p > 0, k, l \in \mathbb{Z}, k \geq 0 \) and \( p_k = p^{2^k}. \) Then
\[
\int \phi_{l1} \phi_{l2} = \delta_{l1} \delta_{l2}, \quad \int \psi_{kl1} \psi_{kl2} = \delta_{l1} \delta_{k1} \delta_{l2} \delta_{k2}, \quad \int \phi_{l1} \psi_{kl2} = 0,
\]
where \( \delta_{ij} \) denotes the Kronecker delta (i.e. \( \delta_{ij} = 1, \) if \( i = j; \) 0, otherwise), and the system \( \{ \phi(y), \psi_{kl}(y), k, l \in \mathbb{Z}, k \geq 0 \} \) is an orthonormal basis for the space \( \mathcal{L}_2(\mathbb{R}). \)

For more on wavelets see Refs. 7 and 16.

Let \( a_L = \inf\{ x : L(x) > 0 \} \) and \( b_L = \sup\{ x : L(x) < 1 \}. \) For current model, as discussed by Woodroofe\(^3\) or Sun and Zhou,\(^28\) we assume that \( a_G \leq a_F \) and \( b_G \leq b_F. \) Throughout this paper, let \( a \) and \( b \) be two real numbers such that \( a < b < b_F. \)

In this paper we estimate \( f_1(x) = f(x)I(a \leq x \leq b), \) i.e. the density function \( f(x) \) for \( x \in [a, b]. \) If \( f_1 \in \mathcal{L}_2(\mathbb{R}), \) then we have the following wavelet expansion:
\[
\hat{f}_1(x) = \sum_l b_l \phi(x) + \sum_{k=0}^{\infty} \sum_l b_{kl} \psi_{kl}(x), \quad (2.4)
\]
where \( b_l = \int f_1(x) \phi_l(x) dx \) and \( b_{kl} = \int f_1(x) \psi_{kl}(x) dx \) are wavelet coefficients of function \( f_1(x). \) The proposed nonlinear wavelet estimator of \( f_1(x) \) is
\[
\hat{f}_1(x) = \sum_l \hat{b}_l \phi_l(x) + \sum_{k=0}^{q-1} \sum_l \hat{b}_{kl} I(\hat{b}_{kl} > \delta) \psi_{kl}(x), \quad (2.5)
\]
where \( \delta > 0 \) is a threshold and \( q \geq 1 \) is another smoothing parameter, and the wavelet coefficients \( \hat{b}_k \) and \( \hat{b}_{kl} \) are defined as follows:
\[
\hat{b}_l = \int_{0}^{\infty} \phi_l(x) I(a \leq x \leq b) dF_n(x), \quad \hat{b}_{kl} = \int_{0}^{\infty} \psi_{kl}(x) I(a \leq x \leq b) dF_n(x).
\]

3. Main Results

In the sequel, let \( C, C_0, C_1, \ldots, c \) and \( c_0 \) denote generic finite positive constants, whose values are unimportant and may change from line to line. \( A_n = O(B_n) \) stands for \( |A_n| \leq C|B_n|. \) All limits are taken as the sample size \( n \) tends to \( \infty, \) unless specified otherwise.

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

(A1) \( a(n) = O(p^n) \) for some \( p \in (0, 1). \)

(A2) For all integers \( j \geq 1, \) the joint density \( f_j^x(\cdot, \cdot) \) of the samples \( X_1 \) and \( X_{j+1} \) exists and satisfies \( f_j^x(x_1, x_2) \leq C \) for \( x_1, x_2 \in \mathbb{R}. \)

(A3) The smoothing parameters \( p, q \) and \( \delta \) are functions of \( n. \) Suppose that \( p \to \infty, \) \( q \to \infty \) in such a manner that \( \delta \geq C_0 (n^{-1} \log n)^{1/2}, \) \( p_q \delta^2 = O(n^{-\mu}) \) for some sufficiently small constant \( \mu > 0 \) and \( p^{2^{r_a+1}} \delta^2 \to \infty. \)
Remark 3.1. The conditions (A1)–(A3) were assumed in Ref. 31 for complete data. In i.i.d. setting, the assumption (A3) except $p_\phi \delta^2 = O(n^{-\mu})$ for some sufficiently small constant $\mu > 0$ had been used by some authors, such as Hall and Patil, Li and Rodríguez-Casal and de Uña-Alvarez.

3.1. Mean integrated squared error

Theorem 3.1. Suppose that the conditions on $\phi$ and $\psi$ stated in Sec. 2 and assumptions (A1)–(A3) hold. If $f(x)$ is bounded and has bounded and continuous $r$th derivative. Then

$$
E \left[ (\hat{f}_1 - f_1)^2 \right] - \left\{ n^{-1} p \gamma \int f_1 G^{-1} + p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int f_1^{(r)} \right\}^2 = o(n^{-1} p + p^{-2r}).
$$

In Theorem 3.1, we assumed that the density function $f$ is $r$th continuously differentiable. However, if $f$ is only piecewise smooth, Theorem 3.1 still holds. That is the following theorem.

Theorem 3.2. In Theorem 3.1, suppose that the $r$th derivative $f^{(r)}$ is only piecewise smooth, i.e. there exist points $x_0 = -\infty < x_1 < \cdots < x_\tau < +\infty = x_{\tau+1}$ such that the first $r$ derivatives of $f$ exist and are bounded and continuous on $(x_i, x_{i+1})$ for $1 \leq i \leq \tau$, with left- and right-limits. In particular $f$ itself may be only piecewise continuous. Also assume that $p_\phi^{2r+1} n^{-2r} \to \infty$. Then the conclusion in Theorem 3.1 remains true.

Remark 3.2. (i) The classical MISE formula in the context of linear, kernel-type estimators may be recognized into the following form $\text{MISE} \sim C_1 (nh_n)^{-1} + C_2 h_n^2$, where $n$ denotes sample size, $h_n$ is the bandwidth of the kernel estimator with complete data, $r$ is the order of the kernel, $C_1$ and $C_2$ are constants depending on both kernel and the unknown density, and "$\sim$" means that the ratio of the left- and right-hand sides converges to 1 as $n \to \infty$. The first term derives from variance, the next from squared bias. Compared with the kernel estimators, the wavelet analogue of the bandwidth $h_n$ of the kernel estimators is $p^{-1}$. As pointed out by Hall and Patil the variance component of the integrated squared error is of size $n^{-1} p$ and the squared bias component is of $p^{-2r}$, the optimal size of $p$ is $c_n^{1/(2r+1)}$.

(ii) By choosing $p \sim n^{1/(2r+1)}$ it can be shown that the MISE satisfies

$$
E \left[ (\hat{f}_1 - f_1)^2 \right] \sim n^{-1} p \gamma \int f_1 G^{-1} + p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int f_1^{(r)} \sim a n^{-2r/(2r+1)}
$$

with $a = \gamma \int f_1 G^{-1} + \kappa^2 (1 - 2^{-2r})^{-1} \int f_1^{(r)}$. 

Nonlinear Wavelet Density Estimation
3.2. Asymptotic normality

Theorem 3.3. Suppose that the assumptions in Theorem 3.1 are satisfied, and that there exists an integer $m$ such that $p = 2^m = O(n^{1/(2r+1)})$. Then

$$\sqrt{np^{-1}(f_1(t_m) - f_1(t_m) - b(t_m))} \rightarrow N(0, \sigma^2(t)), \quad t \in \mathbb{R},$$

where $t_m = [2^m t]/2^m$, $b(t) = f^{(r)}(t)/r! \int u^r \sum l \phi(u + l) \phi(l) du$ and $\sigma^2(t) = \gamma G^{-1}(t)f_1(t)\int(\sum l \phi(u + l) \phi(l))^2 du$.

4. Proof of Main Results

Proof of Theorem 3.1. From the orthogonality of the wavelet basis functions, it follows that

$$\int (f_1 - f_1)^2 = \sum l (b_l - b_l)^2 + \sum_{k=0}^{q-1} \sum l (b_{kl} - b_{kl})^2 I(|b_{kl}| > \delta)$$

$$+ \sum_{k=0}^{q-1} \sum l b_{kl}^2 I(|b_{kl}| \leq \delta) + \sum_{k=q}^\infty \sum l b_{kl}^2$$

$$:= S_1 + S_2 + S_3 + S_4.$$

It suffices to show that

$$E \left| S_1 - n^{-1}p \gamma \int f_1 G^{-1} \right| = o(n^{-1}p), \quad ES_2 = o(n^{-2r/(2r+1)}), \quad (4.1)$$

$$E \left| S_3 - p^{-2r}k(1 - 2^{-2r})^{-1} \int f_1^{(r)} \right| = o(p^{-2r}), \quad S_4 = o(p^{-2r}). \quad (4.2)$$

To prove (4.1) and (4.2), we divide the proof into the following four lemmas.

