Chin. Ann. Math.
32B(4), 2012, 609–624
DOI: 10.1007/s11401-012-0718-z

Chinese Annals of Mathematics, Series B

Convergence Rates of Wavelet Estimators in Semiparametric Regression Models Under NA Samples*

Hongchang HU[1] Li WU[1]

Abstract Consider the following heteroscedastic semiparametric regression model:

$$y_i = X_i^{\mathrm{T}}\beta + g(t_i) + \sigma_i e_i, \quad 1 \le i \le n,$$

where $\{X_i,\ 1 \le i \le n\}$ are random design points, errors $\{e_i,\ 1 \le i \le n\}$ are negatively associated (NA) random variables, $\sigma_i^2 = h(u_i)$, and $\{u_i\}$ and $\{t_i\}$ are two nonrandom sequences on $[0,1]$. Some wavelet estimators of the parametric component β, the nonparametric component $g(t)$ and the variance function $h(u)$ are given. Under some general conditions, the strong convergence rate of these wavelet estimators is $O(n^{-\frac{1}{3}}\log n)$. Hence our results are extensions of those results on independent random error settings.

Keywords Semiparametric regression model, Wavelet estimate, Negatively associated random error, Strong convergence rate
2000 MR Subject Classification 62G05, 62G20

1 Introduction

Consider the following heteroscedastic semiparametric regression model:

$$y_i = X_i^{\mathrm{T}}\beta + g(t_i) + \sigma_i e_i, \quad 1 \le i \le n, \tag{1.1}$$

where $\{y_i\}$ are scalar response variables, $\beta = (\beta_1, \beta_2, \cdots, \beta_d)^{\mathrm{T}}$ is an unknown d-dimensional parameter vector, $\sigma_i^2 = h(u_i)$, $g(\cdot)$ and $h(\cdot)$ are unknown functions defined on $[0,1]$, $\{u_i\}$ and $\{t_i\}$ are two nonrandom sequences on $[0,1]$, $X_i = (x_{i1}, x_{i2}, \cdots, x_{id})^{\mathrm{T}}$ are random design points with $d \le n$, and random errors $\{e_i, 1 \le i \le n\}$ are negatively associated (its definition is given by the following) random variables with $Ee_i = 0$ and $Ee_i^2 = 1$.

Following [1], denote

$$x_{ir} = f_r(t_i) + \eta_{ir}, \quad 1 \le i \le n,\ 1 \le r \le d, \tag{1.2}$$

where $f_r(\cdot)$ is some unknown smooth function defined on $[0,1]$, $\{\overline{\eta}_i\}$ is a stochastic sequence with $\overline{\eta}_i = (\eta_{i1}, \cdots \eta_{id})^{\mathrm{T}}$ i.i.d. and

$$E\overline{\eta}_i = 0, \quad \mathrm{Var}(\overline{\eta}_i) = V, \tag{1.3}$$

Manuscript received September 4, 2010. Revised April 14, 2012.
[1]College of Mathematics and Statistics, Hubei Normal University, Huangshi 435002, Hubei, China.
E-mail: retutome@163.com wuli@163.com
*Project supported by the National Natural Science Foundation of China (No. 11071022), the Key Project of the Ministry of Education of China (No. 209078) and the Youth Project of Hubei Provincial Department of Education of China (No. Q20122202).

where $V = (V_{ij})$ $(i, j = 1, 2, \cdots, d)$ is a positive definite matrix with d-order. Moreover, $\{\eta_{ir}\}$ and $\{e_i\}$ are independent of each other.

Definition 1.1 (see [2]) Random variables $\xi_1, \xi_2, \cdots, \xi_n$, $n \geq 2$ are said to be negatively associated (NA) if for every pair of disjoint subsets B_1 and B_2 of $\{1, 2, \cdots, n\}$,

$$\mathrm{Cov}(l_1(\xi_i, i \in B_1), l_2(\xi_i, i \in B_2)) \leq 0,$$

where l_1 and l_2 are increasing for every variable (or decreasing for every variable) such that this covariance exists. A sequence of random variables $\{\xi_i, i \geq 1\}$ is said to be NA if every finite subfamily is NA.

The definition of NA random variables is introduced by Alam and Saxena [3] and is carefully studied by Joag-Dev and Proschan [4]. Because of its wide applications in the multivariate statistical analysis and the system reliability, the notion of NA has received considerable attention recently. There are some fundamental convergence results for NA random variables, and we refer to Joag-Dev and Proschan [4] for fundamental properties, Matula [5] for the three series theorem, Shao and Su [6] for the law of the iterated logarithm, Wu and Jiang [7] for the law of the iterated logarithm of partial sums, Liang [8] and Baek et al. [9] for the complete convergence, Liang and Zhang [10] for the strong convergence for weighted sums of NA arrays, Su et al. [11] for the moment inequality and the weak convergence, Jin [12] for the convergence rate in the law of logarithm for NA random fields, and Roussas [13] for the central limit theorem of random fields.

When $\{\sigma_i e_i\}$ are independent and identically distributed random variables, the model (1.1) can be reduced to the usual semiparametric regression model. Chen [14], Qian and Cai [15], Xue and Liu [16], Chai and Xu [17], Hu [18], Heckman [19], Green and Silverman [20], Härdle et al. [21], Shi and Li [22] and Bianco and Boente [23] used various estimation methods (the piecewise polynomial method, and the wavelet method, the ridge method, the spline method, the penalized least squares method, the kernel smoothing method, the robust estimate method) to obtain estimators of the unknown parametric components in (1.1) and discussed some properties of these estimators. However, the independence assumption for the errors is not always appropriate in applications. Recently, the semiparametric regression model with correlated errors has attracted attentions of many authors, such as Hu and Hu [24–25], You and Chen [26], Wu and Jiang [7], Liang and Jing [27], Zhou et al. [28], and so on.

When $\{X_i, 1 \leq i \leq n\}$ are fixed design points, a few authors studied the model (1.1) with NA errors. For instance, Baek and Liang [29] discussed the strong consistency and the asymptotic normality of weighted least-squares estimators; Liang and Wang [30] discussed the convergence rate of wavelet estimator; and Ren and Chen [31] obtained the strong consistency of a class of estimators. However, there are few asymptotic results for the estimators of parametric and nonparametric components in model (1.1) with NA errors when $\{X_i, 1 \leq i \leq n\}$ are random design points.

By using the wavelet method, we continue to discuss the semiparametric regression model with an NA error sequence, and obtain the strong convergence rates of wavelet estimators in this paper. The organization of this paper is as follows. In Section 2, the wavelet estimators of parametric component β, the nonparametric component $g(t)$ and the variance function $h(u)$ are given by the wavelet smoothing method. Under some general conditions, the strong convergence rates of the wavelet estimators are investigated in Section 3. The main proofs are presented in Section 4.

