This is to certify that the

dissertation entitled

Nonlinear Wavelet-Based Nonparametric Curve Estimation With Censored Data And Inference on Long Memory Processes

presented by

Linyuan Li

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Statistics

Major professor
Hira L. Koul

Date May 20, 2002

NONLINEAR WAVELET-BASED NONPARAMETRIC CURVE ESTIMATION WITH CENSORED DATA AND INFERENCE ON LONG MEMORY PROCESSES

By

Linyuan Li

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

2002

ABSTRACT

NONLINEAR WAVELET-BASED NONPARAMETRIC CURVE ESTIMATION WITH CENSORED DATA AND INFERENCE ON LONG MEMORY PROCESSES

By

Linyuan Li

In the first two parts of this thesis, we provide asymptotic formulaes for the mean integrated squared error (MISE) of nonlinear wavelet-based density and hazard rate estimators under randomly censored data. We show this MISE formula, when the underlying survival density and hazard rate functions and the censoring distribution function are only piecewise smooth, has the same expansion as analogous kernel density estimators. However, as to the kernel estimators, this MISE formula holds only under the smoothness assumption. In addition, we establish an asymptotic normality of non-linear wavelet estimator of hazard rate function, which is useful to construct a confidence interval of hazard rate function.

In the third part, we discusses the asymptotic behavior of Koul's minimum distance (m.d.) estimators of the regression parameter vector in linear regression models with long memory moving average errors, when the design variables are either known constants or i.i.d. random variables, independent of the errors. It is observed that all these estimators are asymptotically equivalent to the least squares estimator in the first order.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my dissertation advisor, Professor Hira L. Koul, for his constant guidance, generous support and extreme patience shown during the writing of this dissertation. His dedication and contribution to statistics have been my main source of inspiration during my graduate study. I am also very grateful to Professor Winfried Stute for his valuable and constructive suggestion and Professor Donatas Surgailis for providing me a preprint of Lemma 3.3.3 in this thesis.

I would also like to thank Professors LePage, Levental and Salehi for serving on my thesis committee. My special thanks go to Professor Ibragimov who offers wonderful courses during each summer and to Professor Page for training me as a statistical consultant.

The research in this thesis was also partly supported by the NSF Grant DMS 0071619 under the P.I. Professor Hira L. Koul.

TABLE OF CONTENTS

Chapter 1

Nonlinear Wavelet-based Density Estimator

1.1 Introduction

The mathematical theory of wavelets and their applications in statistics have become a well-known technique for non-parametric curve estimation: See e.g., Meyer (1990), Daubechies (1992), Chui (1992), Mallat (1989), Donoho and Johnstone (1994), Donoho, et al. (1995, 1996) and Kerkyacharian and Picard (1992, 1993). For a systematic discussion of wavelets and their applications see recent monograph by Härdle, et al. (1998). The major advantage of the wavelet method is its adaptation to erratic behavior of the density and local adaptation to the degree of smoothness of the unknown density. These wavelet estimators typically achieve the optimal convergence rates over exceptionally large function spaces. They do an excellent job of taking care of discontinuities in the target function, and in consequence they enjoy very good convergence rate even if smoothness conditions are imposed only in a piecewise sense.

Hall and Patil (1995) first explicitly demonstrated that, in the no censorship case, the discontinuities of densities have a negligible effect on the performance of non-linear wavelet density estimators. The mean integrated squared error (MISE) of the kernel estimator of density function f has the form

$$\text{MISE} \sim c_1(nh)^{-1} + c_2 h^{2r},$$

where "$\sim$" means that the ratio of the left- and right-hand sides converges to 1 as the sample size $n \to \infty$, h is the bandwidth of the kernel estimator, r is the order of the kernel and c_1 and c_2 are constants depending on both the kernel and unknown density. The first term derives from the variance and the second from the squared bias. This expansion for kernel estimators generally fails if the underlying density function does not have r derivatives (Hall and Patil, 1995, p.906). However, the MISE expansion of non-linear wavelet estimators is still valid for only piecewise smooth density function, and even has the same constants c_1 and c_2. Patil (1997) provided similar results for non-linear wavelet-based hazard rate estimator with complete data.

In industrial life-testing, medical follow-up research and other studies, the observation of the occurrence of the failure event may be prevented by the previous occurrence of the censoring event. So only part of observations are real failure times. Formally, let $X_1, X_2, \cdots, X_n$ be i.i.d. survival times with a common distribution function F and density function f. Also let $Y_1, Y_2, \cdots, Y_n$ be i.i.d. censoring times with a common distribution function G. It is assumed that X_i is independent of Y_i for every i. Rather than observing $X_1, X_2, \cdots, X_n$, the variables of interest, in the randomly right-censored models, one observes $Z_i = \min(X_i, Y_i) = X_i \wedge Y_i$ and $\delta_i = I(X_i \leq Y_i)$, $i = 1, 2, \cdots, n$, where $I(A)$ denotes the indicator function of the set A.

Antoniadis, et al. (1999) describe a wavelet method for the estimation of density and hazard rate functions from randomly right-censored data. The method is based on dividing the time axis into a dyadic number of intervals and then counting the number of events within each interval. The number of events and survival function of the observations are then separately smoothed over time via linear wavelet smoothers. They provide estimator's asymptotic normality and obtained best possible asymptotic MISE convergence rate under the assumption that survival time density function f is r-times continuously differentiable and the censoring density g is continuous.

The objective of this chapter is to propose a non-linear wavelet estimator of density function with censored data and derive a result similar to the main result, Theorem 2.1 of Hall and Patil (1995). One of the consequence of this extension is that we can show that MISE has the analogous expansion

$$\text{MISE} \sim k_1 n^{-1} p + k_2 p^{-2r}, \tag{1.1.1}$$

where n denotes the sample size, p is the smoothing parameter, a wavelet analogue of the bandwidth h^{-1} for kernel estimators and k_1 and k_2 are constants depending on the wavelet, unknown density and censoring distribution.

Recently Wu and Wells (1999) provided hazard rate estimation by non-linear wavelet methods in the left truncation and right censoring model. They have n observations (X_i, δ_i, V_i) with $X_i \geq V_i$, where $X_i = \min(T_i, U_i)$ and $\delta_i = I(T_i \leq U_i)$. They applied counting process techniques and obtained analogous MISE expansion, but needed further truncation. They provided a wavelet-based estimator for hazard rate function over bounded interval $[\iota, \tau]$ which is chosen such that the size of risk population satisfies the following conditions:

(Y1): $P(Y_{\min} \leq n\alpha) = o(n^{-2})$ for some $\alpha > 0$, where $Y_{\min} = \inf_{t\in[\iota,\tau]} Y(t)$ and $Y(t) = \sum_{i=1}^{n} I(X_i \geq t \geq V_i)$.

(Y2): $E\sup_{t\in[\iota,\tau]}|\frac{1}{Y(t)} - \frac{1}{nC(t)}| = o(n^{-1})$, where $C(s) = E|Y(s)/n|$.

Basically, the condition (Y1) means that the size of the risk population $Y(t)$ is large and the condition Y(2) means that $Y(t)$ is uniformly close to its expectation, for all $t \in [\iota, \tau]$. In addition, they only obtained the approximation (1.1.1) for the MISE, which is weaker than the result (1.3.1) given below.

In this thesis, we apply the method of Stute (1995) that approximates a Kaplan-Meier integral by an average of i.i.d. random variables with a sufficiently small rate. We provide a MISE expansion similar to that of Hall and Patil (1995) for density function over $(-\infty, T]$, for any fixed $T < \tau_H$, where $\tau_H = \inf\{x : H(x) = 1\} \leq \infty$ is the least upper bound for the support of H, the distribution function of Z_1.

In the next section, we give the elements of wavelet transform and provide nonlinear wavelet-based density estimators. The main results are described in Section 3, while their proofs appear in Sections 4 and 5.

1.2 Notations and Estimators

This section contains some facts about wavelets that will be used in the sequel. Let $\phi(x)$ and $\psi(x)$ be father and mother wavelets, having the properties: ϕ and ψ are bounded and compactly supported. $\int \phi^2 = \int \psi^2 = 1$, $\mu_k \equiv \int y^k \psi(y)\, dy = 0$ for $0 \le k \le r-1$ and $\mu_r = r!\kappa \ne 0$, where $\kappa = (r!)^{-1} \int y^r \psi(y)\, dy$. Let

$$\phi_j(x) = p^{1/2}\phi(px - j), \quad \psi_{ij}(x) = p_i^{1/2}\psi(p_i x - j), \quad x \in \mathbb{R}$$

for arbitrary $p > 0, -\infty < j < \infty$ and $p_i = p2^i, \quad i \ge 0$. Then

$$\int \phi_{j_1}\phi_{j_2} = \delta_{j_1 j_2}, \quad \int \psi_{i_1 j_1}\psi_{i_2 j_2} = \delta_{i_1 i_2}\delta_{j_1 j_2}, \quad \int \phi_{j_1}\psi_{i j_2} = 0,$$

where δ_{ij} denotes the Kronecker delta, i.e., $\delta_{ij} = 1$, if $i = j$; 0, otherwise. For the more on wavelets see Daubechies (1992).