In the following proof, set $\varphi_l(x) = \phi_l(x)I(a \leq x \leq b)$ and $\varphi_{kl}(x) = \psi_{kl}(x)I(a \leq x \leq b)$. We assume, without loss of generality, that both $\phi$ and $\psi$ are compactly supported on $[-v, v]$ ($v > 0$), which implies that the numbers of non-zero terms (in $l$) in $S_1$ and $S_2$ (or $S_3$) are of order $O(p)$ and $O(p_k)$, respectively. Note that condition (A3) implies that $p_k = o(n)$, $q = O(\log n)$ and $p^2r n^{-\mu} \rightarrow \infty$.

Lemma 4.1. We verify that $E|S_1 - n^{-1}p \gamma \int f_1 G^{-1}| = o(n^{-1}p)$.

Proof. From Lemma A.4 we write

$$S_1 = \sum l (b_l - b_l)^2 + \sum B_l^2 + \sum \beta_l^2 + 2 \sum l \{B_l(b_l - b_l) + \beta_l(b_l - b_l) + B_l b_l\}$$

$$:= S_{11} + S_{12} + S_{13} + 2S_{14}.$$
It suffices to show that $E|S_{11} - n^{-1}p\gamma \int f_1 G^{-1}| = o(n^{-1}p)$, $ES_{12} = o(n^{-1}p)$, $ES_{13} = o(n^{-1}p)$ and $E|S_{14}| = o(n^{-1}p)$.

Note that $E|S_{11} - n^{-1}p\gamma \int f_1 G^{-1}| \leq |ES_{11} - n^{-1}p\gamma \int f_1 G^{-1}| + \sqrt{\text{Var}(S_{11})}$. So, in view of Lemma A.5, to prove $E|S_{11} - n^{-1}p\gamma \int f_1 G^{-1}| = o(n^{-1}p)$, it is sufficient to verify that

$$ES_{11} = n^{-1}p\gamma \int f_1 G^{-1} + o(n^{-1}p).$$  

(4.3)

Set $Y_i = \gamma G^{-1}(X_i)\varphi_l(X_i)$. Observe

$$nE(\tilde{b}_i - b_i)^2 = nE\left\{\frac{1}{n} \sum_{i=1}^{n} (Y_i - b_i)\right\}^2 = \text{Var}(Y_{11}) + 2 \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) \text{Cov}(Y_{11}, Y_{i+1}).$$  

(4.4)

From (2.1) one can obtain that

$$\text{Var}(Y_{11}) = EY_{11}^2 - (EY_{11})^2 = \gamma^2 \int G^{-2}(x)\varphi_l^2(x)dF^*(x)$$

$$- \left(\gamma \int G^{-1}(x)\varphi_l(x)dF^*(x)\right)^2$$

$$= \gamma \int \phi^2(x)G^{-1}((x + l)/p)f_1((x + l)/p)dx$$

$$- p^{-1} \left(\int \phi(x)f_1((x + l)/p)dx\right)^2.$$  

(4.5)

Therefore

$$\sum_i \text{Var}(Y_{11}) = p\gamma \left(\int \phi^2(x)dx \int G^{-1}(y)f_1(y)dy\right)$$

$$+ O(1) = p\gamma \int G^{-1}(y)f_1(y)dy + O(1).$$  

(4.6)

In addition, $|p^{\frac{1}{2}} Y_{11}| = |\phi(pX_1 - l)(a \leq X_1 < b)G^{-1}(X_1)\gamma| \leq C$, then according to Lemma A.1, we have $|\text{Cov}(Y_{11}, Y_{i+1})| \leq 4C^2\rho\alpha(i)$. On the other hand, from condition (A2) we have

$$|\text{Cov}(Y_{11}, Y_{i+1})| \leq |E(Y_{11} Y_{i+1})| + (EY_{11})^2$$

$$\leq Cp^{-1} \int \phi(x_1)\phi(x_2)f_1^*(x_1 + l)/p, (x_2 + l)/p)dx_1dx_2$$

$$+ O(p^{-1}) = O(p^{-1}).$$
Choose $M_n = [p/\log n]$, then it follows that
\[
\sum_{i=1}^{n-1} |\text{Cov}(Y_{i+1}, Y_{i+i})| \leq C \sum_{i=1}^{n-1} \min\{p^{-1}, p\alpha(i)\}
\]
\[
= C \left( \sum_{i<M_n} + \sum_{i \geq M_n} \right) \min\{p^{-1}, p\alpha(i)\} = o(1). \tag{4.7}
\]

Therefore, (4.3) follows from (4.4), (4.6) and (4.7). Obviously, (4.4), (4.5) and (4.7) imply that $E(\hat{b} - b)^2 = O(n^{-1})$. Similarly, we prove that

\[
EB_2^2 = O\left( \frac{\ln \ln n}{n} \right) E\left( \frac{1}{n} \sum_{i=1}^{n} |Y_{i+1}| \right)^2
\]
\[
= O\left( \frac{\ln \ln n}{n} \right) \left\{ E\left( \frac{1}{n} \sum_{i=1}^{n} (|Y_{i+1}| - E|Y_{i+1}|)^2 + (E|Y_{i+1}|)^2 \right) \right\} = O\left( \frac{\ln \ln n}{n^p} \right).
\]

Since condition (A3) implies that $\ln \ln n/p \to 0$, $ES_2 = \sum_i EB_i^2 = O(p\ln \ln n/n^p) = o(n^{-1})$. Lemma A.4 shows that $ES_2 = \sum_i E\beta_i^2 = O(p\ln \ln n/n^p) = o(n^{-1})$. By applying Cauchy–Schwarz inequality, one can easily get $E|S_{14}| = o(n^{-1})$.

**Lemma 4.2.** We prove that $ES_2 = o(n^{-2r/(2r+1)})$.

**Proof.** According to Lemma A.4, we have

\[
ES_2 = \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} E\{\hat{b}_{k+1} + B_{k+1} + \beta_{k+1} - b_{k+1}\}^2 I(|\hat{b}_{k+1}| > \delta)
\]
\[
\leq 2 \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} E\beta_{k+1}^2 + 2 \sum_{k=0}^{q-1} E\{\hat{b}_{k+1} + B_{k+1} - b_{k+1}\}^2 I(|\hat{b}_{k+1}| > \delta).
\]

Condition (A3) implies that $q = O(\ln n)$ and $\ln n \ln n/p \to 0$, it can be easily obtained that

\[
\sum_{q=1}^{q-1} \sum_{k=0}^{q-1} E\beta_{k+1}^2 = \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} O(p^{1/2}) = O(q^{1/2}) = o(n^{-2r/(2r+1)}).
\]

Put $S_{20} = \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} (\hat{b}_{k+1} + B_{k+1} - b_{k+1})^2 I(|\hat{b}_{k+1}| > \delta)$. Next, we prove $ES_2 = o(n^{-2r/(2r+1)})$. Let $a_i$ and $b_i$ be positive numbers such that $a_i + b_i = 1$, $i = 1, 2$. By Lemma A.4, we find

\[
S_{20} \leq \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} \{\hat{b}_{k+1} + B_{k+1} - b_{k+1}\}^2 I(|\beta_{k+1} + b_{k+1}| > a_1 \delta)
\]
\[
+ \sum_{q=1}^{q-1} \sum_{k=0}^{q-1} \{\hat{b}_{k+1} + B_{k+1} - b_{k+1}\}^2 I(|\hat{b}_{k+1} - b_{k+1} + B_{k+1}| > b_1 \delta)
\]
\[
:= S_{21} + S_{22},
\]
\[
S_{21} \leq 2 \sum_{k=0}^{q-1} \sum_{l} \{(\hat{b}_{kl} - b_{kl})^2 \} I(|\beta_{kl} + b_{kl}| > a_1 \delta) \\
+ 2 \sum_{k=0}^{q-1} \sum_{l} B_{kl}^2 I(|\beta_{kl} + b_{kl}| > a_1 \delta) \\
:= 2\{S_{21,1} + S_{21,2}\},
\]

\[
S_{22} \leq 2 \sum_{k=0}^{q-1} \sum_{l} \{(\hat{b}_{kl} - b_{kl})^2 \} I(|\hat{b}_{kl} - b_{kl} + B_{kl}| > b_1 \delta) \\
+ 2 \sum_{k=0}^{q-1} \sum_{l} B_{kl}^2 I(|\hat{b}_{kl} - b_{kl} + B_{kl}| > b_1 \delta) \\
:= 2\{S_{22,1} + S_{22,2}\},
\]

\[
S_{22,1} \leq \sum_{k=0}^{q-1} \sum_{l} \{(\hat{b}_{kl} - b_{kl})^2 \} I(|\hat{b}_{kl} - b_{kl}| > a_2 b_1 \delta) \\
+ \sum_{k=0}^{q-1} \sum_{l} \{(\hat{b}_{kl} - b_{kl})^2 \} I(|B_{kl}| > b_2 b_1 \delta) \\
:= S_{22,11} + S_{22,12}.
\]