2 Wavelet Estimators

In the following, we apply the wavelet technique to estimate β, $g(t)$ and $h(u)$. Suppose that there exists a scaling function $\phi(x)$ in the Schwartz space S_l and a multiresolution analysis $\{V_m\}$ in the concomitant Hilbert space $L^2(\mathbb{R})$, with its reproducing kernel $E_m(t,s)$ given by

$$E_m(t,s) = 2^m E_0(2^m t, 2^m s) = 2^m \sum_{k\in\mathbb{Z}} \phi(2^m t - k)\phi(2^m s - k). \tag{2.1}$$

Let $A_i = [s_{i-1}, s_i]$ denote intervals that partition $[0,1]$ with $t_i \in A_i$ and $1 \le i \le n$. Then we can define some wavelet estimators as follows:

Firstly, suppose that β is known. We define the estimator of $g(\cdot)$ by

$$\widehat{g}_0(t,\beta) = \sum_{i=1}^{n} (y_i - X_i^{\mathrm{T}}\beta) \int_{A_i} E_m(t,s)\mathrm{d}s. \tag{2.2}$$

In succession, we define the wavelet estimator $\widehat{\beta}_n$ given by

$$\widehat{\beta}_n = \arg\min_{\beta} \sum_{i=1}^{n} (y_i - X_i^{\mathrm{T}}\beta - \widehat{g}_0(t,\beta))^2 = (\widetilde{X}^{\mathrm{T}}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{Y}, \tag{2.3}$$

where $X = (X_{ir})_{n\times d}$, $Y = (y_1, \cdots, y_n)^{\mathrm{T}}$, $S = (S_{ij})_{n\times n}$, $S_{ij} = \int_{A_j} E_m(t_i,s)\mathrm{d}s$, $\widetilde{X} = (I-S)X$, $\widetilde{Y} = (I-S)Y$.

Finally, we define the linear wavelet estimators of $g(\cdot)$ and $h(\cdot)$ given by

$$\widehat{g}_n(t) = \widehat{g}_0(t,\widehat{g}_n) = \sum_{i=1}^{n} (y_i - X_i^{\mathrm{T}}\widehat{\beta}_n) \int_{A_i} E_m(t,s)\mathrm{d}s, \tag{2.4}$$

$$\widehat{h}_n(u) = \sum_{i=1}^{n} (\widetilde{y}_i - \widetilde{X}_i^{\mathrm{T}}\widehat{\beta}_n)^2 \int_{A_i} E_m(u,s)\mathrm{d}s. \tag{2.5}$$

It is well-known that the weighted least squares estimator is superior to the least squares estimator for a linear regression model with the heteroscedastic errors. Thus, the above estimators are modified by the following estimators:

$$\widetilde{\beta}_n = (\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{Y}, \tag{2.6}$$

$$\widetilde{g}_n(t) = \widehat{g}_0(t,\widetilde{\beta}_n) = \sum_{i=1}^{n} (y_i - X_i^{\mathrm{T}}\widetilde{\beta}_n) \int_{A_i} E_m(t,s)\mathrm{d}s, \tag{2.7}$$

$$\widetilde{h}_n(u) = \sum_{i=1}^{n} (\widetilde{y}_i - \widetilde{X}_i^{\mathrm{T}}\widetilde{\beta}_n)^2 \int_{A_i} E_m(u,s)\mathrm{d}s, \tag{2.8}$$

where $\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2)$.

Since $\sigma_i^2 = h(u_i)$ are unknown functions, the above estimators (2.6)–(2.8) can not be directly used. Because $\widehat{h}_n(u)$ is a uniformly strong consistency estimator of $h(u)$ (see the following Theorem 3.1), it is very natural that $h(u)$ is instead of $\widehat{h}_n(u)$ in (2.6)–(2.8). Thus, we obtain the other estimators given by

$$\overline{\beta}_n = (\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{Y}, \tag{2.9}$$

$$\overline{g}_n(t) = \sum_{i=1}^n (y_i - X_i^{\mathrm{T}}\overline{\beta}_n)\int_{A_i} E_m(t,s)\mathrm{d}s, \tag{2.10}$$

$$\overline{h}_n(u) = \sum_{i=1}^n (\widetilde{y}_i - \widetilde{X}_i^{\mathrm{T}}\overline{\beta}_n)^2\int_{A_i} E_m(u,s)\mathrm{d}s. \tag{2.11}$$

To obtain our results, the following five conditions are sufficient:

(A_1) $g(\cdot), f_r(\cdot), h(\cdot) \in H^\alpha$ (Sobolev space), for some $\alpha > \frac{1}{2}$, $1 \le r \le d$.

(A_2) $g(\cdot), f_r(\cdot)$ and $h(u)$ are Lipschitz functions of order $\gamma > 0$, $1 \le r \le d$.

(A_3) $\phi(\cdot)$ belongs to S_l, which is a Schwartz space for $l \ge \alpha$. ϕ is a Lipschitz function of order 1 and has compact support, in addition to $|\widehat{\phi}(\xi) - 1| = O(\xi)$ as $\xi \to 0$, where $\widehat{\phi}$ denotes the Fourier transform of ϕ.

(A_4) s_i $(i = 1, \cdots, n)$ and m satisfy $\max\limits_{1\le i\le n}(s_i - s_{i-1}) = O(n^{-1})$ and $2^m = O(n^{\frac{1}{3}})$, respectively.

(A_5) $0 < m_0 \le \min\limits_{1\le i\le n} h(u_i) \le \max\limits_{1\le i\le n} h(u_i) \le M_0 < \infty$.

3 Statements of the Results

Theorem 3.1 *Suppose that conditions* (A_1)–(A_4) *hold. If* $\sup\limits_{i\ge 1} E|e_i|^p < \infty$ *for some* $p > 3$ *and* $E\eta_{1j}^2 < \infty$ $(j = 1, 2, \cdots, d)$, *then for* $\gamma \ge \frac{1}{3}$ *and* $\alpha > \frac{3}{2}$,

$$\sup_{1\le i\le d} |\widehat{\beta}_{ni} - \beta_i| = O(n^{-\frac{1}{3}}\log n), \quad a.s., \ n \to \infty,$$

where $\widehat{\beta}_{ni}$ *is the* i*th component of* $\widehat{\beta}_n$.

Theorem 3.2 *Under the same assumptions in Theorem* 3.1, *we have*

$$\sup_{0\le t\le 1} |\widehat{g}_n(t) - g(t)| = O(n^{-\frac{1}{3}}\log n), \quad a.s., \ n \to \infty.$$

Theorem 3.3 *Under the same assumptions in Theorem* 3.1, *and* $\sup\limits_{i\ge 1} E|e_i|^4 < \infty$, *we have*

$$\sup_{0\le u\le 1} |\widehat{h}_n(u) - h(u)| = O(n^{-\frac{1}{3}}\log n), \quad a.s., \ n \to \infty.$$

Theorem 3.4 *Under the same assumptions in Theorem* 3.1, *we have*

$$\sup_{1\le i\le d} |\widetilde{\beta}_{ni} - \beta_i| = O(n^{-\frac{1}{3}}\log n), \quad a.s., \ n \to \infty.$$

Theorem 3.5 *Under the same assumptions in Theorem* 3.1, *we have*

$$\sup_{0\le t\le 1} |\widetilde{g}_n(t) - g(t)| = O(n^{-\frac{1}{3}}\log n), \quad a.s., \ n \to \infty.$$

Theorem 3.6 *Under the same assumptions in Theorem* 3.3*, we have*

$$\sup_{0\le u\le 1}|\widetilde{h}_n(u)-h(u)|=O(n^{-\frac{1}{3}}\log n),\quad a.s.,\ n\to\infty.$$

Theorem 3.7 *Under the same assumptions in Theorem* 3.1*, we have*

$$\sup_{1\le i\le d}|\overline{\beta}_{ni}-\beta_i|=O(n^{-\frac{1}{3}}\log n),\quad a.s.,\ n\to\infty.$$

Theorem 3.8 *Under the same assumptions in Theorem* 3.1*, we have*

$$\sup_{0\le t\le 1}|\overline{g}_n(t)-g(t)|=O(n^{-\frac{1}{3}}\log n),\quad a.s.,\ n\to\infty.$$

Theorem 3.9 *Under the same assumptions in Theorem* 3.3*, we have*

$$\sup_{0\le u\le 1}|\overline{h}_n(u)-h(u)|=O(n^{-\frac{1}{3}}\log n),\quad a.s.,\ n\to\infty.$$

Remark 3.1 Theorem 3.1 is the same as [15, Theorem 2], where the errors are i.i.d. random variables. In this paper, however, the errors are NA sequences. Hence the result is an extension of Qian and Cai [15]. In addition, we easily obtain the strong consistencies and a weak convergence rate of the wavelet estimators by the above results. Hence our results are extensions of those results on independent random error settings, such as Qian and Cai [15], Zhou and You [32], and so on.