In our random censorship model, we observe $Z_i = \min(X_i, Y_i)$, and $\delta_i = I(X_i \le Y_i)$, $i = 1, 2, \cdots, n$. Let $T < \tau_H$ be fixed and $f_1(x) = f(x)I(x \le T)$. We estimate $f_1(x)$, i.e. density function $f(x)$ for $x \in (-\infty, T]$. The wavelet expansion of $f_1(x)$, assuming $f_1 \in \mathcal{L}_2$, is

$$\begin{gathered} f_1(x) = \sum_{j=-\infty}^{\infty} b_j\phi_j(x) + \sum_{i=0}^{\infty}\sum_{j=-\infty}^{\infty} b_{ij}\psi_{ij}(x), \\ b_j = \int f_1\phi_j, \quad b_{ij} = \int f_1\psi_{ij}. \end{gathered} \tag{1.2.1}$$

We propose a nonlinear wavelet estimator of $f_1(x)$:

$$\hat{f}_1(x) = \sum_{j=-\infty}^{\infty} \hat{b}_j \phi_j(x) + \sum_{i=0}^{q-1} \sum_{j=-\infty}^{\infty} \hat{b}_{ij} I(|\hat{b}_{ij}| > \delta) \psi_{ij}(x), \tag{1.2.2}$$

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

$$\hat{b}_j = \int \phi_j(x) I(x \leq T)\, d\hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k I(Z_k \leq T) \phi_j(Z_k)}{1 - \hat{G}_n(Z_k -)}, \tag{1.2.3}$$

$$\hat{b}_{ij} = \int \psi_{ij}(x) I(x \leq T)\, d\hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k I(Z_k \leq T) \psi_{ij}(Z_k)}{1 - \hat{G}_n(Z_k -)}. \tag{1.2.4}$$

Here $\hat{F}_n$ and $\hat{G}_n$ denote the Kaplan-Meier estimators of distribution functions F and G, respectively, i.e.,

$$\hat{F}_n(x) = 1 - \prod_{k=1}^{n} \left[1 - \frac{\delta_{(k)}}{n-k+1}\right]^{I(Z_{(k)} \leq x)},$$

$$\hat{G}_n(x) = 1 - \prod_{k=1}^{n} \left[1 - \frac{1-\delta_{(k)}}{n-k+1}\right]^{I(Z_{(k)} \leq x)},$$

where, $Z_{(k)}$ is the k-th ordered Z-value and $\delta_{(k)}$ is the concomitant of the i-th order Z statistic, i.e., $\delta_{(k)} = \delta_j$ if $Z_{(k)} = Z_j$. Note that $\delta_k / n(1 - \hat{G}_n(Z_k-))$ is the jump of the Kaplan-Meier estimator $\hat{F}_n$ at Z_k.

Remark 1.2.1 We can define the wavelet estimator of $f(x)$, say $\hat{f}(x)$, instead of $f_1(x)$, similarly to (1.2.2)-(1.2.4). However, in this case, the MISE, i.e. $E\int(\hat{f} - f)^2 < \infty$ cannot be ensured. Thus we typically consider $E\int_{-\infty}^{T}(\hat{f} - f)^2$ to eliminate the endpoint effects. Since wavelet estimator $\hat{f}$ is the same as $\hat{f}_1$ whenever $Z_{(n)} \leq T$, we have

$$\int_{-\infty}^{T} (\hat{f} - f)^2 = \int (\hat{f}_1 - f_1)^2 - \int_{T}^{\infty} \hat{f}_1^2 \simeq \int (\hat{f}_1 - f_1)^2,$$

provided that $T \le Z_{(n)}$. Thus, our analysis for $\hat{f}_1$ is closely related to that for $\hat{f}$ restricted to $(-\infty, T]$.

Remark 1.2.2 Although here we consider survival times setting, the random variables by no means be necessary restrictedly to positive. Suppose there is no censoring, i.e. $G \equiv 0$ on $(-\infty, \infty)$. Then $\delta_k \equiv 1$, for all $k = 1, 2, \cdots, n$ and upon taking $T = \tau_H = \tau_F$, we see that $f_1 \equiv f$ and the above estimator $\hat{f}_1 = \hat{f}$ of Hall and Patil (1995).

1.3 Main results

We assume that the smoothing parameters p, q and δ satisfy the following condition:

(A): $p \to \infty, \quad q \to \infty, \quad p_q\delta^2 \to 0, \quad p^{2r+1}\delta^2 \to \infty, \quad \delta \ge C\sqrt{n^{-1}\ln n}$, where $C > C_0 \equiv 2\{r(2r+1)^{-1} \sup f_1(1-G)^{-1}\}^{1/2}$.

Theorem 1.3.1 *In addition to the conditions on ϕ and ψ stated in section 1.2, assume that the r-th derivative $f^{(r)}$ is continuous on $(-\infty, \infty)$ and is bounded, monotone on $(-\infty, -u)$ and (u, ∞) for a sufficiently large positive u and the censoring distribution function G is continuous. Also assume that condition (A) holds. Then*

$$E\left|\int (\hat{f}_1 - f_1)^2 - \left\{n^{-1}p \int \frac{f_1}{1-G} + p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int f_1^{(r)^2}\right\}\right| = o(n^{-1}p + p^{-2r}). \tag{1.3.1}$$

Remark 1.3.1 This theorem is an analogue of Theorem 1 of Hall and Patil (1995), where the monotonicity of $f^{(r)}$ on (u, ∞) for large positive u is needed. However, it is not needed for censored data case, because of the effect of truncation at T.

Remark 1.3.2 The result (1.3.1) is stronger than traditional asymptotic formula for MISE. It implies a wavelet version of the MISE formula:

$$E\int(\hat{f}_1 - f_1)^2 \sim n^{-1}p\int\frac{f_1}{1-G} + p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int f_1^{(r)^2}.$$

In the Theorem 1.3.1, we have assumed that survival time density f is r-times continuously differentiable and censoring distribution function G is continuous for simplicity and convenience of exposition. However, if $f^{(r)}$ and G are only piecewise continuous, Theorem 1.3.1 still holds. That is the following:

Theorem 1.3.2 *In addition to the conditions on ϕ and ψ stated in section 1.2, assume that the r-th derivative $f^{(r)}$ and G are only piecewise smooth, i.e. there exist points $x_0 = -\infty < x_1 < x_2 < \cdots < x_N < \infty = x_{N+1}$ such that the first r derivatives of f exist and are bounded and continuous on (x_i, x_{i+1}) for $0 \le i \le N$, with left- and right-hand limits; and that $f^{(r)}$ is monotone on $(-\infty, -u)$ and (u, ∞) for sufficiently large positive u. In particular, f and G themselves may be only piecewise continuous. Also assume that condition (A) holds and $p_q^{2r+1}n^{-2r} \to \infty$. Then also (1.3.1) holds.*

1.4 Proofs of the theorems

The proof of the above theorem follows along the lines in Hall and Patil (1995), combined with Stute (1995) which establishes an asymptotic representation for the Kaplan-Meier integral $\int \varphi d\hat{F}_n$ as an average of i.i.d. random variables with a sufficiently small error. This allows for a more traditional and direct approach to the density estimation problem for the censored data, compared to the martingale approach as, e.g., in the Wu and Wells (1999). We begin with some lemmas. To state these lemmas, we first need some addtional natation. Let

$$\varphi_j(x) = \phi_j(x) I(x \le T), \; j = 0, \pm 1, \pm 2, \cdots , \tag{1.4.1}$$

$$\varphi_{ij}(x) = \psi_{ij}(x) I(x \le T), \; i = 0, 1, \cdots , q-1, \; j = 0, \pm 1, \pm 2, \cdots , \tag{1.4.2}$$

$$\tilde{b}_j = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k \varphi_j(Z_k)}{1 - G(Z_k)}, \; j = 0, \pm 1, \pm 2, \cdots , \tag{1.4.3}$$

$$\tilde{b}_{ij} = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k \varphi_{ij}(Z_k)}{1 - G(Z_k)}, \; i = 0, 1, \cdots , q-1; \; j = 0, \pm 1, \pm 2, \cdots , \tag{1.4.4}$$

and

$$\begin{aligned} W_j(Z_k) &= U_j(Z_k) - V_j(Z_k), & W_{ij}(Z_k) &= U_{ij}(Z_k) - V_{ij}(Z_k), \\ \overline{W}_j &= \frac{1}{n} \sum_{k=1}^{n} W_j(Z_k), & \overline{W}_{ij} &= \frac{1}{n} \sum_{k=1}^{n} W_{ij}(Z_k), \end{aligned} \tag{1.4.5}$$

where

$$\begin{aligned} U_j(Z_k) &= \frac{1 - \delta_k}{1 - H(Z_k)} \int_{Z_k}^{\tau_H} \varphi_j(\omega) \, F(d\omega), \\ U_{ij}(Z_k) &= \frac{1 - \delta_k}{1 - H(Z_k)} \int_{Z_k}^{\tau_H} \varphi_{ij}(\omega) \, F(d\omega), \\ V_j(Z_k) &= \int_{-\infty}^{\tau_H} \int_{-\infty}^{\tau_H} \frac{\varphi_j(\omega) I(v < Z_k \wedge \omega)}{[1 - H(v)][1 - G(v)]} G(dv) F(d\omega), \\ V_{ij}(Z_k) &= \int_{-\infty}^{\tau_H} \int_{-\infty}^{\tau_H} \frac{\varphi_{ij}(\omega) I(v < Z_k \wedge \omega)}{[1 - H(v)][1 - G(v)]} G(dv) F(d\omega). \end{aligned} \tag{1.4.6}$$