Hence \(S_{20} \leq 2\{S_{21,1} + S_{21,2} + S_{22,11} + S_{22,12} + S_{22,2}\}.

By using Taylor expansion and \(\int \psi(y)y' \, dy = 0\), for all \(0 \leq i \leq r-1\), we have

\[
p_{kl}^{1/2} b_{kl} = \int \psi(y)f_1 \left(\frac{y + 1}{p_k}\right) dy \\
= \int \psi(y) \left\{\sum_{i=0}^{r-1} \frac{1}{i!} \left(\frac{y}{p_k}\right)^i f_1^{(i)} \left(\frac{l}{p_k}\right) + \frac{1}{(r-1)!} \left(\frac{y}{p_k}\right)^r\right\} dy \\
\times \int_0^1 (1-t)^{r-1} f_1^{(r)} \left(\frac{l + ty}{p_k}\right) dt \right\} dy \\
= \int \psi(y)y' p_k^{-r} dy \left\{\frac{1}{(r-1)!} (f_{kl} + \xi_{kl}) \int_0^1 (1-t)^{r-1} dt \right\} = \kappa p_k^{-r} (f_{kl} + \xi_{kl}),
\]

(4.8)

where \(f_{kl} = f_1^{(r)} (l/p_k)\) and \(\sup_{0 \leq k \leq q-1, l} |\xi_{kl}| \to 0\). Then \(|\hat{b}_{kl}| = C_1 p_k^{-(r+1/2)}\).

Note that (A3) implies that

\[
\sup_l |\beta_{kl}| \delta^{-1} = O(\gamma_n) \delta^{-1} p_k^{-1/2} \int |(y)| f_1 \left(\frac{y + 1}{p_k}\right) dy = O(\gamma_n \delta^{-1} p_k^{-1/2}) \to 0
\]

and \(\sup_l |b_{kl}| \delta^{-1} \leq C_1 p_k^{-(r+1/2)} \delta^{-1} \to 0\), so \(\sup_l |\hat{b}_{kl} + \beta_{kl}| \delta^{-1} \to 0\). Then for all sufficiently large \(n\), \(I(|\hat{b}_{kl} + \beta_{kl}| > a_1 \delta) = 0\). Thus \(ES_{21,1} = o(n^{-2r/(2r+1)})\) and \(ES_{21,2} = o(n^{-2r/(2r+1)})\).
Let $\tau_1$ and $\tau_2$ be positive numbers such that $\frac{2}{\tau_1} + \frac{1}{\tau_2} = 1$. Then $(E(\tilde{b}_{kl} - b_{kl})^{\tau_1})^{\frac{1}{\tau_1}} = O(n^{-1})$ and $\{P(\tilde{b}_{kl} - b_{kl} > a_2 b_1 \delta)\}^{\frac{1}{\tau_2}} = o(n^{-(2r)/(2r+1)})$ can be obtained by Lemmas A.6 and A.7. As a result,

$$ES_{22.11} \leq \sum_{k=0}^{q-1} \sum_t E(\tilde{b}_{kl} - b_{kl})^{\tau_1} P(\tilde{b}_{kl} - b_{kl} > a_2 b_1 \delta)$$

Let $A = \{b_{kl} - b_{kl} \leq a_2 b_1 \delta\}$. Similarly to the arguments as for $EB_{kl}^2 = O((\ln \ln n)/(np_k))$ in Lemma 4.1, one can verify that $EB_{kl}^2 = O((\ln \ln n)/(np_k))$. Hence we have

$$ES_{22.12} = \sum_{k=0}^{q-1} \sum_t E(\tilde{b}_{kl} - b_{kl})^2 I(\{\tilde{b}_{kl} > b_2 b_1 \delta\}) I_A$$

$$+ \sum_{k=0}^{q-1} \sum_t E(\tilde{b}_{kl} - b_{kl})^2 I(\{\tilde{b}_{kl} > b_2 b_1 \delta\}) I_{A^c}$$

$$\leq \sum_{k=0}^{q-1} \sum_t a_2^2 b_1^2 \delta^2 P(\tilde{b}_{kl} > b_2 b_1 \delta) + \sum_{k=0}^{q-1} \sum_t E(\tilde{b}_{kl} - b_{kl})^2 I_{A^c}$$

$$\leq \sum_{k=0}^{q-1} \sum_t a_2^2 b_1^2 \delta^2 \frac{EB_{kl}^2}{b_2^2 b_1^2 \delta^2} + S_{22.11} = o(n^{-(2r)/(2r+1)})$$

and $ES_{22.2} \leq \sum_{k=0}^{q-1} \sum_t EB_{kl}^2 = O(n^{-1}(\ln \ln n) = o(n^{-(2r)/(2r+1)}))$. Thus the proof of Lemma 4.2 is completed.

**Lemma 4.3.** We prove that $E|S_3| = p^{-2r} \rho^2 (1 - 2^{-2r})^{-1} \int f_{1/2}^{(r)} | = o(p^{-2r})$.

**Proof.** Let $\varepsilon > 0$ and define $S_{31} = \sum_{k=0}^{q-1} \sum_t b_{kl}^2 I(|b_{kl} + \beta_{kl}| \leq (1 + \varepsilon) \delta)$,

$$S_{32} = \sum_{k=0}^{q-1} \sum_t b_{kl}^2 I(|b_{kl} + \beta_{kl}| \leq (1 - \varepsilon) \delta)$$

and

$$\Delta_1 = \sum_{k=0}^{q-1} \sum_t b_{kl}^2 I(|\tilde{b}_{kl} - b_{kl} + B_{kl}| > \varepsilon \delta).$$

Then $S_{32} - \Delta_1 \leq S_3 \leq S_{31} + \Delta_1$.

The proof in Lemma 4.2 shows that $I(|b_{kl} + \beta_{kl}| \leq \delta) = 1$ for sufficient large $n$. Then, from (4.8) we have

$$S_{31} = S_{32} = \sum_{k=0}^{q-1} \sum_t b_{kl}^2 = \kappa^2 \sum_{k=0}^{q-1} p_k^{-2r} \sum_t p_k^{-1} (f_{1/2}^{(r)}(l/p_k) + \xi_{kl})^2$$

$$= p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int f_{1/2}^{(r)} + o(p^{-2r}).$$
Next we need only to verify that $E \Delta_4 = o(S_{31})$. Let $a_3$ and $b_3$ be positive numbers such that $a_3 + b_3 = 1$. Clearly

$$\Delta_1 \leq \sum_{k=0}^{q-1} \sum_l b_{kl}^2 I(|b_{kl} - b_k| > a_3 \epsilon \delta) + \sum_{k=0}^{q-1} \sum_l b_{kl}^2 I(|B_{kl}| > b_3 \epsilon \delta) := \Delta_{11} + \Delta_{12}.$$  

According to Lemma A.7 and $EB_{kl}^2 = O(\ln \ln n/p_k)$, we have $E \Delta_{11} = \sum_{k=0}^{q-1} \sum_l b_{kl}^2 P(|b_{kl} - b_k| > a_3 \epsilon \delta) = o(S_{31})$ and $E \Delta_{12} = \sum_{k=0}^{q-1} \sum_l b_{kl}^2 P(|B_{kl}| > b_3 \epsilon \delta) \leq \sum_{k=0}^{q-1} \sum_l b_{kl}^2 EB_{kl}^2 = o(S_{31})$. This proves Lemma 4.3.