Remark 3.2 From the following proofs, we know that Lemma 4.4 and Lemma 4.5 are crucial results for investigating the above convergence rates. Thus the convergence rates of the parametric and nonparametric components are the same under the same conditions. Qian and Cai [15] obtained different convergence rates under different conditions. We think that we shall obtain different convergence rates under different conditions by their technique.

4 Proofs of Theorems

Throughout this paper, let C denote a generic positive constant which could take different values at each occurrence. To prove the main results, we first introduce some lemmas.

Lemma 4.1 (see [33]) *If condition* (A$_3$) *holds, then*

(I) $|E_0(t,s)|\le \dfrac{C_k}{(1+|t-s|)^k}$ *and* $|E_m(t,s)|\le \dfrac{2^mC_k}{(1+2^m|t-s|)^k}$, *where k is a positive integer and C_k is a constant depending on k only.*

(II) $\sup_{0\le s\le 1}|E_m(t,s)|=O(2^m)$.

(III) $\sup_t\int_0^1|E_m(t,s)|\mathrm{d}s\le C$, *where C is a positive constant.*

Lemma 4.2 (see [24]) *Let $\tau_m=2^{-m(\alpha-\frac{1}{2})}$ when $\frac{1}{2}<\alpha<\frac{3}{2}$, $\tau_m=\sqrt{m}\cdot 2^{-m}$ when $\alpha=\frac{3}{2}$, and $\tau_m=2^{-m}$ when $\alpha>\frac{3}{2}$. If conditions* (A$_1$)–(A$_4$) *hold, then*

$$\sup_t\Big|f_j(t)-\sum_{k=1}^n\Big(\int_{A_k}E_m(t,s)\mathrm{d}s\Big)f_j(t_k)\Big|=O(n^{-\gamma})+O(\tau_m),$$

$$\sup_t \Big|g(t) - \sum_{k=1}^n \Big(\int_{A_k} E_m(t,s)\mathrm{d}s\Big) g(t_k)\Big| = O(n^{-\gamma}) + O(\tau_m).$$

Lemma 4.3 (see [34]) *Let $\{e_i, i \geq 1\}$ be an NA sequence with $\sup_{i\geq 1} E|e_i|^p < \infty$ for some $p > 3$. Assume that $a_{ni}(t)$ are nonnegative functions on $[0,1]$ such that $\sum_{i=1}^n a_{ni}(t) \leq c$ and $\max_{1\leq i\leq n} a_{ni}(t) n^{\frac{2}{3}} \leq c$ for any $t \in [0,1]$, and $\max_{1\leq i\leq n} |a_{ni}(t_1) - a_{ni}(t_2)| \leq c|t_1 - t_2|$ for any $t_1, t_2 \in [0,1]$. Then*

$$\sup_{0\leq t\leq 1} \Big|\sum_{i=1}^n a_{ni}(t)\varepsilon_i\Big| = O(n^{-\frac{1}{3}} \log n), \quad a.s.,\ n \to \infty.$$

Lemma 4.4 *Let $\{e_i, i \geq 1\}$ be an NA sequence with $\sup_{i\geq 1} E|e_i|^p < \infty$ for some $p > 3$. Suppose that conditions* (A_3) *and* (A_4) *hold. Then*

$$\sup_{0\leq t\leq 1} \Big|\sum_{i=1}^n e_i \int_{A_i} E_m(t,s)\mathrm{d}s\Big| = O(n^{-\frac{1}{3}} \log n), \quad a.s.,\ n \to \infty.$$

Proof Let $a_i(t) = \int_{A_i} E_m(t,s)\mathrm{d}s$. Then by Lemma 4.3,

$$\sup_{0\leq t\leq 1} \sum_{i=1}^n \Big|\int_{A_i} E_m(t,s)\mathrm{d}s\Big| \leq \sup_{0\leq t\leq 1} \int_0^1 |E_m(t,s)|\mathrm{d}s \leq c$$

and

$$\sup_{0\leq t\leq 1} \max_{1\leq i\leq n} \Big|\int_{A_i} E_m(t,s)\mathrm{d}s\Big| n^{\frac{2}{3}} \leq cn^{-1}2^m n^{\frac{2}{3}} \leq c.$$

Since $E_0(t,s)$ satisfies Lipschitz condition with order 1 for any t, we obtain

$$\begin{aligned}\max_{1\leq i\leq n} \Big|\int_{A_i} E_m(t_1,s)\mathrm{d}s - \int_{A_i} E_m(t_2,s)\mathrm{d}s\Big| &\leq \int_{A_i} |E_m(t_1,s) - E_m(t_2,s)|\mathrm{d}s \\ &\leq cn^{-1}2^m |E_0(2^m t_1, 2^m s) - E_0(2^m t_2, 2^m s)| \\ &\leq cn^{-1}2^m |2^m t_1 - 2^m t_2| \leq c|t_1 - t_2|.\end{aligned}$$

By Lemma 4.3, we obtain Lemma 4.4.

Lemma 4.5 (see [35]) *Suppose that conditions* (A_3) *and* (A_4) *hold, and $E\eta_{1j}^2 < \infty$ $(j = 1, 2, \cdots, d)$. Then*

$$\sup_t \Big|\sum_{k=1}^n \eta_{kj} \int_{A_k} E_m(t,s)\mathrm{d}s\Big| = O(n^{-\frac{1}{3}} \log n), \quad a.s.,\ n \to \infty.$$

Lemma 4.6 (see [36]) *Let $\{e_i, i \geq 1\}$ be a strongly mixing sequence with $Ee_i = 0$ and $\sup_{i\geq 1} E|e_i|^p < \infty$ for some $p > 2$. Assume that $\sum_{n=1}^\infty \Big(\sum_{i=1}^\infty a_{ni}^2 \log n\Big)^{\frac{p}{2}} < \infty$ and $\sum_{n=1}^\infty \alpha(n)^{\frac{p-2}{p}} < \infty$. Then*

$$\sum_{i=1}^\infty a_{ni} e_i = o(1), \quad a.s.,\ n \to \infty.$$

where $\alpha(n)$ are their mixing coefficients, and $\{a_{ni}, i = 1, 2, \cdots\}$ is a real number sequence.