We also define

$$\tilde{H}^0(z) = P(Z \le z, \delta = 0) = \int_{-\infty}^{z} (1 - F(y))G(dy),$$

$$\tilde{H}^1(z) = P(Z \le z, \delta = 1) = \int_{-\infty}^{z} (1 - G(y-))F(dy),$$

$$H_n(z) = n^{-1} \sum_{i=1}^{n} I(Z_i \le z),$$

$$\tilde{H}_n^j(z) = n^{-1} \sum_{i=1}^{n} I(Z_i \le z, \delta_i = j), \quad j = 0, 1, \quad z \in \mathbb{R},$$

and

$$B_{kn} \equiv n \int_{-\infty}^{Z_k-} \ln \left[1 + \frac{1}{n(1 - H_n(z))} \right] \tilde{H}_n^0(dz) - \int_{-\infty}^{Z_k-} \frac{\tilde{H}_n^0(dz)}{1 - H_n(z)}, \tag{1.4.7}$$

$$C_{kn} \equiv \int_{-\infty}^{Z_k-} \frac{\tilde{H}_n^0(dz)}{1 - H_n(z)} - \int_{-\infty}^{Z_k-} \frac{\tilde{H}^0(dz)}{1 - H(z)}, \quad 1 \le k \le n. \tag{1.4.8}$$

Lemma 1.4.1 *Let $\hat{b}_j$ and $\hat{b}_{ij}$ be defined as in equations (1.2.3) and (1.2.4). Then the following equations hold.*

$$\hat{b}_j = \tilde{b}_j + \overline{W}_j + R_{n,j}, \qquad E(R_{n,j}^2) = O\left(\frac{1}{n^2}\right) \int \varphi_j^2 \, dF, \tag{1.4.9}$$

$$\hat{b}_{ij} = \tilde{b}_{ij} + \overline{W}_{ij} + R_{n,ij}, \qquad E(R_{n,ij}^2) = O\left(\frac{1}{n^2}\right) \int \varphi_{ij}^2 \, dF. \tag{1.4.10}$$

Proof. Because the proofs of (1.4.9) and (1.4.10) are similar, details are given only for (1.4.10). The proof below uses many ideas of the proof of the main theorem of Stute (1995). The main difference is that here the integrand φ_{ij} depend on n, instead of a fixed function in Stute (1995). Another difference is that we need the rate of reminder in second moment, instead of in probability.

Write $\hat{b}_{ij} = \int \varphi_{ij}(x)d\hat{F}_n(x)$. From a result in Stute (1995, p.434) and details of the proof of the theorem in Stute (1995), we obtain

$$\hat{b}_{ij} = \int \varphi_{ij}(w)\gamma_0(w)\tilde{H}_n^1(dw) + \iint \frac{I(v < w)\varphi_{ij}(w)\gamma_0(w)}{1 - H(v)}\tilde{H}^1(dw)\tilde{H}_n^0(dv)$$
$$- \iiint \frac{I(v < u, v < w)\varphi_{ij}(w)\gamma_0(w)}{[1 - H(v)]^2}\tilde{H}^0(dv)\tilde{H}^1(dw)H_n(du) + R_{n,ij},$$

where

$$R_{n,ij} = S_{n1,ij} + S_{n2,ij} + R_{n1,ij} - R_{n2,ij} + 2R_{n3,ij}\,, \tag{1.4.11}$$

H_n, $\tilde{H}_n^0$ and $\tilde{H}_n^1$ are the empirical (sub-) distribution function estimators of H , $\tilde{H}^0$ and $\tilde{H}^1$, respectively, $\gamma_0(Z_k) = 1/(1 - G(Z_k))$, and

$$S_{n1,ij} = \frac{1}{n}\sum_{k=1}^{n} \varphi_{ij}(Z_k)\gamma_0(Z_k)\delta_k B_{kn}, \tag{1.4.12}$$

$$S_{n2,ij} = \frac{1}{2n}\sum_{k=1}^{n} \varphi_{ij}(Z_k)\delta_k e^{\Delta_k}\{B_{kn} + C_{kn}\}^2, \tag{1.4.13}$$

$$R_{n1,ij} = \iint \varphi_{ij}(w)\gamma_0(w)I(z < w)\frac{[H_n(z) - H(z)]^2}{[1 - H(z)]^2[1 - H_n(z)]}\tilde{H}_n^0(dz)\tilde{H}_n^1(dw), \tag{1.4.14}$$

$$\begin{aligned} R_{n2,ij} = &\iiint \frac{I(v < u, v < w)\varphi_{ij}(w)\gamma_0(w)}{[1 - H(v)]^2}H_n(du) \\ &\qquad \times [\tilde{H}_n^0(dv) - \tilde{H}^0(dv)][\tilde{H}_n^1(dw) - \tilde{H}^1(dw)] \\ &+ \iiint \frac{I(v < u, v < w)\varphi_{ij}(w)\gamma_0(w)}{[1 - H(v)]^2} \\ &\qquad \times [H_n(du) - H(du)]\tilde{H}^0(dv)[\tilde{H}_n^1(dw) - \tilde{H}^1(dw)] \\ &+ \iiint \frac{I(v < u, v < w)\varphi_{ij}(w)\gamma_0(w)}{[1 - H(v)]^2} \\ &\qquad \times [H_n(du) - H(du)][\tilde{H}_n^0(dv) - \tilde{H}^0(dv)]\tilde{H}^1(dw), \end{aligned} \tag{1.4.15}$$

$$R_{n3,ij} = \iint \frac{I(v < w)\varphi_{ij}(w)\gamma_0(w)}{1 - H(v)}[\tilde{H}_n^0(dv) - \tilde{H}^0(dv)][\tilde{H}_n^1(dw) - \tilde{H}^1(dw)]. \tag{1.4.16}$$

Writing the first three integrals of $\hat{b}_{ij}$ as sum and using the definitions of $U_{ij}(Z_k)$, $V_{ij}(Z_k)$ and $W_{ij}(Z_k)$, we have

$$\hat{b}_{ij} = \tilde{b}_{ij} + \frac{1}{n}\sum_{k=1}^{n} U_{ij}(Z_k) - \frac{1}{n}\sum_{k=1}^{n} V_{ij}(Z_k) + R_{n,ij}$$

$$= \tilde{b}_{ij} + \overline{W}_{ij} + R_{n,ij}.$$

Thus, to prove (1.4.10), it suffices to bound the five terms of the RHS of (1.4.11). Hence, the lemma follows from the following five propositions. All proofs are given in the next section.

Proposition 1.4.1 *Under the assumptions of Theorem 1.3.1,*

$$E\left(S_{n1,ij}^2\right) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF.$$

Proposition 1.4.2 *Under the assumptions of Theorem 1.3.1,*

$$E\left(S_{n2,ij}^2\right) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF.$$

Proposition 1.4.3 *Under the assumptions of Theorem 1.3.1,*

$$E\left(R_{n1,ij}^2\right) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF.$$

Proposition 1.4.4 *Under the assumptions of Theorem 1.3.1,*

$$E\left(R_{n2,ij}^2\right) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF.$$

Proposition 1.4.5 *Under the assumptions of Theorem 1.3.1,*

$$E\left(R_{n3,ij}^2\right) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF.$$

Lemma 1.4.2 *Under the assumptions of Theorem 1.3.1,*

$$s_1 \equiv E\left|\sum_j (\hat{b}_j - b_j)^2 - n^{-1}p\int \frac{f_1}{1-G}\right| = o(n^{-1}p).$$

Proof. In view of (1.4.9),

$$\begin{aligned} s_1 \leq & E\left|\sum_j (\tilde{b}_j - b_j)^2 - n^{-1}p\int \frac{f_1}{1-G}\right| + E\sum_j \overline{W}_j^2 + E\sum_j R_{n,j}^2 \\ & + 2E\sum_j |\tilde{b}_j - b_j||R_{n,j}| + 2E\sum_j |\tilde{b}_j - b_j||\overline{W}_j| + 2E\sum_j |R_{n,j}||\overline{W}_j| \\ = & s_{11} + s_{12} + s_{13} + s_{14} + s_{15} + s_{16}, \quad (say). \end{aligned}$$

Noticing that

$$nE(\tilde{b}_j - b_j)^2 = \int \phi^2(y)\frac{f_1((y+j)/p)}{1-G((y+j)/p)}\,dy - b_j^2,$$

we obtain

$$\sum_j E(\tilde{b}_j - b_j)^2 = \frac{p}{n}\int \phi^2(y)\sum_j p^{-1}\frac{f_1((y+j)/p)}{1-G((y+j)/p)}\,dy - \frac{1}{n}\sum_j b_j^2.$$

Since $\int \phi^2 = 1$,

$$\sum_j p^{-1}f_1((y+j)/p)/(1-G((y+j)/p)) \to \int f_1/(1-G),$$

and $\sum_j b_j^2 = O(\int f_1^2)$, then

$$E\sum_j (\tilde{b}_j - b_j)^2 = n^{-1}p\int f_1/(1-G) + o(n^{-1}p).$$

As a consequence of $Z_k \leq T$ for all $k = 1, 2, \cdots, n$, all denominators appearing in (1.4.3) and (1.4.4) are bounded away from below. Thus they may be handled along the same lines as those in Hall and Patil's paper p.922 to show that $\mathrm{Var}(\sum_j(\tilde{b}_j - b_j)^2) = o(n^{-2}p^2)$. So we obtain $s_{11} = o(n^{-1}p)$.