**Lemma 4.4.** We prove that $S_4 = o(p^{-2r})$.

**Proof.** In view of (4.8) and $q \rightarrow \infty$,

$$S_4 = \sum_{k=q}^{\infty} \sum_l \kappa^2 p_k^{-2(r+1)} (f_l + \xi_{kl})^2 \leq 2 \sum_{k=q}^{\infty} \kappa^2 p_k^{-2r} \sum_{l} p_k^{-1} (f_l^{(r)} (l/p_k))^2 = o(p^{-2r}) = o(p^{-2r}).$$

**Proof of Theorem 3.2.** It suffices to show that Lemmas 4.1–4.4 remain true in the context of Theorem 3.2. Lemma 4.1 holds since the piecewise smoothness does not affect the integrality of $f_1$.

As for Lemmas 4.2 and 4.3, define $\mathcal{H} = \{x | x$ is a discontinuity point of $f^{(s)}, 0 \leq s \leq r \}$. Unless $l \in R_k = \{l : l \in (p_k x - v, p_k x + v), x \in \mathcal{H}\}$, both $b_{kl}$ and $\hat{b}_{kl}$, constructed entirely from an interval where $f^{(r)}$ exists and is bounded, have precisely properties claimed in Theorem 3.1. Then the methods in the proof of Theorem 3.1 can be employed to obtain that

$$E \left\{ \sum_{k=0}^{q} \sum_{l \in R_k} (b_{kl} - \hat{b}_{kl})^2 I(|b_{kl}| > \delta) \right\} = o(n^{-2r/(2r+1)})$$

and $E \{ \sum_{k=0}^{q} \sum_{l \in R_k} b_{kl}^2 I(|b_{kl}| \leq \delta) \} = p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int f^{(r)}^2$ where $R_k$ denotes the complement of $R_k$. Since $\# \mathcal{H}$ is finite set of points and $R_k$ has no more than $(2v+1)(\# \mathcal{H})$ elements for each $k$, according to $E(b_{kl} - \hat{b}_{kl})^2 = O(n^{-1})$, $EB_{kl}^2 = O(\ln \ln n/(np_k))$ and $|\beta_{kl}| = O(\gamma n p_k^{-1/2})$ we have

$$\sum_{k=0}^{q} \sum_{l \in R_k} E(b_{kl} - \hat{b}_{kl})^2 \leq \sum_{k=0}^{q} \sum_{l \in R_k} (E(b_{kl} - \hat{b}_{kl})^2 + EB_{kl}^2 + \beta_{kl}^2)$$

$$= O \left( \frac{q}{n} + \frac{2 \ln \ln n}{np} \right) = o(n^{-2r/(2r+1)}).$$

Note that $\sup_{k,l} |\beta_{kl}| \delta^{-1} \rightarrow 0$ and $E |B_{kl}| \delta^{-1} \rightarrow 0$, and similarly to the proof as in Lemma A.7 one can prove that $P(|\hat{b}_{kl} - b_{kl} + B_{kl}| > \epsilon \delta) = o(n^{-2r/(2r+1)})$ for some
\( \epsilon > 0 \). Therefore

\[
E \left\{ \sum_{k=0}^{q} \sum_{t \in R_k} b_{kt} I(|\tilde{b}_{kt}| \leq \delta) \right\}
\]

\[
\leq \sum_{k=0}^{q} \sum_{t \in R_k} b_{kt}^2 I(|\tilde{b}_{kt} + \beta_{kt}| \leq (1 + \epsilon)\delta) + E \left\{ \sum_{k=0}^{q} \sum_{t \in R_k} b_{kt}^2 I(|\tilde{b}_{kt} - b_{kt} + B_{kt}| > \epsilon \delta) \right\}
\]

\[
\leq \sum_{k=0}^{q} \sum_{t \in R_k} b_{kt}^2 I(|\tilde{b}_{kt}| \leq |\beta_{kt}| + (1 + \epsilon)\delta) + \sum_{k=0}^{q} O(p_k^{-1}) P(|\tilde{b}_{kt} - b_{kt} + B_{kt}| > \epsilon \delta)
\]

\[
= O(\delta^2) + o(n^{-2r/(2r+1)}) = o(n^{-2r/(2r+1)})
\]

by \( p_q q^2 = O(n^{-\mu}) \) and \( p_q^{2r+1} n^{-2r} \to +\infty \). As to Lemma 4.4, in view of \(|\tilde{b}_{kt}| = O(p_k^{-1/2}) \) and \( p_q^{2r+1} n^{-2r} \to +\infty \) we have \( \sum_{k=0}^{q} \sum_{t \in R_k} b_{kt}^2 = O\left( \sum_{k=0}^{q} O(p_k^{-1}) \right) = o(n^{-2r/(2r+1)}) \). \( \square \)

**Proof of Theorem 3.3.** Set \( K(t, x) = \sum_{j=-\infty}^{\infty} \phi(t - j) \phi(x - j) \). Then \( K(t, x) \) has the following properties:

(P1) \( K(t, x) \) is uniformly bounded;

(P2) \( K(t, x) = 0 \) for \( |t - x| > 4v \) if support \( \phi(x) \subset [-v, v] \);

(P3) \( K(t + j, x + j) = K(t, x) \) for any integer \( j \);

(P4) \( K(t, x) \) satisfies the moment condition (cf. Ref. 16, Theorem 8.3, p. 93), i.e.

\[
\int (t - x)^k K(t, x) dt = \delta_{0k} \text{ for } k = 0, 1, \ldots, r - 1.
\]

We write

\[
f_1(t_m) - f_1(t_m) - b(t_m) = \sum_{l} (\tilde{b}_l - b_l) \phi_l(t_m) + \sum_{k=0}^{q-1} b_{kt} \psi_{kt}(t_m) I(|\tilde{b}_{kt}| > \delta)
\]

\[
+ \left( \sum_{l} b_l \phi_l(t_m) - f_1(t_m) - b(t_m) \right) := J_1 + J_2 + J_3.
\]

It suffices to show that \( \sqrt{np^{-1}} J_1 \to_{\mathbb{P}} \mathcal{N}(0, \sigma^2(t)) \), \( EJ_3^2 = o(n^{-1}p) \), \( J_3 = o(p^{-r}) = o(\sqrt{n^{-1}p}) \). The proofs are divided into the following three lemmas.

**Lemma 4.5.** We prove \( \sqrt{np^{-1}} J_1 \to_{\mathbb{P}} \mathcal{N}(0, \sigma^2(t)) \).

**Proof.** In view of Lemma A.4, we have

\[
\sum_{l} (\tilde{b}_l - b_l) \phi_l(t_m) = \sum_{l} (\tilde{b}_l - b_l) \phi_l(t_m) + \sum_{l} B_l \phi_l(t_m) + \sum_{l} \beta_l \phi_l(t_m).
\]

Since the support of \( \phi \) is compact, there exists only finite number of non-zero \( l \) terms \( \phi_l(t_m) = p^{1/2} \phi(p^l t_m - l) \) for each fixed \( t_m \). Hence, from (2.1) and \( \beta_l = O(\gamma_n p^{-1/2}) \)
we have

\[
E \left( \sum_i B_i \phi_i(t_m) \right)^2 = O \left( \frac{\ln \ln n}{np} \right) = o(n^{-p}),
\]

\[
E \left( \sum_i \beta_i \phi_i(t_m) \right)^2 = O(\gamma n^2 p^{-1}) = o(n^{-1}p).
\]

Therefore, we need only to prove that \( \sqrt{np^{-1}} \sum_i (\hat{b}_i - b_i) \phi_i(t_m) \rightarrow_d N(0, \sigma^2(t)) \).