Lemma 4.7 (see [37]) *Let $\{b_n, n \geq 1\}$ be a sequence of positive nondecreasing real numbers and let $\{e_i, i \geq 1\}$ be NA random variables with $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{b_n^2} < \infty$, where $\sigma_n^2 = \mathrm{Var}(e_n)$. Assume that $0 < b_n \uparrow \infty$. Then*

$$\sum_{i=1}^{n} \frac{e_i}{b_n} = o(1), \quad a.s.,\ n \to \infty.$$

Proof of Theorem 3.1 Let $\widetilde{\varepsilon} = (I - S)\varepsilon = (I - S)\Sigma e$, $e = (e_1, e_2, \cdots, e_n)^{\mathrm{T}}$, $g = (g(t_1), \cdots, g(t_n))^{\mathrm{T}}$, and $\widetilde{g} = (I - S)g$. Then

$$\widehat{\beta}_n - \beta = (n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{X})^{-1}(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g} + n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{\varepsilon}). \tag{4.1}$$

By the proof of Theorem 2.1 in [24], we obtain that

$$n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{X} \xrightarrow{\text{a.s.}} V \quad (n \to \infty). \tag{4.2}$$

Next we shall prove that

$$(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g})_i = O(n^{-\frac{1}{3}} \log n), \quad \text{a.s.},\ n \to \infty. \tag{4.3}$$

where $(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g})_i$ is the ith component of $n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g}$. In fact,

$$\begin{aligned}(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g})_i &= \frac{1}{n}\sum_{h=1}^{n} \eta_{hi}\Big(g(t_h) - \sum_{r=1}^{n} S_{hr}g(t_r)\Big) \\ &\quad - \frac{1}{n}\sum_{h=1}^{n}\sum_{k=1}^{n} S_{hk}\eta_{ki}\Big(g(t_h) - \sum_{r=1}^{n} S_{hr}g(t_r)\Big) \\ &\quad + \frac{1}{n}\sum_{h=1}^{n}\Big(f_i(t_h) - \sum_{k=1}^{n} S_{hk}f_i(t_k)\Big)\Big(g(t_h) - \sum_{r=1}^{n} S_{hr}g(t_r)\Big) \\ &=: \mathrm{J}_1 + \mathrm{J}_2 + \mathrm{J}_3. \end{aligned} \tag{4.4}$$

By using Markov inequality and Lemma 4.2, we obtain that

$$\sum_{n=1}^{\infty} P(|\mathrm{J}_1| \geq n^{-\frac{1}{3}} \log n) \leq \sum_{n=1}^{\infty} \frac{E\mathrm{J}_1^2}{(n^{-\frac{1}{3}} \log n)^2} \leq C\sum_{n=1}^{\infty} n^{-1}(\log n)^{-2} < \infty.$$

Hence, by Borel-Cantelli lemma,

$$\mathrm{J}_1 = O(n^{-\frac{1}{3}} \log n), \quad \text{a.s.}, n \to \infty. \tag{4.5}$$

By Lemma 4.2 and Lemma 4.5, it is easy to obtain

$$\mathrm{J}_2 = o(n^{-\frac{1}{3}} \log n), \quad \text{a.s.}, n \to \infty. \tag{4.6}$$

By Lemma 4.2, we obtain

$$\mathrm{J}_3 = O(n^{-2\gamma}) + O(\tau_m^2) = o(n^{-\frac{1}{3}} \log n), \quad \text{a.s.}, n \to \infty. \tag{4.7}$$

Hence, (4.3) follows from (4.5)–(4.7).

In the end, we shall show

$$(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{\varepsilon})_i = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n \to \infty. \tag{4.8}$$

In fact,

$$\begin{aligned}(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{\varepsilon})_i &= \frac{1}{n}\sum_{h=1}^{n}\Big(\eta_{hi} - \sum_{k=1}^{n} S_{hk}\eta_{ki}\Big)\Big(\varepsilon_h - \sum_{r=1}^{n} S_{hk}\varepsilon_r\Big)\\ &\quad + \frac{1}{n}\sum_{h=1}^{n}\Big(f_i(t_h) - \sum_{k=1}^{n} S_{hk}f_i(t_k)\Big)\Big(\varepsilon_h - \sum_{r=1}^{n} S_{hk}\varepsilon_r\Big)\\ &=: \mathrm{T}_1 + \mathrm{T}_2.\end{aligned} \tag{4.9}$$

Write

$$\begin{aligned}\mathrm{T}_1 &= \frac{1}{n}\sum_{h=1}^{n}\Big(\eta_{hi}\varepsilon_h - \Big(\sum_{k=1}^{n} S_{hk}\eta_{ki}\Big)\varepsilon_h - \Big(\sum_{r=1}^{n} S_{hr}\varepsilon_r\Big)\eta_{hi}\Big)\\ &\quad + \frac{1}{n}\sum_{h=1}^{n}\Big(\sum_{k=1}^{n} S_{hk}\eta_{ki}\Big)\Big(\sum_{r=1}^{n} S_{hr}\varepsilon_r\Big)\\ &=: \mathrm{T}_1^{(1)} - \mathrm{T}_1^{(2)} - \mathrm{T}_1^{(3)} + \mathrm{T}_1^{(4)}.\end{aligned} \tag{4.10}$$

Note that $\{\eta_{hi}\}$ and $\{\varepsilon_h\}$ are independent of each other, $E\overline{\eta}_i = 0$ and $Ee_i = 0$. It is easy to show that $\mathrm{Cov}(\eta_{hi}e_h, \eta_{ki}e_k) = 0$. Thus $\{\eta_{hi}e_h, h = 1, 2, \cdots, n\}$ is a ρ-mixing random sequence, which implies an α-mixing sequence (note that $0 \le \alpha(n) \le \frac{1}{4}\rho(n) = 0$, see [2]).

Let $a_{ni} = n^{-\frac{2}{3}}\log^{-1} n$. By Lemma 4.6 and $\alpha(n) = 0$, we can easily obtain

$$\mathrm{T}_1^{(1)} = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n \to \infty. \tag{4.11}$$

It is easy to show that $\{e_i, i = 1, 2, \cdots, n\}$ being NA implies $\{e_i^+, i = 1, 2, \cdots, n\}$ and $\{e_i^-, i = 1, 2, \cdots, n\}$ being NA also. By Lemma 4.7, we obtain that $n^{-1}\sum\limits_{h=1}^{n} e_h^+ = o(1)$, a.s., $n \to \infty$, and $n^{-1}\sum\limits_{h=1}^{n} e_h^- = o(1)$, a.s., $n \to \infty$. Therefore, by $|e_i| = e_i^+ + e_i^-$,

$$n^{-1}\sum_{h=1}^{n}|e_h| = o(1), \quad \text{a.s., } n \to \infty. \tag{4.12}$$

By Lemma 4.5 and (4.12), we obtain

$$|\mathrm{T}_1^{(2)}| \le \max_h\Big|\sum_{k=1}^{n} S_{hk}\eta_{ki}\Big| \cdot \max_h|h^{\frac{1}{2}}(u_h)| \cdot \Big(\frac{1}{n}\sum_{k=1}^{n}|e_h|\Big) = o(n^{-1}\log n), \quad \text{a.s., } n \to \infty. \tag{4.13}$$

By Lemma 4.4 and the strong law of large numbers, we obtain

$$|\mathrm{T}_1^{(3)}| \le \max_h\Big|\sum_{r=1}^{n} S_{hr}e_r\Big| \cdot \max_h|h^{\frac{1}{2}}(u_h)| \cdot \Big(\frac{1}{n}\sum_{h=1}^{n}|\eta_{hi}|\Big) = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n \to \infty. \tag{4.14}$$

By Lemma 4.4 and Lemma 4.5, we obtain

$$|\mathrm{T}_1^{(4)}| \le \max_h \Big| \Big(\sum_{k=1}^n S_{hk}\eta_{ki}\Big) \cdot \max_h |h^{\frac{1}{2}}(u_h)| \cdot \Big(\sum_{r=1}^n S_{hr}e_r\Big)\Big| = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty. \tag{4.15}$$

Hence, from (4.10)–(4.15), we obtain

$$\mathrm{T}_1 = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty. \tag{4.16}$$

By Lemma 4.2, (4.12) and Lemma 4.4, we obtain that

$$\begin{aligned}|\mathrm{T}_2| &\le \max_h \Big| f_i(t_h) - \sum_{k=1}^n S_{hk} f_i(t_k)\Big| \Big\{ \max_h |h^{\frac{1}{2}}(u_h)| \cdot \frac{1}{n}\sum_{h=1}^n |e_h| + \max_r |h^{\frac{1}{2}}(u_r)| \sum_{r=1}^n S_{hr}e_r \Big\} \\ &\le Cn^{-\frac{1}{3}}.\end{aligned} \tag{4.17}$$

Thus, (4.8) follows from (4.9), (4.16)–(4.17).