From (1.4.5), we have

$$s_{12} \le n^{-1} \sum_j EW_j^2(Z_1) \le 2n^{-1} \sum_j (EU_j^2(Z_1) + EV_j^2(Z_1)).$$

In view of (1.4.6), applying the Cauchy-Schwarz inequality and using the compact support of ϕ, we finally can obtain

$$EU_j^2(Z_1) \le \frac{1}{[1-H(T)][1-G(T)]} p^{-1} \int \phi^2(u) f_1^2((u+j)/p) du. \tag{1.4.17}$$

So we obtain

$$n^{-1} \sum_j EU_j^2(Z_1) = O\left(n^{-1} \int \phi^2(u) \sum_j p^{-1} f_1^2((u+j)/p) du\right) = o(n^{-1}p).$$

By applying the same argument, we can obtain

$$EV_j^2(Z_1) \le \frac{1}{[1-H(T)]^2[1-G(T)]^2} p^{-1} \int \phi^2(u) f_1^2((u+j)/p) du, \tag{1.4.18}$$

thus, $n^{-1} \sum_j EV_j^2(Z_1) = o(n^{-1}p)$ too. Hence $s_{12} = o(n^{-1}p)$. By (1.4.9),

$$s_{13} = O\left(\frac{1}{n^2}\right) \sum_j \int \varphi_j^2 dF = O\left(\frac{p}{n^2}\right) = o(n^{-1}p).$$

Finally, by applying Cauchy-Schwarz inequality twice to s_{14}, together with s_{11} and s_{13}, we obtain

$$s_{14} \le 2\left(\sum_j E(\tilde{b}_j - b_j)^2 \cdot \sum_j ER_{n,j}^2\right)^{1/2} = o(n^{-1}p).$$

By applying the same argument, we can show $s_{15} = s_{16} = o(n^{-1}p)$ too. This completes the proof of the lemma.

Lemma 1.4.3 *Under the assumptions of Theorem 1.3.1,*

$$s_2 \equiv \sum_{i=0}^{q-1} \sum_j E\Big\{(\hat{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta)\Big\} = o(n^{-2r/(2r+1)}).$$

Proof. The proof is analogous to that of Theorem 2.1 of Hall and Patil (1995, p.916-918). The difference is that we need to take care of two additional terms, $\overline{W}_{ij}$ and $R_{n,ij}$. Thus we only provide the detail for these two parts. As there, let α and β denote positive numbers satisfying $\alpha + \beta = 1$, and set

$$s_{21} = \sum_{i=0}^{q-1} \sum_j E\{(\hat{b}_{ij} - b_{ij})^2\} I(|b_{ij}| > \alpha\delta),$$
$$s_{22} = \sum_{i=0}^{q-1} \sum_j E\{(\hat{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij} - b_{ij}| > \beta\delta)\}.$$

So $s_2 \le s_{21} + s_{22}$. By (1.4.10), we have

$$\begin{aligned} s_{21} \le& 3\sum_{i=0}^{q-1} \sum_j E\{(\tilde{b}_{ij} - b_{ij})^2\} I(|b_{ij}| > \alpha\delta) \\ &+ 3\sum_{i=0}^{q-1} \sum_j E\overline{W}_{ij}^2 I(|b_{ij}| > \alpha\delta) \\ &+ 3\sum_{i=0}^{q-1} \sum_j ER_{n,ij}^2 I(|b_{ij}| > \alpha\delta) \\ =& 3(s_{21,1} + s_{21,2} + s_{21,3}), \quad (say). \end{aligned}$$

Define $f_{1,ij} = \sup_{y \in \text{supp}\,\psi} f_1((y+j)/p_i)/(1 - G((y+j)/p_i))$. Since the denominator $1 - G((y+j)/p_i) \ge 1 - G(T)$, for all i, j and y, we have $f_{1,ij} \le \sup_{y \in \text{supp}\,\psi} (1 - G(T))^{-1} f_1((y+j)/p_i)$. Because f_1 is bounded and monotone in the extreme tails and ψ has a compact support $(-v, v)$, we have, for suffiently large K,

$$\begin{aligned} \sup_{n;i\ge 0} \frac{1}{p_i} \sum_j f_{1,ij} &\le (1 - G(T))^{-1} \sup_{n;i\ge 0} \frac{1}{p_i} \sum_j \sup_{y \in (-v,v)} f_1((y+j)/p_i) \\ &\le (1 - G(T))^{-1} \left[\sup_{n;i\ge 0} \frac{1}{p_i} \sum_{|j/p_i| > K} f_1((u+j)/p_i) + K \sup f_1 \right] \\ &\le (1 - G(T))^{-1} \left[\sup_{n;i\ge 0} \int_{|x|>K} f_1(u/p_i + x)dx + K \sup f_1 \right], \end{aligned}$$

where $u = v$ or $-v$, depending on the monotony of f_1. Hence we have

$$\sup_{n;i\geq 0} p_i^{-1} \sum_j f_{1,ij} < \infty \quad \text{and} \quad \sup_{n;i\geq 0} p_i^{-1} \sum_j \int \varphi_{ij}^2 dF < \infty. \tag{1.4.19}$$

Use this fact and an argument as in Hall and Patil (p.916-917) to obtain $s_{21.1} = o(n^{-2r/(2r+1)})$. As to the $s_{21.2}$, from (1.4.5),

$$s_{21.2} \leq \frac{2}{n} \sum_{i=0}^{q-1} \sum_j \left[EU_{ij}^2(Z_1) + EV_{ij}^2(Z_1) \right]. \tag{1.4.20}$$

By applying an argument similar to (1.4.17) in Lemma 1.4.2, we have

$$EU_{ij}^2(Z_1) = O\Big(p_i^{-1} \int \phi^2(u) f_1^2((u+j)/p_i) du \Big),$$

so that

$$\begin{aligned} \frac{2}{n} \sum_{i=0}^{q-1} \sum_j EU_{ij}^2(Z_1) &= O\left(\frac{1}{n}\right) \sum_{i=0}^{q-1} \int \phi^2(u) \sum_j p_i^{-1} f_1^2((u+j)/p_i) du \\ &= O\left(\frac{\ln n}{n}\right) = o(n^{-2r/(2r+1)}), \end{aligned}$$

where the second equality holds from $q = O(\ln n)$, and $\sum_j p_i^{-1} f_1^2((y+j)/p_i) \to \int f_1^2$.

By applying an argument similar to (1.4.18) to the second term of (1.4.20), we have $s_{21,2} = o(n^{-2r/(2r+1)})$. Next, from (1.4.10),

$$\begin{aligned} s_{21,3} &\leq \sum_{i=0}^{q-1} \sum_j ER_{n,ij}^2 = O\left(\frac{1}{n^2}\right) \sum_{i=0}^{q-1} p_i \sum_j p_i^{-1} \int \varphi_{ij}^2 dF \\ &= O\left(\frac{p_q}{n^2}\right) \qquad \text{by (1.4.19)}, \\ &= o(n^{-2r/(2r+1)}) \qquad \text{by } n^{-1} p_q \to 0. \end{aligned} \tag{1.4.21}$$

Thus, $s_{21} = o(n^{-2r/(2r+1)})$.

As to the s_{22}, by (1.4.10) we have

$$\begin{aligned} s_{22} \leq & 3\sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij} - b_{ij}| > \beta\delta)\} \\ & + 3\sum_{i=0}^{q-1}\sum_{j} E\{\overline{W}_{ij}^2 I(|\hat{b}_{ij} - b_{ij}| > \beta\delta)\} \\ & + 3\sum_{i=0}^{q-1}\sum_{j} E\{R_{n,ij}^2 I(|\hat{b}_{ij} - b_{ij}| > \beta\delta)\} \\ = & 3(s_{22,1} + s_{22,2} + s_{22,3}), \quad (say). \end{aligned}$$

By applying the similar argument as that in $s_{21,2}$ and $s_{21,3}$, it is obvious that $s_{22,2} = s_{22,3} = o(n^{-2r/(2r+1)})$. To complete the proof of this lemma, it thus suffices to prove $s_{22,1} = o(n^{-2r/(2r+1)})$. In view of (1.4.10), we have

$$\begin{aligned} s_{22,1} \leq & \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|\tilde{b}_{ij} - b_{ij}| > \alpha_1\beta\delta)\} \\ & + \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|\overline{W}_{ij}| > \alpha_2\beta\delta)\} \\ & + \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|R_{n,ij}| > \alpha_3\beta\delta)\} \\ = & s_{22,11} + s_{22,12} + s_{22,13}, \end{aligned}$$

where α_1, α_2 and α_3 are positive numbers such that $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

The term $s_{22,11}$ is similar to s_{12} in Hall and Patil, following the argument there, noticing $\tilde{b}_{ij}$ and $f_{1,ij}$ play the roles of $\hat{b}_{ij}$ and f_{ij} there, we can show that

$$s_{22,11} = o\left(n^{-2r/(2r+1)}\right). \tag{1.4.22}$$

As to the $s_{22,13}$, let $A = \{|\tilde{b}_{ij} - b_{ij}| \le \alpha_1\beta\delta\}$, then,

$$\begin{aligned}
s_{22,13} =& \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|R_{n,ij}| > \alpha_3\beta\delta) I_A\} \\
&+ \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|R_{n,ij}| > \alpha_3\beta\delta) I_{A^c}\} \\
\le& \sum_{i=0}^{q-1}\sum_{j} \alpha_1^2\beta^2\delta^2 P(|R_{n,ij}| > \alpha_3\beta\delta) \\
&+ \sum_{i=0}^{q-1}\sum_{j} E\{(\tilde{b}_{ij} - b_{ij})^2 I(|\tilde{b}_{ij} - b_{ij}| > \alpha_1\beta\delta)\} \\
\le& \sum_{i=0}^{q-1}\sum_{j} \alpha_1^2\beta^2\delta^2 \frac{ER_{n,ij}^2}{\alpha_3^2\beta^2\delta^2} + s_{22,11} \\
=& O\left(\sum_{i=0}^{q-1}\sum_{j} ER_{n,ij}^2\right) + s_{22,11} \\
=& o\left(n^{-2r/(2r+1)}\right),
\end{aligned}$$

the last equality follows from (1.4.21) and (1.4.22). By applying the same argument to $s_{22,12}$, we can show that $s_{22,12} = o\left(n^{-2r/(2r+1)}\right)$ too. This completes the proof of the Lemma 1.4.3.