We observe that

\[
\sqrt{np^{-1}} \sum_i (\hat{b}_i - b_i) \phi_i(t_m) = \frac{1}{\sqrt{n}} \sum_{i=1}^n p^{1/2} \left( \gamma I(a \leq X_i \leq b) \frac{K(pX_i, pt_m)}{G(X_i)} \right) - \int f_1(x) K(px, pt_m) dx
\]

\[
:= \frac{1}{\sqrt{n}} \sum_{i=1}^n u_i.
\]

Then \( EU_1 = 0 \) and from (2.1) and (P3) above we find

\[
EU_1^2 = p \int \frac{\gamma}{G(x)} K^2(px, pt_m) f_1(x) dx - p \left( \int f_1(x) K(px, pt_m) dx \right)^2
\]

\[
= \int G(t_m + u/p) f_1 t_m + \frac{u}{p} K^2(u, 0) du + O(p^{-1})
\]

\[
= \int G(t_m + u/p) f_1 t_m + \frac{u}{p} K^2(u, 0) du + O(p^{-1}) \rightarrow \sigma^2(t).
\]

Let \( n_1 = \lfloor \frac{1}{\log \log n} \rfloor \), \( n_2 = \lfloor \frac{1}{\log \log n} \rfloor \) and \( n_3 = \lfloor \frac{n}{n_1 + n_2} \rfloor \). It is easy to verify that

\[
n_1 n_3 n^{-1} \rightarrow 1, \quad n_2 n_3 n^{-1} \rightarrow 0 \quad \text{and} \quad n_3 \alpha(n_2) \rightarrow 0.
\]

Set \( \sum_{i=1}^n U_i = J_{11} + J_{12} + J_{13} \), where \( J_{11} = \sum_{k=1}^{n_3} V_k \), \( J_{12} = \sum_{k=n_3+1}^{n_1} \sum_{i \in \Delta_k} U_i \),

\[
J_{13} = \sum_{i \in \Delta'_n} U_i,
\]

\( V_k = \sum_{i \in \Delta_k} U_i \), \( \Delta_k = \{(k-1)(n_1 + n_2) + 1, \ldots, (k-1)(n_1 + n_2) + n_1\} \),

\( \Delta'_n = \{(k-1)(n_1 + n_2) + 1, \ldots, n_1 + n_2\} \) and \( \Delta'_n = \{(n_3(n_1 + n_2) + 1, \ldots, n\} \).

Hence, it is sufficient to prove that

\[
n^{-\frac{1}{2}} J_{11} \rightarrow_d N(0, \sigma^2(t)), \quad EU_{12}^2 = o(n) \quad \text{and} \quad EU_{13}^2 = o(n).
\]
By Lemma A.1 we have $|\text{Cov}(U_1, U_{1+k})| = O(\alpha(k))$, which together with (4.9) and (4.10) gives us

$$
\frac{1}{n} \text{Var}\left(\sum_{k=1}^{n_3} V_k\right) = n_3 n^{-1} \text{Var}(V_1) + n^{-1} \sum_{k=1}^{n_3-1} (n_3 - k) \text{Cov}(V_1, V_{1+k})
$$

$$
= n_3 n_1 n^{-1} E \sum_{i=1}^{n_1} (U_1, U_{1+k})
$$

$$
+ O(n_3 n^{-1}) \sum_{k=1}^{n_3-1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_3} \text{Cov}(U_1, U_{j+k(n_1+n_2)})
$$

$$
= n_3 n_1 n^{-1} E \sum_{i=1}^{n_1} (U_1, U_{1+k}) + O(n_3 n_1) + O(n^{-1} n_3^2 n_1^2 \rho^{n_1+n_2}) \to \sigma^2(t).
$$

(4.12)

According to Lemma A.2, from (4.10) we have

$$
\left| E e^{it \sum_{k=1}^{n_3} n^{-1/2} V_k} - \prod_{k=1}^{n_3} E e^{i t n^{-1/2} V_k} \right| \leq 16(n_3 - 1) \alpha(n_2) \to 0. \quad (4.13)
$$

According to $p = O(n^{1/(2r+1)})$ and (4.10), from (P1) we have $V_n = \frac{V_1}{\sqrt{n}} \sum_{i=1}^{n_3} U_i \leq C n^{-1/2} n_1 p^{1/2} = O(\log \log n)^{-1}$ a.s., which leads that the set $\{|V_1| \geq \epsilon \sqrt{n_1} \sqrt{\sigma(t)}\}$ is an empty set for any $\epsilon > 0$ and large $n$. Therefore

$$
\sum_{k=1}^{n_3} E \left\{ \left( \frac{V_k}{\sqrt{n_1} \sqrt{\sigma(t)}} \right)^2 I \left( \left| \frac{V_k}{\sqrt{n_1} \sqrt{\sigma(t)}} \right| \geq \epsilon \right) \right\}
$$

$$
= n_3 E \left\{ \left( \frac{V_1}{\sqrt{n_1} \sqrt{\sigma(t)}} \right)^2 I \left( \left| \frac{V_1}{\sqrt{n_1} \sqrt{\sigma(t)}} \right| \geq \epsilon \right) \right\} \to 0. \quad (4.14)
$$

(4.12)–(4.14) imply that $n^{-1/2} J_{11} \to_d N(0, \sigma^2(t))$.

We observe that

$$
E J_{12}^2 = n_3 E \left( \sum_{i=1}^{n_2} U_i \right)^2 + \sum_{k=1}^{n_3} (n_3 - k) \text{Cov} \left( \sum_{i=1}^{n_2} U_i, \sum_{j=1}^{n_2} U_{j+k(n_1+n_2)} \right)
$$

$$
\leq n_3 n_2 E U_1^2 + n_3 \sum_{i=1}^{n_3} (n_2 - i) |\text{Cov}(U_1, U_{1+i})|
$$

$$
+ n_3 \sum_{k=1}^{n_3-1} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} |\text{Cov}(U_1, U_{j+k(n_1+n_2)})|
$$

$$
= O(n_3 n_2) + n_3 n_2^2 \sum_{k=1}^{n_3-1} \alpha(k(n_1 + n_2)) = o(n).
$$

Similarly, it can be easily obtained that $E J_{13}^2 = o(n)$. Then (4.11) is proved. \(\square\)
Lemma 4.6. We prove $EJ_2^2 = o(n^{-1}p)$. 

Proof. Since $\psi$ has compact support, there are only finite non-zero $l$ terms $\psi_{kl}(t_m)$ for each $k$. Therefore

$$EJ_2^2 = O(q) \sum_{k=0}^{q-1} \sum_{l} E\{\hat{b}_{kl}^2 I(|\hat{b}_{kl}| > \delta)\} \psi_{kl}(t_m)$$

$$= O(q) \sum_{k=0}^{q-1} \sum_{l} E\{(\hat{b}_{kl} - b_{kl})^2 I(|\hat{b}_{kl}| > \delta)\} p_k + O(q) \sum_{k=0}^{q-1} b_{kl}^2 P(|\hat{b}_{kl}| > \delta) p_k$$

$$:= J_{21} + J_{22}.$$ 

By using similar proof method as in Lemma 4.2, it is easy to obtain that $J_{21} = o(n^{-1}p)$. Let $a_3$, $b_3$ and $c_3$ be positive numbers such that $a_3 + b_3 + c_3 = 1$. Note that $|b_{kl}| = O(p_k^{(r+1)/2})$ and $|b_{kl}|\delta^{-1} \to 0$. Then

$$J_{22} \leq O(q) \sum_{k=0}^{q-1} \sum_{l} p_k^{-2r} \{P(|\hat{b}_{kl} - b_{kl}| > a_3\delta)$$

$$+ P(|\hat{b}_{kl} - b_{kl}| > b_3\delta) + P(|\hat{b}_{kl} - b_{kl}| > c_3\delta)\}$$

$$= O(q) \sum_{k=0}^{q-1} \sum_{l} p_k^{-2r} \left(n^{-2r/(2r+1)} \frac{\ln \ln n}{np_k\delta^2}\right) = o(n^{-1}p). \quad \Box$$

Lemma 4.7. We prove $J_3 = o(p^{-r}) = o(\sqrt{n^{-1}p})$.

Proof. In view of (P3) and (P4), by using the Taylor expansion, it follows that

$$J_3 = \sum_{l} \phi_l(t_m) \int \phi_l(x)f_1(x)dx - \int K(pt_m,x)dx f_1(t_m) - b(t_m)$$

$$= \int p(f_1(x) - f_1(t_m))K(px,pt_m)dx - b(t_m)$$

$$= \int(f_1(t_m + u/p) - f_1(t_m))K(pt_m + u,pt_m)du - b(t_m)$$

$$= \int \sum_{i=1}^{r} \frac{f_1^{(i)}(t_m)}{i!} u^{i-1} p^{-i} K(u,0)du - b(t_m) + o(p^{-r})$$

$$= \frac{f_1^{(r)}(t_m)}{r! p^r} \int u^r K(u,0)du - b(t_m) + o(p^{-r}) = o(p^{-r}). \quad \Box$$

Acknowledgments

This work was supported by the National Natural Science Foundation of China (10871146), and also by the Grant MTM2008-03129 of the Spanish Ministry of Science and Innovation and project 10PXIB300068PR of the Xunta de Galicia, Spain.
Appendix

In this section, we give some preliminary lemmas, which have been used in Sec. 4.