By (4.1), we obtain

$$\sup_{1\le i\le d} |\widehat{\beta}_{ni} - \beta_i| \le d \cdot \sup_{1\le j\le d} \Big|(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{X})^{-1}_{ij}\Big| \sup_{1\le j\le d} \Big|(n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{g})_j + (n^{-1}\widetilde{X}^{\mathrm{T}}\widetilde{\varepsilon})_j\Big|. \tag{4.18}$$

Hence, Theorem 3.1 follows from (4.2)–(4.3), (4.8) and (4.18).

Proof of Theorem 3.2 Note that

$$\begin{aligned}\sup_t |\widehat{g}_n(t) - g(t)| &\le \sup_t |\widehat{g}_0(t,\beta) - g(t)| + \sup_t \Big| \sum_{j=1}^n X_j^{\mathrm{T}}(\beta - \widehat{\beta}_n) \int_{A_j} E_m(t,s)\mathrm{d}s \Big| \\ &\le \sup_t \Big| \sum_{j=1}^n g(t_j) \int_{A_j} E_m(t,s)\mathrm{d}s - g(t)\Big| + \sup_t \Big| \sum_{j=1}^n \varepsilon_j \int_{A_j} E_m(t,s)\mathrm{d}s\Big| \\ &\quad + \sum_{j=1}^d \Big(|\widehat{\beta}_{nj} - \beta_j| \sup_t \Big| \sum_{i=1}^n f_j(t_i) \int_{A_i} E_m(t,s)\mathrm{d}s\Big|\Big) \\ &\quad + \sum_{j=1}^d \Big(|\widehat{\beta}_{nj} - \beta_j| \sup_t \Big| \sum_{i=1}^n \eta_{ij} \int_{A_i} E_m(t,s)\mathrm{d}s\Big|\Big) \\ &=: \mathrm{K}_1 + \mathrm{K}_2 + \mathrm{K}_3 + \mathrm{K}_4.\end{aligned} \tag{4.19}$$

By Lemma 4.2, we obtain

$$\mathrm{K}_1 = O(n^{-\gamma}) + O(\tau_m) = o(n^{-\frac{1}{3}}\log n). \tag{4.20}$$

By Lemma 4.4, we obtain

$$\mathrm{K}_2 = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty. \tag{4.21}$$

Let $M = \sup_{j,t_i} |f_j(t_i)|$. Then by Lemma 4.1 and Theorem 3.1, we obtain

$$\mathrm{K}_3 \le d \cdot \sup_{j,t_i} |f_j(t_i)| \cdot \sup_t \int_0^1 |E_m(t,s)|\mathrm{d}s \cdot \max_j |\widehat{\beta}_{nj} - \beta_j|$$

$$\leq C \max_{1\leq j\leq d}|\widehat{\beta}_{nj}-\beta_j| = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty. \tag{4.22}$$

By Lemma 4.5 and Theorem 3.1, we obtain

$$\begin{aligned}
\mathrm{K}_4 &\leq d\sup_{1\leq j\leq d}|\widehat{\beta}_{nj}-\beta_j|\max_{1\leq j\leq d}\sup_t\Big|\sum_{j=1}^n \eta_{ij}\int_{A_i}E_m(t,s)\mathrm{d}s\Big| \\
&= o(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty.
\end{aligned}\tag{4.23}$$

From (4.19)–(4.23), the desired conclusion follows.

Proof of Theorem 3.3 By (2.5) and $\widetilde{y}_i = \widetilde{X}_i^{\mathrm{T}}\beta + \widetilde{g}(t_i) + \widetilde{\varepsilon}_i$, we obtain that

$$\begin{aligned}
\widehat{h}_n(u) &= \sum_{i=1}^n (\widetilde{X}_i^{\mathrm{T}}(\beta-\widehat{\beta}_n))^2\int_{A_i}E_m(u,s)\mathrm{d}s + \sum_{i=1}^n \widetilde{g}^2(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s \\
&\quad + \sum_{i=1}^n \widetilde{\varepsilon}_i^2\int_{A_i}E_m(u,s)\mathrm{d}s + 2\sum_{i=1}^n \widetilde{X}_i^{\mathrm{T}}(\beta-\widehat{\beta}_n)\widetilde{g}(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s \\
&\quad + 2\sum_{i=1}^n \widetilde{X}_i^{\mathrm{T}}(\beta-\widehat{\beta}_n)\widetilde{\varepsilon}_i\int_{A_i}E_m(u,s)\mathrm{d}s + 2\sum_{i=1}^n \widetilde{g}(t_i)\widetilde{\varepsilon}_i\int_{A_i}E_m(u,s)\mathrm{d}s \\
&=: \mathrm{I}_1+\mathrm{I}_2+\mathrm{I}_3+2\mathrm{I}_4+2\mathrm{I}_5+2\mathrm{I}_6.
\end{aligned}\tag{4.24}$$

By Theorem 3.1, Lemma 4.1, Lemma 4.2 and Lemma 4.5,

$$\begin{aligned}
|\mathrm{I}_1| &= \Big|\sum_{i=1}^n\Big(\sum_{r=1}^d \widetilde{x}_{ir}(\beta_r-\widehat{\beta}_{nr})\Big)^2\int_{A_i}E_m(u,s)\mathrm{d}s\Big| \\
&\leq \max_{1\leq r\leq d}|\beta_r-\widehat{\beta}_{nr}|^2\Big|\sum_{i=1}^n\Big(\sum_{r=1}^d(\widetilde{f}_r(t_i)+\widetilde{\eta}_{ir})\Big)^2\int_{A_i}E_m(u,s)\mathrm{d}s\Big| \\
&\leq C\max_{1\leq r\leq d}|\beta_r-\widehat{\beta}_{nr}|^2\Big|\sum_{i=1}^n\sum_{r=1}^d(\widetilde{f}_r^2(t_i)+\widetilde{\eta}_{ir}^2)\int_{A_i}E_m(u,s)\mathrm{d}s\Big| \\
&\leq Cn^{-\frac{2}{3}}\log^2 n\Big(Cn^{-\frac{1}{3}}+\sum_{i=1}^n\sum_{r=1}^d\widetilde{\eta}_{ir}^2\int_{A_i}E_m(u,s)\mathrm{d}s\Big) \\
&= o(n^{-\frac{1}{3}}\log n), \quad \text{a.s., } n\to\infty.
\end{aligned}\tag{4.25}$$

By Lemma 4.1 and Lemma 4.2, we have that

$$\begin{aligned}
\sup_{0\leq u\leq 1}|\mathrm{I}_2| &\leq \sup_{0\leq u\leq 1}\sum_{i=1}^n\Big|\int_{A_i}E_m(u,s)\mathrm{d}s\Big|\max_{1\leq i\leq n}|\widetilde{g}^2(t_i)| \\
&= O(n^{-2\gamma})+O(\tau_m^2) = O(n^{-\frac{2}{3}}), \quad \text{a.s., } n\to\infty.
\end{aligned}\tag{4.26}$$

In the following, we shall prove that

$$\sup_{0\leq u\leq 1}|\mathrm{I}_3-h(u)| = O(n^{-\frac{1}{3}}\log n), \quad \text{a.s.,, } n\to\infty. \tag{4.27}$$