Lemma 1.4.4 *Under the assumptions of Theorem 1.3.1,*

$$s_3 \equiv E\left|\sum_{i=0}^{q-1}\sum_{j} b_{ij}^2 I(|\hat{b}_{ij}| \le \delta) - p^{-2r}\kappa^2(1 - 2^{-2r})^{-1}\int f_1^{(r)^2}\right| = o(p^{-2r}).$$

Proof. The proof here is again analogous to that of Hall and Patil (1995) and the details are omitted. We only provide special treatment for the additional two terms related to $\overline{W}_{ij}$ and $R_{n,ij}$. As there, let $\epsilon > 0$, and define

$$s_{30} = \sum_{i=0}^{q-1}\sum_{j} b_{ij}^2 I(|\hat{b}_{ij}| \le \delta), \qquad s_{31} = \sum_{i=0}^{q-1}\sum_{j} b_{ij}^2 I\{|b_{ij}| \le (1+\epsilon)\delta\},$$

$$s_{32} = \sum_{i=0}^{q-1}\sum_{j} b_{ij}^2 I\{|b_{ij}| \le (1-\epsilon)\delta\}, \qquad \Delta = \sum_{i=0}^{q-1}\sum_{j} b_{ij}^2 I(|\hat{b}_{ij} - b_{ij}| > \epsilon\delta).$$

Then

$$s_{32} - \Delta \le s_{30} \le s_{31} + \Delta\,. \tag{1.4.23}$$

By applying the arguments analogous to those of Hall and Patil (1995, p.918-921) to s_{31} and s_{32}, we can obtain

$$s_{31} = s_{32} = p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int f_1^{(r)^2} + o(p^{-2r}). \tag{1.4.24}$$

Now,

$$\begin{aligned}
E\Delta \le & \sum_{i=0}^{q-1}\sum_j b_{ij}^2 P(|\tilde{b}_{ij} - b_{ij}| > \gamma_1\epsilon\delta) + \sum_{i=0}^{q-1}\sum_j b_{ij}^2 P(|\overline{W}_{ij}| > \gamma_2\epsilon\delta) \\
& + \sum_{i=0}^{q-1}\sum_j b_{ij}^2 P(|R_{n,ij}| > \gamma_3\epsilon\delta) \\
= & \Delta_1 + \Delta_2 + \Delta_3,\ (say),
\end{aligned}$$

where γ_1, γ_2 and γ_3 are positive numbers such that $\gamma_1 + \gamma_2 + \gamma_3 = 1$.

$$\begin{aligned}
\Delta_1 &= O\left(\sum_{i=0}^{q-1}\sum_j b_{ij}^2 \exp\left[-\frac{1}{2}(1-\epsilon)\gamma_1^2\beta^2 f_{1,ij}^{-1} n\delta^2\right]\right) \\
&= o\left(\sum_{i=0}^{q-1}\sum_j b_{ij}^2\right) = o(s_{31})\,.
\end{aligned}$$

The first equality follows from Bernstein's inequality (see Härdle, et al. 1998, p.244), while the second follows from $n\delta^2 \to \infty$.

$$\begin{aligned}
\Delta_2 &= O\left(\sum_{i=0}^{q-1}\sum_j b_{ij}^2 \frac{E\overline{W}_{ij}^2}{\gamma_2^2\epsilon^2\delta^2}\right) = O\left(\sum_{i=0}^{q-1}\sum_j b_{ij}^2 \frac{1}{n\delta^2}\int \varphi_{ij}^2 dF\right) \\
&= o\left(\sum_{i=0}^{q-1}\sum_j b_{ij}^2\right) = o(s_{31})\,.
\end{aligned}$$

The second equality follows from the arguments in (1.4.17) and (1.4.18), while the third follows from $n\delta^2 \to \infty$. Similarly, we can show $\Delta_3 = o(s_{31})$ too. Thus, $E\Delta = o(s_{31})$. Combining (1.4.23) and (1.4.24), the proof of lemma follows.

Lemma 1.4.5 *Under the assumptions of Theorem 1.3.1,*

$$s_4 \equiv \sum_{i=q}^{\infty} \sum_{j} b_{ij}^2 = o(p^{-2r}).$$

Proof. The proof follows from the step 3 of Theorem 2.1 of Hall and Patil (1995).

We are now in the position to give the proof of the Theorem 1.3.1 and 1.3.2.

Proof of the Theorem 1.3.1. Proof follows from the bound

$$E\left|\int (\hat{f}_1 - f_1)^2 - \left\{n^{-1}p \int \frac{f_1}{1-G} + p^{-2r}\kappa^2(1-2^{-2r})^{-1} \int f_1^{(r)^2}\right\}\right|$$
$$\le s_1 + s_2 + s_3 + s_4,$$

and Lemmas 1.4.2, 1.4.3, 1.4.4 and 1.4.5.

Proof of the Theorem 1.3.2. We use the same notations as in Hall and Patil (1995). Noticing that, by the orthogonality properties of ϕ and ψ,

$$\int (\hat{f}_1 - f_1)^2 = I_q(\mathbb{Z}, \mathbb{Z}, \dots),$$

where $\mathbb{Z}$ denotes the set of all intergers and

$$I_q(\Psi, \Psi_0, \Psi_1, \dots) = \sum_{j \in \Psi} (\hat{b}_j - b_j)^2 + \sum_{i=0}^{q-1} \sum_{j \in \Psi_i} (\hat{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta)$$
$$+ \sum_{i=0}^{q-1} \sum_{j \in \Psi_i} b_{ij}^2 I(|\hat{b}_{ij}| \le \delta) + \sum_{i=q}^{\infty} \sum_{j \in \Psi_i} b_{ij}^2.$$

By (1.4.9) and (1.4.10),

$$
\begin{aligned}
&I_q(\Psi,\Psi_0,\Psi_1,\dots)\\
&=\sum_{j\in\Psi}(\tilde{b}_j-b_j)^2+\sum_{j\in\Psi}\overline{W}_j^2+\sum_{j\in\Psi}R_{n,j}^2+2\sum_{j\in\Psi}(\tilde{b}_j-b_j)\overline{W}_j\\
&\quad+2\sum_{j\in\Psi}(\tilde{b}_j-b_j)R_{n,j}+2\sum_{j\in\Psi}\overline{W}_jR_{n,j}+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\tilde{b}_{ij}-b_{ij})^2I(|\hat{b}_{ij}|>\delta)\\
&\quad+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}\overline{W}_{ij}^2I(|\hat{b}_{ij}|>\delta)+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}R_{n,ij}^2I(|\hat{b}_{ij}|>\delta)\\
&\quad+2\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\tilde{b}_{ij}-b_{ij})\overline{W}_{ij}I(|\hat{b}_{ij}|>\delta)+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}(\tilde{b}_{ij}-b_{ij})R_{n,ij}I(|\hat{b}_{ij}|>\delta)\\
&\quad+2\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}\overline{W}_{ij}R_{n,ij}I(|\hat{b}_{ij}|>\delta)+\sum_{i=0}^{q-1}\sum_{j\in\Psi_i}b_{ij}^2I(|\hat{b}_{ij}|\le\delta)+\sum_{i=q}^{\infty}\sum_{j\in\Psi_i}b_{ij}^2\\
&=I_1+I_2+I_3+I_4+I_5+I_6+I_7+I_8+I_9+I_{10}\\
&\qquad+I_{11}+I_{12}+I_{13}+I_{14},\quad (say).
\end{aligned}
$$

By the Theorem 1.1 of Stute (1995), when F and G are only piecewise continuous, a quartile transformation may be applied so as to trace everything back to uniformly distributed Z's. Thus the above Lemma 1.4.1 still holds in this case (see also Stute and Wang (1993, p.1605).

From (1.4.17) and (1.4.18) in the Lemma 1.4.2, we obtain $EI_2=o(n^{-1}p)$. From (1.4.9) in the Lemma 1.4.1, it is easy to see $EI_3=O(n^{-2}p)=o(n^{-1}p)$. From the proof of Lemma 1.4.3, we have $EI_8=o(n^{-2r/(2r+1)})$. From (1.4.21), we also have $EI_9=o(n^{-2r/(2r+1)})$. Applying the Cauchy-Schwarz inequality, we can show I_4, I_5, I_6, I_{10}, I_{11} and I_{12} are all of the order $o(n^{-2r/(2r+1)})$.