Lemma A.1 (Ref. 14, Corollary A.1). Suppose that $X$ and $Y$ are random variables such that $E|X| \leq c_1$, $E|Y| \leq c_2$, then $|EXY - EXEY| \leq 4c_1c_2\sup_{A \in \sigma(X)B \in \sigma(Y)}\{P(A \cup B) - P(A)P(B)\}$.  

Lemma A.2 (Ref. 34). Let $V_1, \ldots, V_m$ be $\alpha$-mixing random variables measurable with respect to the $\sigma$-algebra $F_{i_1}^{n_1}, \ldots, F_{j_m}^{n_m}$, 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 $l, j = 1, 2, \ldots, m$. Then $|E(\prod_{j=1}^{m} V_j) - \prod_{j=1}^{m} EV_j| \leq 16(m - 1)\alpha(w)$, where $F^b_a = \sigma\{V_i, a \leq i \leq b\}$ and $\alpha(w)$ is the mixing coefficient.

Lemma A.3 (Ref. 24, Proposition 5.1). Let $\{X_i\}$ be $\alpha$-mixing random variables with mixing coefficients $\{\alpha(n)\}$. Assume that $EX_i = 0$ and $|X_i| \leq S < \infty$ a.s. ($i = 1, 2, \ldots, n$). Set $D_m = \max_{1 \leq i \leq 2m} \text{Var}(\sum_{i=1}^{j} X_i)$. Then for $n, m \in \mathbb{N}$, $0 < m < n/2$, $\epsilon > 0$,

$$P\left(\frac{\sum_{i=1}^{n} X_i}{n} > \epsilon\right) \leq 4\exp\left\{-\frac{\epsilon^2}{16} \left(nm^{-1}D_m + \frac{1}{3}\epsilon mS\right)^{-1}\right\} + \frac{32S}{\epsilon} \alpha(m).$$

Lemma A.4. Let $\tilde{b}_1$ and $\tilde{b}_{kl}$ be defined in Sec. 2. Set $\gamma_n = (\ln \ln n/n)^{1/2}$, $\varphi_l(x) = \phi_l(x)I(a \leq x \leq b)$ and $\varphi_{kl}(x) = \psi_{kl}(x)I(a \leq x \leq b)$. If (A1) holds, then

$$\tilde{b}_1 = \frac{1}{n} \sum_{i=1}^{n} \varphi_l(X_i) \frac{\gamma}{G(X_i)} + O(\gamma_n) \frac{1}{n} \sum_{i=1}^{n} |\varphi_l(X_i)| \frac{\gamma}{G(X_i)} + O(\gamma_n) \int |\varphi_l(x)| f(x)dx$$

$$:= \tilde{b}_1 + B_1 + \beta_1 \quad \text{a.s.,} \quad (A.1)$$

$$\tilde{b}_{kl} = \frac{1}{n} \sum_{i=1}^{n} \varphi_{kl}(X_i) \frac{\gamma}{G(X_i)} + O(\gamma_n) \sum_{i=1}^{n} |\varphi_{kl}(X_i)| \frac{\gamma}{G(X_i)} + O(\gamma_n) \int |\varphi_{kl}(x)| f(x)dx$$

$$:= \tilde{b}_{kl} + B_{kl} + \beta_{kl} \quad \text{a.s.} \quad (A.2)$$

and

$$\tilde{E}\tilde{b}_1 = b_1, \quad \beta_1 = O(\gamma_n p^{\frac{-\theta}{2}}) \quad \text{a.s.,} \quad (A.3)$$

$$\tilde{E}\tilde{b}_{kl} = b_{kl}, \quad \beta_{kl} = O(\gamma_n p^{\frac{-\theta}{2}}) \quad \text{a.s.} \quad (A.4)$$

Proof. Let $\Lambda(x) = \int_{-\infty}^{x} \frac{dF(u)}{F(u)}$ denote the cumulative hazard function of $F$. From (2.1) and (2.2) it follows that $\Lambda(x) = \int_{0}^{x} \frac{dF^*(u)}{C(u)}$. Hence, a natural estimator of $\Lambda$ is given by $\Lambda_n(x) = \int_{0}^{x} \frac{dF_n^*(u)}{C(u)}$. The proof of Theorem 2.1 in Ref. 28 (see (A.6), (A.11), there) shows that

$$\sup_{\alpha \leq x \leq b} |C_n(x) - C(x)| = O(\gamma_n) \quad \text{a.s.,} \quad \sup_{\alpha \leq x \leq b} |\Lambda_n(x) - \Lambda(x)| = O(\gamma_n) \quad \text{a.s.} \quad (A.5)$$
(A.8) and (A.9) of Ref. 28 give that $1 - F_n(x) - e^{-\Lambda_n(x)} = O(n^{-1}) \int_0^x \frac{dF^\ast(u)}{C^2(u)}$. The definition of $\Lambda$ implies that $F(x) = 1 - e^{-\Lambda(x)}$. Therefore, by the Taylor expansion we have

$$F_n(x) - F(x) = (1 - e^{-\Lambda_n(x)}) - (1 - e^{-\Lambda(x)}) + O(n^{-1}) \int_0^x \frac{dF^\ast(u)}{C^2(u)}$$

$$= (1 - F(x))(\Lambda_n(x) - \Lambda(x)) + O(1)(\Lambda_n(x) - \Lambda(x))^2$$

$$+ O(n^{-1}) \int_0^x \frac{dF^\ast(u)}{C^2(u)}. \quad \text{(A.6)}$$

Now we are ready to prove (A.1) and (A.3), the proof of (A.2) and (A.4) is analogous. Note that $\Lambda_n(x) - \Lambda(x) = \int_0^x C_n^{-1}(u)dF^\ast(u) - \int_0^x C_n^{-1}(u)dF^\ast(u) = \int \varphi(x)(1 - F(x))C^{-1}(x)dF^\ast(x) = \int \varphi(x)f(x)dx$ and $C(x) = \gamma^{-1}G(x)[1 - F(x)]$. Therefore, by the Taylor expansion

$$\tilde{b}_1 = \int \varphi(x)dF_n(x)$$

$$= \int \varphi(x)(1 - F(x))\frac{dF^\ast(x)}{C_n(x)} + O(\gamma_n) \int |\varphi(x)||f(x)|dx + O(\gamma_n) \int |\varphi(x)|\frac{dF^\ast(x)}{C(x)}$$

$$= \frac{1}{n} \sum_{i=1}^n \varphi(X_i) \frac{1 - F(X_i)}{C_n(X_i)} + O(\gamma_n) \int |\varphi(x)||f(x)|dx + O(\gamma_n) \int |\varphi(x)|\frac{dF^\ast(x)}{C(x)}$$

$$= \frac{1}{n} \sum_{i=1}^n \varphi(X_i) \frac{1 - F(X_i)}{C(X_i)} + \frac{1}{n} \sum_{i=1}^n \varphi(X_i)(1 - F(X_i)) \left( \frac{1}{C_n(X_i)} - \frac{1}{C(X_i)} \right)$$

$$+ O(\gamma_n) \frac{1}{n} \sum_{i=1}^n \left| \varphi(X_i) \right| \frac{1}{C(X_i)} + O(\gamma_n) \int |\varphi(x)||f(x)|dx$$

$$= \frac{1}{n} \sum_{i=1}^n \varphi(X_i) \frac{\gamma}{G(X_i)} + O(\gamma_n) \frac{1}{n} \sum_{i=1}^n \left| \varphi(X_i) \right| \frac{\gamma}{G(X_i)} + O(\gamma_n) \int |\varphi(x)||f(x)|dx,$$

which is (A.1). From (2.1) we have

$$E\tilde{b}_1 = E(\varphi(X_1)G^{-1}(X_1)\gamma) = \int \varphi(x)G^{-1}(x)\gamma dF^\ast(x) = \int \phi(x)f_1(x)dx = b_n.$$ 

Since $f$ is bounded and $\phi$ is bounded and compactly supported, we have

$$\beta_t = O(\gamma_n) \int |\varphi(x)||f(x)|dx$$

$$= O(\gamma_n p^{-\frac{1}{2}}) \int |\phi(y)||f_1((y + l)/p)|dy = \Omega(\gamma_n p^{-\frac{1}{2}}) \quad \text{a.s.}$$

Then (A.3) is proved. The proof of Lemma A.4 is completed.