By Cauchy inequality and Lemma 4.1, we obtain that

$$\sup_{0\leq u\leq 1}|\mathrm{I}_3-h(u)|$$

$$
\begin{aligned}
&\le \sup_{0\le u\le 1}\Big|\sum_{i=1}^n h(u_i)(e_i^2-1)\int_{A_i}E_m(u,s)\mathrm{d}s\Big|\\
&\quad+\sup_{0\le u\le 1}\Big|\sum_{i=1}^n h(u_i)\int_{A_i}E_m(u,s)\mathrm{d}s-h(u)\Big|\cdot\max_{1\le i\le n}\Big(\sum_{j=1}^n h^{\frac12}(u_j)e_j\int_{A_j}E_m(t_i,s)\mathrm{d}s\Big)^2\\
&\quad+2\sup_{0\le u\le 1}\Big|\Big\{\sum_{i=1}^n h(u_i)(e_i^2-1)\int_{A_i}E_m(u,s)\mathrm{d}s+\sum_{i=1}^n h(u_i)\int_{A_i}E_m(u,s)\mathrm{d}s\Big\}^{\frac12}\Big|\\
&\quad\cdot\sup_{0\le u\le 1}\Big|\Big\{\sum_{i=1}^n\int_{A_i}E_m(u,s)\mathrm{d}s\Big(\sum_{j=1}^n h^{\frac12}(u_j)e_j\int_{A_j}E_m(t_i,s)\mathrm{d}s\Big)^2\Big\}^{\frac12}\Big|\\
&=:\sup_{0\le u\le 1}|\mathrm{I}_{31}|+\sup_{0\le u\le 1}|\mathrm{I}_{32}|+\max_{1\le i\le n}|\mathrm{I}_{33}|\\
&\quad+2\sup_{0\le u\le 1}\{\mathrm{I}_{31}+\mathrm{I}_{34}\}^{\frac12}\sup_{0\le u\le 1}\Big\{\sum_{i=1}^n\int_{A_i}E_m(u,s)\mathrm{d}s|\mathrm{I}_{33}|\Big\}^{\frac12}. \qquad (4.28)
\end{aligned}
$$

Let $\xi_i=e_i^2-1=(e_i^+)^2-E(e_i^+)^2-((e_i^-)^2-E(e_i^-)^2)=\xi_i^+-\xi_i^-$. Then $\{\xi_i^+,i\ge 1\}$ and $\{\xi_i^-,i\ge 1\}$ are NA with $E\xi_i^\pm=0$ and $\mathrm{Var}(\xi_i^\pm)<\infty$. By Lemma 4.5, we obtain that

$$\sup_{0\le u\le 1}|\mathrm{I}_{31}|=O(n^{-\frac13}\log n),\quad \text{a.s., } n\to\infty. \qquad (4.29)$$

By Lemma 4.2, we have

$$\sup_{0\le u\le 1}|\mathrm{I}_{32}|=O(n^{-\gamma})+O(\tau_m)=O(n^{-\frac13}),\quad \text{a.s., } n\to\infty. \qquad (4.30)$$

By Lemma 4.4, we have

$$\max_{0\le u\le 1}|\mathrm{I}_{33}|=O(n^{-\frac23}\log^2 n),\quad \text{a.s., } n\to\infty. \qquad (4.31)$$

By Lemma 4.1, we obtain that

$$|\mathrm{I}_{34}|\le\max_{1\le i\le n}|h(u_i)|\int_0^1|E_m(u,s)|\mathrm{d}s\le C. \qquad (4.32)$$

Hence, (4.27) follows from (4.28)–(4.32).

By Cauchy inequality, (4.25) and (4.26), we obtain that

$$\mathrm{I}_4^2\le\sum_{i=1}^n(\widetilde{X}_i^{\mathrm{T}}(\beta-\widehat{\beta}_n))^2\int_{A_i}E_m(u,s)\mathrm{d}s\cdot\sum_{i=1}^n\widetilde{g}^2(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s=\mathrm{I}_1\cdot\mathrm{I}_2.$$

Hence,

$$|\mathrm{I}_4|=o(n^{-\frac13}\log n),\quad \text{a.s., } n\to\infty. \qquad (4.33)$$

By Cauchy inequality, (4.25) and (4.27), we obtain that

$$\mathrm{I}_5^2\le\sum_{i=1}^n(\widetilde{X}_i^{\mathrm{T}}(\beta-\widehat{\beta}_n))^2\int_{A_i}E_m(u,s)\mathrm{d}s\cdot\sum_{i=1}^n\widetilde{\varepsilon}_i^2\int_{A_i}E_m(u,s)\mathrm{d}s=\mathrm{I}_1\cdot\mathrm{I}_3.$$

Thus

$$|\mathrm{I}_5|=o(n^{-\frac13}\log n),\quad \text{a.s., } n\to\infty. \qquad (4.34)$$

By Cauchy inequality, (4.26) and (4.27), we obtain that

$$\begin{aligned}\sup_{0\le u\le 1}|\mathrm{I}_6| &\le \sup_{0\le u\le 1}\Big|\Big(\sum_{i=1}^n \widetilde{\varepsilon}_i^2\int_{A_i}E_m(u,s)\mathrm{d}s\Big)^{\frac12}\Big(\sum_{i=1}^n \widetilde{g}^2(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s\Big)^{\frac12}\Big|\\ &\le \sup_{0\le u\le 1}|\mathrm{I}_2^{\frac12}\mathrm{I}_3^{\frac12}| = o(n^{-\frac13}\log n),\quad \text{a.s.}, n\to\infty. \end{aligned}\tag{4.35}$$

Therefore, Theorem 3.3 follows from (4.24)–(4.27) and (4.33)–(4.35).

Proof of Theorem 3.4 Note (A_5) and

$$\widetilde{\beta}_n-\beta=(n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}(n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{g}+n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{\varepsilon}),$$

the proof of Theorem 3.4 is analogous with the proof of Theorem 3.1.

Proof of Theorem 3.5 Note that

$$\begin{aligned}&\sup_t|\widetilde{g}_n(t)-g(t)|\\ &\le \sup_t\Big|\sum_{i=1}^n(y_i-X_i^{\mathrm{T}}\beta)\int_{A_i}E_m(t,s)\mathrm{d}s-g(t)\Big|+\sup_t\Big|\sum_{i=1}^n X_i^{\mathrm{T}}(\beta-\widetilde{\beta}_n)\int_{A_i}E_m(u,s)\mathrm{d}s\Big|\\ &\le \sup_t\Big|\sum_{i=1}^n g(t_i)\int_{A_i}E_m(t,s)\mathrm{d}s-g(t)\Big|+\sup_t\Big|\sum_{i=1}^n h^{\frac12}(u_i)e_i\int_{A_i}E_m(u,s)\mathrm{d}s\Big|\\ &\quad+\sum_{j=1}^d\Big(|\widetilde{\beta}_{nj}-\beta_j|\sup_t\Big|\sum_{i=1}^n f_j(t_i)\int_{A_i}E_m(t,s)\mathrm{d}s\Big|\Big)\\ &\quad+\sum_{j=1}^d\Big(|\widetilde{\beta}_{nj}-\beta_j|\sup_t\Big|\sum_{i=1}^n \eta_{ij}\int_{A_i}E_m(t,s)\mathrm{d}s\Big|\Big).\end{aligned}$$

From Lemma 4.1, Lemma 4.2, Lemma 4.4, Lemma 4.5 and Theorem 3.4, the desired conclusion follows.