When f_1 is only piecewise smooth, let Π denote the finite set of points where $f_1^{(s)}$ has discontinuities for some $0\le s\le r$. Suppose supp $\phi\subseteq(-v,v)$, supp

$\psi \subseteq (-v, v)$ and let

$$\mathbb{K} = \{k : k \in (px - v, px + v) \quad \text{for some} \quad x \in \Pi\},$$

$$\mathbb{K}_i = \{k : k \in (p_i x - v, p_i x + v) \quad \text{for some} \quad x \in \Pi\}.$$

Also let $\mathbb{K}^c$, $\mathbb{K}_i^c$ denote their complements. Then, unless $j \in \mathbb{K}_i$, b_{ij} and $\tilde{b}_{ij}$ are constructed entirely from an integral over or an average of data values from an interval where $f_1^{(r)}$ exists and is bounded. Also, unless $j \in \mathbb{K}$, b_j and $\tilde{b}_j$ are constructed solely from such regions. Thus we may write

$$\begin{aligned} I_q(\Psi, \Psi_0, \Psi_1, \dots) = {} & I_1(\mathbb{K}) + I_2 + I_3 + I_4 + I_5 + I_6 + I_7(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots) \\ & + I_8 + I_9 + I_{10} + I_{11} + I_{12} \\ & + I_{13}(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots) + I_{14}(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots) \\ & + I_1(\mathbb{K}^c) + I_7(\mathbb{K}_0^c, \mathbb{K}_1^c, \mathbb{K}_2^c, \dots) \\ & + I_{13}(\mathbb{K}_0^c, \mathbb{K}_1^c, \mathbb{K}_2^c, \dots) + I_{14}(\mathbb{K}_0^c, \mathbb{K}_1^c, \mathbb{K}_2^c, \dots), \end{aligned} \tag{1.4.25}$$

where

$$I_1(\mathbb{K}) = \sum_{j \in \mathbb{K}} (\tilde{b}_j - b_j)^2, \qquad I_1(\mathbb{K}^c) = \sum_{j \in \mathbb{K}^c} (\tilde{b}_j - b_j)^2,$$

$$I_7(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots) = \sum_{i=0}^{q-1} \sum_{j \in \mathbb{K}_i} (\tilde{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta),$$

$$I_7(\mathbb{K}_0^c, \mathbb{K}_1^c, \mathbb{K}_2^c, \dots) = \sum_{i=0}^{q-1} \sum_{j \in \mathbb{K}_i^c} (\tilde{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta),$$

the rest of the terms are defined similarly. However, for our compactly supported wavelet ϕ and ψ, both $\mathbb{K}$ and $\mathbb{K}_i$ have no more than $(2v + 1)(\#\Pi)$ elements for each i. Considering $q = O(\ln n)$, we can show $I_1(\mathbb{K})$, $I_7(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots)$, and

$I_{13}(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots)$ are of the lower order $o(n^{-2r/(2r+1)})$. Thus it is negligible compared to the main terms of MISE. Although b_{ij} is only of the order $p_i^{-1/2}$ when f_1 is not r-times smooth, based on theorem's additional assumption $p_q^{2r+1}n^{-2r} \to \infty$, we readily see that $I_{14}(\mathbb{K}_0, \mathbb{K}_1, \mathbb{K}_2, \dots) = o(n^{-2r/(2r+1)})$. By tracing the whole proof of Theorem 1.3.1 carefully, we will see the rest of the terms of the right hand side of (1.4.25) have precisely the asymptotic properties claimed for $\int(\hat{f}_1 - f_1)^2$ in Theorem 1.3.2.

1.5 Proofs of the propositions

Proof of the Proposition 1.4.1. In view of (1.4.12), applying the moment inequality to $S_{n1,ij}$ and taking expectation yields

$$\begin{aligned} E\left(S_{n1,ij}^2\right) &\le E\Big(\frac{1}{n}\sum_{k=1}^{n}\varphi_{ij}^2(Z_k)\gamma_0^2(Z_k)\delta_k B_{kn}^2\Big) \\ &= \frac{1}{n}\sum_{k=1}^{n} E\Big(\varphi_{ij}^2(Z_k)\gamma_0^2(Z_k)\delta_k E(B_{kn}^2|(Z_k,\delta_k))\Big). \end{aligned} \tag{1.5.1}$$

In view of (1.4.7), the definition of B_{kn} and note that

$$x - \frac{x^2}{2} \le \ln(1+x) \le x, \quad \text{for} \quad x \ge 0,$$

we have

$$|B_{kn}| \le \frac{1}{2n}\int_{-\infty}^{Z_k-} \frac{\tilde{H}_n^0(dz)}{[1-H_n(z)]^2} \le \frac{1}{n(1-H_n(Z_k-))}. \tag{1.5.2}$$

Thus, conditionally on $\{Z_k = z\}$ and $\{\delta_k = d\}$, noticing $nH_n(z-) = \sum_{j=1}^{n} I(Z_j < z)$ is a binomial random variable with parameters $n-1$ and $p := H(z-)$, we have

$$E\left(B_{kn}^2|Z_k = z, \delta_k = d\right) \le E\left(\frac{1}{n^2(1-H_n(z-))^2}\right) \le \frac{2}{n^2(1-p)^2}.$$

Thus, from (1.5.1), notice (Z_1, δ_1), $(Z_2, \delta_2), \cdots, (Z_n, \delta_n)$ are i.i.d., we have

$$E(S^2_{n1,ij}) \leq E\left(\varphi^2_{ij}(Z_1)\gamma^2_0(Z_1)\delta_1 \frac{2}{n^2(1 - H(Z_1-))^2}\right) = O\left(\frac{1}{n^2}\right) \int \varphi^2_{ij}\, dF.$$

The last equality follows, because $\{Z_1 \leq T\}$, $T < \tau_H$, imply that $1 - H(Z_1-) \geq 1 - H(T) > 0$.

Remark 1.5.1 If we consider the estimation of f, instead of trunction f_1. Then we have

$$E\left(S^2_{n1,ij}\right) = O\left(\frac{1}{n^2}\right) \int \frac{\psi^2_{ij}(x) F(dx)}{[1 - H(x)]^2 [1 - G(x)]}.$$

The above integral will be infinity for some i and j, such that $\psi^2_{ij} \geq K > 0$ on a neighborhood of τ_H. Since $q \to \infty$ and $j \in (-\infty, \infty)$, there always exist some i and j satisfing above condition. Thus we could not show $\sum_{i=0}^{q-1} \sum_j E(S^2_{n1,ij}) = o(n^{-2r/(2r+1)})$ without the truncation.

Proof of the Proposition 1.4.2. In view of (1.4.13),

$$|S_{n2,ij}| \leq \frac{1}{n} \sum_{k=1}^{n} |\varphi_{ij}(Z_k)| \delta_k e^{\Delta_k} B^2_{kn} + \frac{1}{n} \sum_{k=1}^{n} |\varphi_{ij}(Z_k)| \delta_k e^{\Delta_k} C^2_{kn}.$$

Again applying the moment inequality to the average and taking expectation yields

$$E\left(S^2_{n2,ij}\right) \leq \frac{2}{n} \sum_{k=1}^{n} E\left(\varphi^2_{ij}(Z_k) \delta_k e^{2\Delta_k} B^4_{kn}\right) + \frac{2}{n} \sum_{k=1}^{n} E\left(\varphi^2_{ij}(Z_k) \delta_k e^{2\Delta_k} C^4_{kn}\right).$$

Because the proof of the two terms are similar and the second term is more involved and require more details, we here only need to prove, for any k,

$$E\left(\varphi^2_{ij}(Z_k) \delta_k e^{2\Delta_k} C^4_{kn}\right) = O\left(\frac{1}{n^2}\right) \int \varphi^2_{ij}\, dF. \qquad (1.5.3)$$

Writing LHS of (1.5.3) as

$$E\Big(\varphi_{ij}^2(Z_k)\delta_k E\big(e^{2\Delta_k}C_{kn}^4|(Z_k,\delta_k)\big)\Big)$$
$$\leq E\Big[\varphi_{ij}^2(Z_k)\delta_k E^{1/2}\big(e^{4\Delta_k}|(Z_k,\delta_k)\big)E^{1/2}\big(C_{kn}^8|(Z_k,\delta_k)\big)\Big].$$

Next, we want to show, conditionally on $\{Z_k = z\}$ and $\{\delta_k = d\}$, uniformly in k, z and d,

$$E(e^{4\Delta_k}|Z_k = z, \delta_k = d) = O(1), \tag{1.5.4}$$

$$E(C_{kn}^8|Z_k = z, \delta_k = d) = O\left(\frac{1}{n^4}\right). \tag{1.5.5}$$

Then (1.5.3) follows, thus Proposition 1.4.2 follows.