**Lemma A.5.** Let the definition of $\tilde{b}_1$ be as in Lemma A.4, then under the assumptions of Theorem 3.1 we have $\text{Var} \left( \sum (\tilde{b}_1 - b_1)^2 \right) = o(n^{-2}p^2).$
Proof. Put $Z_l = \varphi_l(X_l)G^{-1}(X_l)\gamma - b_l$. We observe that
\[
n^2 \sum_l \left(\tilde{b}_l - b_l\right)^2 = n \sum_{i=1}^n \sum_l Z_{li}^2 + \sum_{i \neq j} \sum_l Z_{li}Z_{lj}.
\] (A.7)
Note that \(\{Z_{li}, i = 1, \ldots, n\}\) is a sequence of identically distributed random variables for any \(l \geq 0\).

\[
\text{Var}\left(\sum_{i=1}^n \sum_l Z_{li}^2\right) = n \text{Var}\left(\sum_l Z_{li}^2\right) + 2 \sum_{i=1}^{n-1} (n - i) \text{Cov}\left(\sum_l Z_{li}^2, \sum_l Z_{li+1}^2\right).
\] (A.8)

Since \(\phi(x)\) has compact support \([-v, v]\) and \(b_l^2 = O(p^{-1})\), we have
\[
\sum_l Z_{li}^2 \leq 2 \left(\sum_l \varphi_l^2(X_l)G^{-2}(X_l)\gamma^2 + \sum_l b_l^2\right) \\
\leq 2(p \sup x \varphi^2 ([|x|] + 2) \sup G^{-2}(x)\gamma^2 + O(1)) = O(p), \quad (A.9)
\]
Lemma A.1 implies that \(\sum_{i<j} \text{Cov}(\sum_l Z_{li}, \sum_l Z_{l+i}^2) = O(p^2 \sum_{i} \alpha(i))\), which together with (A.8) and (A.9), yields that
\[
\text{Var}\left(n^{-2} \sum_i \sum_l Z_{li}^2\right) = n^{-4} \text{Var}\left(\sum_i \sum_l Z_{li}^2\right) = O(n^{-3}p^2). \quad (A.10)
\]

Next we evaluate an upper bound of
\[
E\left(\sum_{l \neq j} Z_{li}Z_{lj}\right)^2 = \sum_{l_1,l_2,i \neq j_1,i \neq j_2} E(Z_{l_1,i}Z_{l_1,j}Z_{l_2,j}Z_{l_2,j}).
\]
This can be obtained by considering several cases of the indices in the sums. Lemma A.1 will be repeatedly used here.

Case 1. Suppose the indices satisfy \(i_1 = i_2 = i\) and \(j_1 = j_2 = j\).
\[
\sum_{l_1,l_2,i \neq j} E(Z_{l_1,i}Z_{l_1,j}Z_{l_2,j}) \\
= \sum_{|l_1 - l_2| \leq 2v,i} (n - i) E(Z_{l_1,i}Z_{l_1,i}Z_{l_2,j+1}) \\
= n \sum_{|l_1 - l_2| \leq 2v,i} \left(\{E(Z_{l_1,i}Z_{l_1,j})/O(p^2\alpha(i))\} = O(n^2p) + O(np^3) = O(n^2p).
\]

Case 2. Suppose the indices satisfy \(i_1 = i_2 = i\) and \(i < j_1 < j_2\). First we have \(p^{-1}|Z_{l_1,i}Z_{l_1,j_1}Z_{l_2,j_2}| \leq C\), \(E|Z_{l_1,i}Z_{l_2,j}| = O(p^{-1})\), then \(E(|Z_{l_1,i}Z_{l_1,j}Z_{l_2,j}|) \leq C\),
\[
\sum_{l_1,l_2,1 \leq j_1 < j_2 \leq n} E(Z_{l_1,i}Z_{l_1,j_1}Z_{l_2,j_2}) \\
= \sum_{|l_1 - l_2| \leq 2v,1 \leq m_1 < m_2 < n} (n - m_1 - 1) E(Z_{l_1,i}Z_{l_2,j_1}Z_{l_1+m_1}Z_{l_2+m_1+m_2})
\]
\[
\begin{align*}
&\leq \sum_{|l_1-l_2|\leq 2v, 1 \leq m_1 \leq m_1 + m_2 < n} n \min\{C_5, E(Z_{l_1l_1}Z_{l_2l_2})E(Z_{l_1l_1+m_1}Z_{l_2l_2+m_1+m_2})
+ 4p^2\alpha(m_1)\} \\
&= O(n^2) \sum_{|l_1-l_2|\leq 2v} \left\{ \sum_{1 \leq m_1 \leq N_n} C + \sum_{1 < m_2 < n} \min\{p^{-1}, \rho\alpha(m_2)\}\right\} O(1) \\
&+ O\left( \sum_{N_n < m_1 < n} p^2\alpha(m_1) \right) = O(n^2 p N_n) = o(n^2 p^2),
\end{align*}
\]

where \(N_n = o(p)\) and \(N_n \to \infty\).

**Case 3.** Suppose the indices satisfy \(i_1 < j_1 < i_2 < j_2\). Taking \(N = O(p^{1/3})\) such that \(N^3/n \to 0\), \(N/p \to 0\) and \(np^2 \rho^N \to 0\),

\[
\sum_{i_1,i_2,j_1,j_2} E(Z_{l_1l_1}Z_{l_2l_1}Z_{l_2l_2})
\leq \sum_{i_1,i_2,m_1+m_2+m_3+1 < n} nE(Z_{l_1l_1}Z_{l_2l_1+m_1+1}Z_{l_2l_1+m_2+2}Z_{l_2l_1+m_3+3})
\]

\[
= \sum_{i_1,i_2} \left( \sum_{D_1} + \sum_{D_2} + \sum_{D_3} + \sum_{D_4} \right) E(Z_{l_1l_1}Z_{l_2l_1+m_1+1}Z_{l_2l_1+m_2+2}Z_{l_2l_1+m_3+3})
\]

\[
:= d_1 + d_2 + d_3 + d_4,
\]

where \(D_1 = \{(m_1,m_2,m_3): \sum_{i=1}^3 m_i < n, 0 < m_1, m_2, m_3 \leq N\}\),

\[
D_2 = \left\{(m_1,m_2,m_3): \sum_{i=1}^3 m_i < n, m_1 > N\right\},
\]

\[
D_3 = \left\{(m_1,m_2,m_3): \sum_{i=1}^3 m_i < n, m_1 < N, m_2 > N\right\}
\]

and \(D_4 = \{(m_1,m_2,m_3): \sum_{i=1}^3 m_i < n, m_1 < N, m_2 \leq N, m_3 > N\}\). It is easy to see that

\[
d_1 \leq \sum_{i_1,i_2,D_1} E|Z_{l_1l_1}Z_{l_1l_1+m_1+1}Z_{l_2l_1+m_2+2}Z_{l_2l_1+m_3+3}| \leq O(p^2 n N^3) = o(p^2 n^2),
\]

\[
d_2 \leq \sum_{i_1,i_2,D_2} \{E(Z_{l_1l_1})E(Z_{l_1l_1+m_1+1}Z_{l_2l_1+m_2+2}Z_{l_2l_1+m_3+3}) + O(p^2 \alpha(m_1))\}
\]

\[
\leq O\left( p^2 n^3 \sum_{m_2 > N} p^2 \alpha(m_2) \right) = O(n^3 p^4 \rho^N) = o(p^2 n^2),
\]
Lemma A.6. Let
\[ d_3 \leq \sum_{l_1, l_2, d_3} \{ E(Z_{l_1}Z_{l_1+m_1+1})E(Z_{l_2}Z_{l_2+m_2+1}Z_{l_2+m_2+m_3+1}) + O(p^2 \alpha(m_3)) \} \]
\[ \leq n \sum_{l_1, l_2} \left\{ \sum_{m_1 < N} E(Z_{l_1}Z_{l_1+m_1}) \sum_{i \neq j} E(Z_{i_1}Z_{i_2}) \right\} + O(p^4 N n^2 \rho N) \]
\[ = O(pn^2 N) + O(p^4 N n^2 \rho N) = o(p^2 n^2). \]

Similarly, one can verify that \( d_4 = o(p^2 n^2). \) So \( E(\sum_{i \neq j} Z_{l_i}Z_{l_j})^2 = o(n^2 p^2), \) which together with (A.7) and (A.10), yields that \( \text{Var}(\sum_i (\tilde{b}_i - b_i)) = o(n^{-2} p^2). \)

\[ \square \]

**Lemma A.6.** Let \( \tau \) be a positive integer satisfying \( \tau > 2, \) then under the assumptions of Theorem 3.1 we have \( E|b_{kl} - b_k|^\tau = O(n^{-\tau/2}), \) where the definition of \( \{ b_{kl}, 0 \leq k \leq q - 1, l \in \mathbb{Z} \} \) is in Lemma A.4.