Proof of Theorem 3.6 Note that

$$\begin{aligned}\widetilde{h}_n(u)&=\sum_{i=1}^n(\widetilde{X}_i^{\mathrm{T}}(\beta-\widetilde{\beta}_n))^2\int_{A_i}E_m(u,s)\mathrm{d}s+\sum_{i=1}^n\widetilde{g}^2(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s\\ &\quad+\sum_{i=1}^n\widetilde{\varepsilon}_i^2\int_{A_i}E_m(u,s)\mathrm{d}s+2\sum_{i=1}^n\widetilde{X}_i^{\mathrm{T}}(\beta-\widetilde{\beta}_n)\widetilde{g}(t_i)\int_{A_i}E_m(u,s)\mathrm{d}s\\ &\quad+2\sum_{i=1}^n\widetilde{X}_i^{\mathrm{T}}(\beta-\widetilde{\beta}_n)\widetilde{\varepsilon}_i\int_{A_i}E_m(u,s)\mathrm{d}s+2\sum_{i=1}^n\widetilde{g}(t_i)\widetilde{\varepsilon}_i\int_{A_i}E_m(u,s)\mathrm{d}s.\end{aligned}$$

The proof is similar to that of Theorem 3.3, and hence we omit it here.

Proof of Theorem 3.7 From Theorem 3.4, we only prove that

$$\widetilde{\beta}_n-\overline{\beta}_n=O(n^{-\frac13}\log n),\quad \text{a.s.}, n\to\infty.\tag{4.36}$$

In fact, by (2.6) and (2.9), we obtain that

$$\widetilde{\beta}_n-\overline{\beta}_n=(\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}(\widetilde{g}+\widetilde{\varepsilon})-(\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}(\widetilde{g}+\widetilde{\varepsilon})$$

$$
\begin{aligned}
&= (\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}\widetilde{X}^{\mathrm{T}}(\Sigma^{-1}-\widehat{\Sigma}^{-1})(\widetilde{g}+\widetilde{\varepsilon}) \\
&\quad + ((\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}-(\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X})^{-1})\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}(\widetilde{g}+\widetilde{\varepsilon}) \\
&=: \mathrm{I}_1+\mathrm{I}_2+\mathrm{I}_3+\mathrm{I}_4.
\end{aligned} \tag{4.37}
$$

We first prove that

$$
n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X} \xrightarrow{\text{a.s.}} \Sigma_1 \quad (n\to\infty) \tag{4.38}
$$

and

$$
n^{-1}\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X} \xrightarrow{\text{a.s.}} \Sigma_2 \quad (n\to\infty), \tag{4.39}
$$

where $(\Sigma_1)_{ij} = \lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{h=1}^{n} h^{-1}(u_h)\eta_{hi}\eta_{hj}$, $(\Sigma_2)_{ij} = \lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{h=1}^{n}\widehat{h}_n^{-1}(u_h)\eta_{hi}\eta_{hj}$. The two limits exist because $\{\eta_{hi}\eta_{hj}, h=1,2,\cdots,n\}$ are independent random variables, and $\{h^{-1}(u_h), h\geq 1\}$ are bound and $\widehat{h}_n^{-1}(u_h)\to h^{-1}(u_h)$.

$$
\begin{aligned}
n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X} &= \frac{1}{n}\sum_{h=1}^{n} h^{-1}(u_h)\Big(f_i(t_h)-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}sf_i(t_k)\Big)\Big) \\
&\quad \cdot\Big(\eta_{hj}-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}s\Big)\eta_{kj}\Big) \\
&\quad + \frac{1}{n}\sum_{h=1}^{n} h^{-1}(u_h)\Big(f_j(t_h)-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}sf_j(t_k)\Big)\Big) \\
&\quad \cdot\Big(\eta_{hi}-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}s\Big)\eta_{ki}\Big) \\
&\quad + \frac{1}{n}\sum_{h=1}^{n} h^{-1}(u_h)\Big(f_i(t_h)-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}sf_i(t_k)\Big)\Big) \\
&\quad \cdot\Big(f_j(t_h)-\sum_{k=1}^{n}\Big(\int_{A_k}E_m(t_h,s)\mathrm{d}s\Big)f_j(t_k)\Big) \\
&\quad + \frac{1}{n}\sum_{h=1}^{n} h^{-1}(u_h)\Big(\eta_{hi}-\sum_{k=1}^{n}\int_{A_k}E_m(t_h,s)\mathrm{d}s\eta_{ki}\Big) \\
&\quad \cdot\Big(\eta_{hj}-\sum_{k=1}^{n}\int_{A_k}E_m(t_h,s)\mathrm{d}s\eta_{kj}\Big) \\
&=: \mathrm{U}_1+\mathrm{U}_2+\mathrm{U}_3+\mathrm{U}_4.
\end{aligned} \tag{4.40}
$$

It is easy to show that

$$
\begin{aligned}
&\mathrm{U}_1 = O(n^{-\gamma})+O(\tau_m), \quad \mathrm{U}_2 = O(n^{-\gamma})+O(\tau_m), \\
&\mathrm{U}_3 = O(n^{-2\gamma})+O(\tau_m^2), \quad \text{a.s.},\ n\to\infty.
\end{aligned} \tag{4.41}
$$

By Lemma 4.5, we obtain that

$$
\mathrm{U}_4 = \frac{1}{n}\sum_{h=1}^{n} h^{-1}(u_h)\eta_{hi}\eta_{hj}+o(1)\to(\Sigma_1)_{ij}, \quad \text{a.s.},\ n\to\infty. \tag{4.42}
$$

Then (4.38) follows from (4.40)–(4.42). With arguments in a way similar to (4.38), we get (4.39).

By (4.3) and Theorem 3.3, we obtain that

$$\begin{aligned}(n^{-1}\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X}\mathrm{I}_1)_i &= n^{-1}\sum_{k=1}^{n}\widetilde{x}_{ki}(h^{-1}(u_k)-\widehat{h}_n^{-1}(u_k))\widetilde{g}(t_k)\\ &= o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \end{aligned}\tag{4.43}$$

By (4.8) and Theorem 3.3, we obtain that

$$(n^{-1}(\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})\mathrm{I}_2)_i = (n^{-1}\widetilde{X}^{\mathrm{T}}(\Sigma^{-1}-\widehat{\Sigma}^{-1})\widetilde{\varepsilon})_i = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \tag{4.44}$$

Note that $\widehat{h}_n(u) \xrightarrow{\text{a.s.}} h(u)$ $(n\to\infty)$, and with arguments similar to (4.43) and (4.44), we get

$$\mathrm{I}_3 = ((\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}-(\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X})^{-1})\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{g} = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \tag{4.45}$$

$$\mathrm{I}_4 = ((\widetilde{X}^{\mathrm{T}}\Sigma^{-1}\widetilde{X})^{-1}-(\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{X})^{-1})\widetilde{X}^{\mathrm{T}}\widehat{\Sigma}^{-1}\widetilde{\varepsilon} = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \tag{4.46}$$

The desired conclusion follows from (4.37)–(4.39) and (4.43)–(4.46).