We first prove (1.5.4). Since (Z_1,δ_1), $(Z_2,\delta_2),\cdots,(Z_n,\delta_n)$ are i.i.d., we here only consider $k=1$ for the convenience of exposition. Recalling Δ_1 is between two terms: $\tilde{x}_1$ and x_1, where

$$\tilde{x}_1 = n\int_{-\infty}^{Z_1-} \ln\left[1+\frac{1}{n(1-H_n(z))}\right]\tilde{H}_n^0(dz), \qquad x_1 = \int_{-\infty}^{Z_1-}\frac{\tilde{H}^0(dz)}{1-H(z)}.$$

Thus there exists $0 \leq \lambda \leq 1$, such that

$$\Delta_1 = x_1 + \lambda(\tilde{x}_1 - x_1) = x_1 + \lambda(B_{1n} + C_{1n}),$$

where B_{1n} is as in (1.4.7) and C_{1n} is defined as in (1.4.8). Notice $0 \leq x_1 = -\ln(1-G(Z_1-)) \leq -\ln(1-G(T))$, which is bounded, and

$$0 \leq |\lambda B_{1n}| \leq |B_{1n}| \leq [n(1-H_n(Z_1-)]^{-1} \leq 1,$$

by $n(1-H_n(Z_1-) \geq 1$. So we obtain

$$e^{4\Delta_1} = \exp\{4x_1 + 4\lambda B_{1n} + 4\lambda C_{1n}\} \leq (1-G(T))^{-4}e^4 e^{4\lambda C_{1n}} = C\cdot e^{4\lambda C_{1n}},$$

where C is a positive constant, the concrete value of which may changes from line to line in the sequel. Thus, in order to prove (1.5.4), we only need to prove, uniformly in z_1 and d_1,

$$E\left(e^{4\lambda C_{1n}}|Z_1 = z_1, \delta_1 = d_1\right) = O(1). \tag{1.5.6}$$

Note that

$$\begin{aligned}\frac{1}{1 - H_n(z)} = &-\frac{1 - H_n(z)}{[1 - H(z)]^2} + \frac{2}{1 - H(z)} \\ &+ \frac{[H_n(z) - H(z)]^2}{[1 - H(z)]^2[1 - H_n(z)]},\end{aligned}$$

we can rewrite C_{1n} as

$$\begin{aligned}C_{1n} = &-\int_{-\infty}^{Z_1-} \frac{1 - H_n(z)}{[1 - H(z)]^2}\tilde{H}_n^0(dz) + \int_{-\infty}^{Z_1-} \frac{2}{1 - H(z)}\tilde{H}_n^0(dz) \\ &- \int_{-\infty}^{Z_1-} \frac{1}{1 - H(z)}\tilde{H}^0(dz) + \int_{-\infty}^{Z_1-} \frac{[H_n(z) - H(z)]^2}{[1 - H(z)]^2[1 - H_n(z)]}\tilde{H}_n^0(dz) \\ =&I_1 + I_2 + I_3 + I_4, (say).\end{aligned}$$

The first three terms of C_{1n} are bounded w.p.1, because $\{Z_1 \leq T\}$ and $T < \tau_H$, all denominators are bounded away from below. Thus we need only to deal with the fourth term I_4. Because of $\lambda \leq 1$, to prove (1.5.6), it suffices to prove, uniformly in z_1 and d_1,

$$E\left(e^{4I_4}|Z_1 = z_1, \delta_1 = d_1\right) = O(1). \tag{1.5.7}$$

Let $J(x) = [(n-1)^{-1}\sum_{j=2}^{n} I(Z_j \le x) - H(x)][1-H(x)]^{-1}$, writing I_4 as a sum, conditionally on $\{Z_1 = z_1\}$, $\{\delta_1 = d_1\}$, we have

$$
\begin{aligned}
I_4 &\le 2\sum_{k=2}^{n} \frac{J^2(Z_k)I(Z_k < z_1)}{n[1-H_n(Z_k)]} + \frac{2}{n^2}\sum_{k=2}^{n} \frac{H^2(Z_k)I(Z_k < z_1)}{[1-H(Z_k)]^2 n[1-H_n(Z_k)]} \\
&\le 2\sup_{z<T} J^2(z) \sum_{k=2}^{n} \frac{I(Z_k < z_1)}{n[1-H_n(Z_k)]} + \frac{2H^2(T)}{n^2[1-H(T)]^2}\sum_{k=2}^{n}\frac{I(Z_k < z_1)}{n[1-H_n(Z_k)]} \\
&\le 2\sup_{z<T} J^2(z)\ \ln n + \frac{C\ln n}{n^2}.
\end{aligned}
$$

Since the second term goes to zero, we only need to bound the first term. From the above inequality, we have $e^{4I_4} \le \sup_{z<T} Cn^{8J^2(z)}$. Since $J(z)$ is a martingale in z (see Koul, 1992, p.42), $n^{8J^2(z)}$ is a sub-martingale. Thus, from the property of the sub-martingale, we have, uniformly in z_1, δ_1,

$$
E\left(e^{4I_4} | Z_1 = z_1, \delta_1 = d_1\right) \le CE\left(n^{8J^2(T)}\right).
$$

Noticing $\sum_{j=2}^{n} I(Z_j \le T)$ is a binomial random variable with parameter $n-1$ and $p := H(T)$. Let $b(k;n,p) = \binom{n}{k}p^k q^{n-k}$, through direct calculation and enlargerment, we have

$$
En^{8J^2(T)} = \sum_{k=0}^{n-1} n^{\frac{8(k-np+p)^2}{(n-1)^2(1-p)^2}} b(k;n-1,p) \le \frac{C}{p}\sum_{k=0}^{n} n^{\frac{16(k-np)^2}{n^2(1-p)^2}} b(k;n,p) =: p^{-1}J.
$$

Since $p = H(T) > 0$, we need to bound the J. The idea is to divide the sum into two parts according to the magnitude of $n^{\frac{16(k-np)^2}{n^2(1-p)^2}}$. To make it clear, let $\mathbf{A} = \{k; |k-np| \le n^d, d \in (1/2, 1)\}$. We write $J = \sum_{k\in\mathbf{A}} + \sum_{k\in\mathbf{A}^c} =: J_1 + J_2$. It is easy to see $J_1 \le n^{16n^{2d-2}(1-p)^{-2}} = O(1)$. As to the J_2, when $k > np + n^d$, we have

(see Feller, 1957 p.163)

$$b(k;n,p) \le b(m;n,p)\, e^{-\frac{1}{2}p\xi_k^2+\delta^2}, \qquad |\delta| < \frac{1}{2},$$

where $\xi_k = \frac{k-(n+1)p+\frac{1}{2}}{\sqrt{(n+1)pq}}$, $(n+1)p - 1 < m \le (n+1)p$ and $b(m;n,p)$ is the central term, which is $O((2\pi npq)^{-\frac{1}{2}})$ (see Feller, 1957 p.140). Thus, when $k - np > n^d$, we have $\xi_k \ge (pq)^{-1/2}[n^{d-1/2} + (1/2-p)n^{-1/2}]$ and

$$\sum_{k>np+n^d} n^{\frac{16(k-np)^2}{n^2(1-p)^2}} b(k;n,p) = O\Big(nn^{16}b(m;n,p)e^{-\frac{1}{2}q^{-1}n^{2d-1}}\Big) = O(1).$$

By applying the same argument to the $k < np - n^d$, we have $J_2 = O(1)$. Thus we prove (1.5.7) and hence (1.5.6).

As to the (1.5.5), in view of (1.4.8), we write C_{1n} , conditionally on $\{Z_1 = z_1\}$ and $\{\delta_1 = d_1\}$, as

$$\begin{aligned} C_{1n} =& \frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < z_1, \delta_k = 0)}{1-H_n(Z_k)} - \frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < z_1, \delta_k = 0)}{1-H(Z_k)} \\ &+ \frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < z_1, \delta_k = 0)}{1-H(Z_k)} - \int_{-\infty}^{z_1 -} \frac{\tilde{H}^0(dz)}{1-H(z)} \\ =& \frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < z_1, \delta_k = 0)[H_n(Z_k) - H(Z_k)]}{[1-H_n(Z_k)][1-H(Z_k)]} \\ &+ \frac{1}{n}\sum_{k=1}^{n} \left[\frac{I(Z_k < z_1, \delta_k = 0)}{1-H(Z_k)} - \int_{-\infty}^{z_1 -} \frac{\tilde{H}^0(dz)}{1-H(z)}\right] \\ =&: D_{1n} + D_{2n}. \end{aligned} \tag{1.5.8}$$

Applying the moment inequality to D_{1n} and taking expectation yields, conditionally on $\{Z_1 = z_1\}$ and $\{\delta_1 = d_1\}$,

$$\begin{aligned} E(D_{1n}^8 | Z_1 = z_1, \delta_1 = d_1) &\le E\bigg(\frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < z_1)[H_n(Z_k) - H(Z_k)]^8}{[1-H_n(Z_k)]^8[1-H(Z_k)]^8}\bigg| Z_1 = z_1, \delta_1 = d_1\bigg) \\ &\le CE\bigg(\frac{I(Z_2 < z_1)[H_n(Z_2) - H(Z_2)]^8}{[1-H_n(Z_2)]^8}\bigg| Z_1 = z_1, \delta_1 = d_1\bigg). \end{aligned}$$

Conditionally on $\{Z_1 = z_1\}$, $\{\delta_1 = d_1\}$, $\{Z_2 = z_2\}$ and $\{\delta_2 = d_2\}$, $nH_n(z_2) := 1+Q$, Q is a binomial random variable with parameter $n-2$ and $p := H(z_2)$. Thus, through directly calculation, we have

$$E\left(\frac{I(Z_2 < z_1)[H_n(z_2) - H(z_2)]^8}{[1 - H_n(z_2)]^8}\middle| Z_1 = z_1, \delta_1 = d_1, Z_2 = z_2, \delta_2 = d_2\right)$$
$$\leq \sum_{k=0}^{n-2} \left(\frac{1 + k - np}{n - 1 - k}\right)^8 \binom{n-2}{k} p^k (1-p)^{n-2-k} \leq C\frac{1}{n^4}.$$

Thus, uniformly in z_1, d_1,

$$E(D_{1n}^8 | Z_1 = z_1, \delta_1 = d_1) = O\left(\frac{1}{n^4}\right). \tag{1.5.9}$$

As to the D_{2n}, which is the sum of i.i.d. centered bounded random variables. By Rosenthal's inequality, it is easy to show

$$E(D_{2n}^8 | Z_1 = z_1, \delta_1 = d_1) = O\left(\frac{1}{n^4}\right). \tag{1.5.10}$$

Combining (1.5.8), (1.5.9) and (1.5.10), we proved (1.5.5). Together with (1.5.4), we proved (1.5.3). Thus Proposition 1.4.2 follows.