**Proof.** According to the Hölder inequality, it suffices to show that \( E(\tilde{b}_{kl} - b_{kl})^\tau = O(n^{-\tau/2}) \) for even number \( \tau. \) The proof follows the line as in the proof of Lemma 9 in Ref. 30.

Set \( Z_i = \varphi_{kl}(X_i)G^{-1}(X_i)\gamma - b_{kl}. \) Then \( (\tilde{b}_{kl} - b_{kl})^\tau = (n^{-1} \sum_{i=1}^n Z_i)^\tau \) and
\[ E \left( \sum_{i=1}^n Z_{kt} \right)^\tau = \sum_{t} \sum_{\tau_i} \sum_{I_2} |E(Z_{i_1}^{\tau_1}Z_{i_1+i_2}^{\tau_2} \cdots Z_{i_1+i_n}^{\tau_n})|, \]
where \( T = \{ t: t = 1, 2, \ldots, \tau \}, \) \( \tau_i = (\tau_1, \ldots, \tau_t): \tau_1, \ldots, \tau_t > 0, \sum_{i=1}^t \tau_i = \tau \) and \( I_t = \{ (i_1, \ldots, i_t): i_1, \ldots, i_t > 0, \sum_{j=1}^t i_j \leq n \}. \) Since \( \tau_i \) and \( T \) have finite elements, we only need to prove that
\[ \sum_{I_t} |E(Z_{i_1}^{\tau_1}Z_{i_1+i_2}^{\tau_2} \cdots Z_{i_1+i_n}^{\tau_n})| = O(n^{\tau/2}). \] (A.11)

It is easy to verify that (A.11) is valid for \( \tau = 1, 2. \) Choosing \( N_2 = O(p^{1/(\tau - 1)}) \) such that \( p_k^{-1} N_2^{\tau - 1} = O(1) \) and \( n^{\tau - 1} \rho N_2 = o(1). \) By induction on \( \tau, \) where \( \sum_{i=1}^t \tau_i = \tau, \) assume that (A.11) holds for \( \tau - 1, \tau \geq 2. \) Set \( I_t = \cup_{d=1}^t I_d, \) where \( I_1 = \{ (i_1, \ldots, i_t): 0 < i_2, \ldots, i_t < N_2 \} \) and
\[ I_d = \{ (i_1, \ldots, i_t): 0 < i_2, \ldots, i_{d-1} < N_2, i_d > N_2 \} \]
for \( 2 \leq d \leq t. \)

Since \( |Z_i| = O(p_k^{-1/2}) \) a.s. and \( E|Z_iZ_j| = O(p_k^{-1}) \) for \( i \neq j, \)
\[ \sum_{I_t} |E(Z_{i_1}^{\tau_1}Z_{i_1+i_2}^{\tau_2} \cdots Z_{i_1+i_n}^{\tau_n})| = O(p_k^{-1} p_k^{-1} N_2^{\tau - 1}) \]
\[ = o(n^\tau p_k^{-1} N_2^{\tau - 1}) = o(n^\tau). \] (A.12)
For \( \{I_d, 2 \leq d \leq t \} \), since the \( Z_i \)'s are bounded by \( p_k^\frac{\delta}{2} \), it follows from Lemma A.1 that

\[
|E(Z_{i_1}^{\tau_1} Z_{i_2}^{\tau_2} \cdots Z_{i_d}^{\tau_d})|
\]

\[
\leq |E(Z_{i_1}^{\tau_1} \cdots Z_{i_d}^{\tau_d})| |E(Z_{i_1}^{\tau_1} \cdots Z_{i_d}^{\tau_d})| + 4p_k^\delta \alpha(i_d).
\]

Consequently, by the inductive hypothesis, for \( 2 \leq d \leq t \),

\[
\sum_{I_d} |E(Z_{i_1}^{\tau_1} Z_{i_2}^{\tau_2} \cdots Z_{i_d}^{\tau_d})|
\]

\[
\leq \sum_{I_d} |E(Z_{i_1}^{\tau_1} \cdots Z_{i_d}^{\tau_d})| |E(Z_{i_1}^{\tau_1} \cdots Z_{i_d}^{\tau_d})| + 4p_k^\delta \sum_{I_d} \alpha(i_d)
\]

\[
= O(n^{\tau_1/2} + O(p_k^{\tau_2/2} n^{-1} \rho N_z) = O(n^{\tau_1/2}).
\]

(13)

Therefore, from (12) and (13) one obtains (A.11).

\[ \square \]

Lemma A.7. Under the assumptions of Theorem 3.1, for \( \delta > C_0 \sqrt{\frac{\log n}{n}} \) with \( C_0 \) properly chosen and positive constants \( \theta_1 \) and \( \theta_2 \) we have \( \{ |b_{kl} - b_k| > \theta_1 \delta \} \) \( \subset \Omega \subset O(n^{-2r/(2r+1)}) \), where the definition of \( \{ b_{kl}, 0 \leq k \leq q-1, l \in Z \} \) is in Lemma A.4.

Proof. Here, we apply Lemma A.3 to evaluate \( P(|b_{kl} - b_k| > \theta_1 \delta) \). Set

\[
W_i = n^{-1/2}(\varphi_{kl}(X_i) G^{-1}(X_i) \gamma - b_k)
\]

and

\[
M_i = \sup_{l,k,y \in [a,b]} \{ \phi(y_n^{1/2} y - l) G^{-1}(y) \gamma \}.
\]

Write \( S = \sqrt{p_k n^{-1} M_1} \) then \( |W_1| \leq S \) a.s. By similar estimation as for \( E(\hat{b}_t - b_t)^2 = O(n^{-1}) \) we have \( D_{m_1} := \max_{1 \leq j \leq 2m_1} \var{ \sum_{i=1}^{d} W_i } = O(m_1 n^{-1}) \). Then, in view of Lemma A.3 and choosing \( m_1 = \lfloor \delta \sqrt{p_k} \rfloor \), we have

\[
P(|b_{kl} - b_k| \geq \theta_1 \delta)
\]

\[
= P\left( \sum_{i=1}^{d} W_i \geq \sqrt{n} \theta_1 \delta \right)
\]

\[
\leq 4 \exp \left( -\theta_1^2 \frac{\theta_1^2}{16} \left( \frac{n m_1^{-1} D_{m_1} + \frac{1}{3} \sqrt{n} \theta_1 \delta S M_1}{16} \right)^{-1} \right) + \frac{S}{\theta_1} \delta \sqrt{n} \alpha(m_1)
\]

\[
\leq 4 \exp \left( -\theta_1^2 \frac{C_0}{16} \right) + O(p_k^{1/2} \delta^{-1} \rho m_1) = O\left( n^{-2r/2r+1} \right),
\]

where \( C_0 = \theta_1^2 (nm_1^{-1} D_{m_1} + \frac{1}{3} \sqrt{p_k} \delta m_1 M_1)^{-1} \), now select \( \delta \geq C_0^{-1/2} \log n \) with \( C_0 \) such that \( \frac{1}{10} C_0 \delta^2 > 2r/(2r+1) \), then \( P(|b_{kl} - b_k| > \theta_1 \delta) \subset O(n^{-2r/(2r+1)}) \).
References