Proof of Theorem 3.8 From Theorem 3.5, we only prove that

$$\sup_{0\le t\le 1}|\overline{g}_n(t)-\widetilde{g}_n(t)| = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \tag{4.47}$$

In fact, by (2.7), (2.10), (4.25) and (4.36), we obtain that

$$\begin{aligned}\sup_{0\le t\le 1}|\overline{g}_n(t)-\widetilde{g}_n(t)| &\le C\sup_{0\le t\le 1}\sum_{i=1}^{n}\Big|X_i^{\mathrm{T}}(\widetilde{\beta}_n-\overline{\beta}_n)\int_{A_i}E_m(t,s)\mathrm{d}s\Big|\\ &= o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty.\end{aligned}$$

Proof of Theorem 3.9 From Theorem 3.6, we only prove that

$$\sup_{0\le u\le 1}|\overline{h}_n(u)-\widetilde{h}_n(u)| = o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty. \tag{4.48}$$

In fact, by (2.8) and (2.11), we obtain that

$$\begin{aligned}&\sup_{0\le u\le 1}|\overline{h}_n(u)-\widetilde{h}_n(u)|\\ &= \sup_{0\le u\le 1}\Big|\sum_{i=1}^{n}(2\widetilde{y}_i-\widetilde{X}_i^{\mathrm{T}}(\overline{\beta}_n-\widetilde{\beta}_n))(\widetilde{X}_i^{\mathrm{T}}(\widetilde{\beta}_n-\overline{\beta}_n))\int_{A_i}E_m(u,s)\mathrm{d}s\Big|\\ &= o(n^{-\frac{1}{3}}\log n), \quad \text{a.s.}, \ n\to\infty.\end{aligned}$$

Acknowledgement The authors wish to thank the referee for some comments and suggestions which helped to improve the presentation of this paper.

References

[1] Speckman, P., Kernel smoothing in partial linear models, *J. R. Statist. Soc. B*, **50**, 1988, 413–436.

[2] Fan, J. and Yao, Q., Nonlinear Time Series: Nonparametric and Parametric Methods, Springer-Verlag, New York, 2003.

[3] Alam, K. and Saxena, K. M. L., Positive dependence in multivariate distributions, *Comm. Statist. Theory Methods Ser. A*, **10**, 1981, 1183–1196.

[4] Joag-Dev, K. and Proschan, F., Negative association of random variables with applications, *Ann. Statist.*, **11**, 1983, 286–295.

[5] Matula, P., A note on the almost sure convergence of sums of negatively dependences random variables, *Statist. Probab. Lett.*, **15**, 1992, 209–213.

[6] Shao, Q. M. and Su, C., The law of the iterated logarithm for negatively associated random variables, *Stochastic Process Appl.*, **83**, 1999, 139–148.

[7] Wu, Q. and Jiang, Y., A law of the iterated logarithm of partial sums for NA random variables, *J. Korean Statist. Soc.*, **39**, 2010, 199–206.

[8] Liang, H. Y., Complete convergence for weighted sums of negatively associated random variables, *Statist. Probab. Lett.*, **48**, 2000, 317–325.

[9] Baek, J. I., Kim, T. S. and Liang, H. Y., On the convergence of moving average processes under dependent conditions, *Austral. New Zealand J. Statist.*, **45**(3), 2003, 331–342.

[10] Liang, H. and Zhang, J., Strong convergence for weighted sums of negatively associated arrays, *Chin. Ann. Math.*, **31B**(2), 2010, 273–288.

[11] Su, C., Zhao, L. and Wang, Y., The moment inequality for NA random variables and the weak convergence (in Chinese), *Sci. China Ser. A*, **26**(12), 1996, 1091–1099.

[12] Jin, J., Convergence rate in the law of logarithm for NA random fields (in Chinese), *Acta Math. Sciencia Ser. A*, **29**(4), 2009, 1138–1143.

[13] Roussas, G. G., Asymptotic normality of random fields of positively or negatively associated processes, *J. Multivariate Anal.*, **50**, 1994, 152–173.

[14] Chen, H., Convergence rates for parametric components in a partly linear model, *Ann. Statist.*, **16**(1), 1988, 136–146.

[15] Qian, W. M. and Cai, G. X., Strong approximability of wavelet estimate in semiparametric regression model (in Chinese), *Sci. China Ser. A*, **29**(3), 1999, 233–240.

[16] Xue, L. and Liu, Q., Bootstrap approximation of wavelet estimates in a semiparametric regression model, *Acta Math. Sin.* (*Engl. Ser.*), **26**(4), 2010, 763–778.

[17] Chai, G. X. and Xu, K. J., Wavelet smoothing in semiparametric regression model (in Chinese), *Chinese J. Appl. Probab. Statist.*, **15**(1), 1999, 97–105.

[18] Hu, H. C., Ridge estimation of a semiparametric regression model, *J. Comput. and Appl. Math.*, **176**, 2005, 215–222.

[19] Heckman, N., Spline smoothing in partly linear models, *J. R. Statist. Soc. B*, **48**, 1986, 244–248.

[20] Green, P. J. and Silverman, B. W., Nonparametric Regression and Generalized Linear Models, Chapman and Hall, London, 1995.

[21] Härdle, W., Liang, H. and Gao, J., Partially Linear Models, Heidelberg, Physica-Verlag, 2000.

[22] Shi, P. and Li, G., A note on the convergent rates of M-estimates for a partly linear model, *Statistics*, **26**, 1995, 27–47.

[23] Bianco, A. and Boente, G., Robust estimators in semiparametric partly linear regression models, *J. Statist. Plann. Inference*, **122**, 2004, 229–252.

[24] Hu, H. C. and Hu, D. H., Strong consistency of wavelet estimation in semiparametric regression models (in Chinese), *Acta Math. Sinca*, **49**(5), 2006, 1417–1424.

[25] Hu, H. C. and Hu, D. H., Wavelet estimation in semiparametric regression models with martingale differences time series errors (in Chinese), *Adv. Math.*, **39**(1), 2010, 23–30.

[26] You, J. and Chen, G., Semiparametric generalized least squares estimation in partially linear regression models with correlated errors, *J. Statist. Plann. Inference*, **137**(1), 2007, 117–132.

[27] Liang, H. and Jing, B., Asymptotic normality in partial linear models based on dependent errors, *J. Statist. Plann. Inference*, **139**, 2009, 1357–1371.

[28] Zhou, X., Liu, X. and Hu, S., Moment consistency of estimators in partially linear models under NA samples, *Metrika*, 2009. DOI:10.10071/s00184-009-0260-5

[29] Baek, J. and Liang, H., Asymptotics of estimators in semi-parametric model under NA samples, *J. Statist. Plann. Inference*, **136**, 2006, 3362–3382.

[30] Liang, H. and Wang, X., Convergence rate of wavelet estimator in semiparametric models with dependent MA(∞) error process, *Chinese J. Appl. Probab. Statist.*, **26**(1), 2010, 35–46.

[31] Ren, Z. and Chen, M., Strong consistency of a class of estimators in partial linear model for negative associated samples (in Chinese), *Chinese J. Appl. Probab. Statist.*, **18**(1), 2002, 60–66.

[32] Zhou, X. and You, J., Wavelet estimation in varying-coefficient partially linear regression models, *Statist. Probab. Lett.*, **68**, 2004, 91–104.

[33] Antoniads, A., Gregoire, G. and Mckeague, I. W., Wavelet methods for curve estimation, *J. Amer. Statist. Assoc.*, **89**, 1994, 1340–1353.

[34] Zhu, C., The convergence rate of the estimate of the regression function under NA sequence, *J. Xinjiang Norm. Univ. Nat. Sci. Ed.*, **27**(2), 2008, 5–11.

[35] Hu, H. C., Convergence rates of wavelet estimators in semiparametric regression models with martingale difference error, *J. Math. Sci. Adv. Appl.*, **3**(1), 2009, 4–19.

[36] Xu, B., Convergence of the weighted sums with applications for strongly mixing sequence (in Chinese), *Acta Math. Sinica*, **45**(5), 2002, 1025–1034.

[37] Liu, J., Gan, S. and Chen, P., The Hjeck-Rnyi inequality for the NA random variables and its application, *Statist. Probab. Lett.*, **43**, 1999, 99–105.