Proof of the Proposition 1.4.3. In view of (1.4.10),

$$R_{n1,ij} = \int \varphi_{ij}(w)\gamma_0(w)\frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < w)(1-\delta_k)[H_n(Z_k) - H(Z_k)]^2}{[1 - H(Z_k)]^2[1 - H_n(Z_k)]}\tilde{H}_n^1(dw).$$

So

$$R_{n1,ij}^2 \leq \int \varphi_{ij}^2(w)\gamma_0^2(w)\left(\frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < w)[H_n(Z_k) - H(Z_k)]^2}{[1 - H(Z_k)]^2[1 - H_n(Z_k)]}\right)^2 \tilde{H}_n^1(dw).$$
$$\leq \int \varphi_{ij}^2(w)\gamma_0^2(w)\frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < w)[H_n(Z_k) - H(Z_k)]^4}{[1 - H(Z_k)]^4[1 - H_n(Z_k)]^2}\tilde{H}_n^1(dw).$$
$$= \frac{1}{n}\sum_{l=1}^{n} \varphi_{ij}^2(Z_l)\gamma_0^2(Z_l)\delta_l\frac{1}{n}\sum_{k=1}^{n} \frac{I(Z_k < Z_l)[H_n(Z_k) - H(Z_k)]^4}{[1 - H(Z_k)]^4[1 - H_n(Z_k)]^2}.$$

Thus

$$E(R^2_{n1,ij}) \le E\Big(\varphi^2_{ij}(Z_1)\gamma^2_0(Z_1)\delta_1 E\Big(\frac{1}{n}\sum_{k=1}^{n}\frac{I(Z_k < Z_1)[H_n(Z_k) - H(Z_k)]^4}{[1 - H(Z_k)]^4[1 - H_n(Z_k)]^2}\Big| Z_1, \delta_1\Big)\Big). \tag{1.5.11}$$

By applying an argument similar to that in Proposition 1.4.2, we can show, conditionally on $\{Z_1 = z_1\}$ and $\{\delta_1 = d_1\}$,

$$E\Big(\frac{1}{n}\sum_{k=1}^{n}\frac{I(Z_k < z_1)[H_n(Z_k) - H(Z_k)]^4}{[1 - H(Z_k)]^4[1 - H_n(Z_k)]^2}\Big| Z_1 = z_1, \delta_1 = d_1\Big) = O\left(\frac{1}{n^2}\right),$$

uniformly in z_1 and d_1. Thus by (1.5.11), the proposition follows.

Proof of the Proposition 1.4.4. In view of (1.4.15), we write $R_{n2,ij}$ as

$$R_{n2,ij} = R_{n2,ij}(1) + R_{n2,ij}(2) + R_{n2,ij}(3).$$

Because of the similarity of the proof, we here only provide the details for the first term. We write

$$R_{n2,ij}(1) = \int K(u)H_n(du),$$

where

$$K(u) = \iint \frac{I(v < u \wedge w)\varphi_{ij}(w)\gamma_0(w)}{[1 - H(v)]^2}[\tilde{H}^0_n(dv) - \tilde{H}^0(dv)][\tilde{H}^1_n(dw) - \tilde{H}^1(dw)].$$

Hence

$$E(R^2_{n2,ij}(1)) \le E(K^2(Z_1)) = E\Big(E(K^2(Z_1)|Z_1)\Big). \tag{1.5.12}$$

Conditionally on $\{Z_1 = z_1\}$ and $\{\delta_1 = d_1\}$, we rewrite $K(z_1)$ as sum, we have

$$K(z_1) = \frac{1}{n}\sum_{k=2}^{n}\int [h(Z_k, w) - u(w)][\tilde{H}^1_n(dw) - \tilde{H}^1(dw)],$$

where

$$h(Z_k, w) = \frac{I(Z_k < z_1 \wedge w)\varphi_{ij}(w)\gamma_0(w)(1-\delta_k)}{[1-H(Z_k)]^2},$$

$$u(w) = E[h(Z_k, w)], \quad k = 2, 3, \cdots, n. \tag{1.5.13}$$

Again continuing to write $K(z_1)$ as sum, we have

$$\begin{aligned}
K(z_1) =& \frac{1}{n}\sum_{k=2}^{n}\left\{\frac{1}{n}\sum_{l=1}^{n}\left[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l\right] - \int [h(Z_k, w) - u(w)]\tilde{H}^1(dw)\right\} \\
=& \frac{1}{n}\sum_{k=2}^{n}\left\{\frac{1}{n}\sum_{l\neq k}^{n}\left[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l\right] - \frac{1}{n}u(Z_k)\delta_k\right. \\
&\left. - \int [h(Z_k, w) - u(w)]\tilde{H}^1(dw)\right\} \\
=& -\frac{1}{n^2}\sum_{k=2}^{n}u(Z_k)\delta_k + \frac{1}{n}\sum_{k=2}^{n}\left\{\frac{1}{n}\sum_{l\neq k}^{n}[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l]\right. \\
&\left. - \int [h(Z_k, w) - u(w)]\tilde{H}^1(dw)\right\} \\
=& -\frac{1}{n^2}\sum_{k=2}^{n}u(Z_k)\delta_k - \frac{1}{n^2(n-1)}\sum_{k=2}^{n}\sum_{l\neq k}^{n}\left[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l\right] \\
&+ \frac{1}{n(n-1)}\sum_{k=2}^{n}\sum_{l\neq k}^{n}\left\{[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l] - \int [h(Z_k, w) - u(w)]\tilde{H}^1(dw)\right\} \\
=& I_1 + I_2 + I_3, \ (say).
\end{aligned}$$

As to the I_1, in view of (1.5.13), we have

$$EI_1^2 = O\left(\frac{1}{n^2}\right)E(u^2(Z_2)\delta_2) = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF. \tag{1.5.14}$$

As to the I_2, in view of (1.5.13), we have

$$|I_2| = O\left(\frac{1}{n^3}\right)\sum_{k=2}^{n}\sum_{l\neq k}^{n}|\varphi_{ij}(Z_l)|\gamma_0(Z_l)\delta_l = O\left(\frac{1}{n^2}\right)\sum_{l=1}^{n}|\varphi_{ij}(Z_l)|\gamma_0(Z_l)\delta_l.$$

Thus

$$EI_2^2 = O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF. \tag{1.5.15}$$

As to the I_3, let $H(Z_k, Z_l) = [h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l] - \int [h(Z_k, w) - u(w)]\tilde{H}^1(dw)$, thus $I_3 = O(n^{-2}) \sum_{k=2}^n \sum_{l \neq k}^n H(Z_k, Z_l)$. Noticing

$$EH(Z_k, Z_l) = E(E(H(Z_k, Z_l)|Z_k)) = 0, \ k \neq l. \tag{1.5.16}$$

$$EH(z_k, Z_l) = 0, \quad EH(Z_k, z_l) = 0, \ k \neq l.$$

Hence

$$\begin{aligned} EI_3^2 =& O\left(\frac{1}{n^4}\right) \sum_{k=2}^n \sum_{l \neq k}^n EH^2(Z_k, Z_l) \\ &+ O\left(\frac{1}{n^4}\right) \sum_{k \neq l_1} \sum_{k \neq l_2} \sum_{l_1 \neq l_2} EH(Z_k, Z_{l_1})H(Z_k, Z_{l_2}) \\ &+ O\left(\frac{1}{n^4}\right) \sum_{k \neq l_1} \sum_{k \neq l_2} \sum_{l_1 \neq l_2} EH(Z_{l_1}, Z_k)H(Z_{l_2}, Z_k) \\ &+ O\left(\frac{1}{n^4}\right) \sum_{k \neq l_1} \sum_{k \neq l_2} \sum_{l_1 \neq l_2} EH(Z_k, Z_{l_1})H(Z_{l_2}, Z_k) \\ &+ O\left(\frac{1}{n^4}\right) \sum_{k_1 \neq l} \sum_{k_2 \neq l} \sum_{k_1 \neq k_2} EH(Z_{k_1}, Z_l)H(Z_l, Z_{k_2}) \\ =& I_3(1) + I_3(2) + I_3(3) + I_3(4) + I_3(5), \ (say). \end{aligned}$$

The first term

$$\begin{aligned} I_3(1) &= O\left(\frac{1}{n^4}\right) \sum_{k=2}^n \sum_{l \neq k}^n E\big(E(H^2(Z_k, Z_l)|Z_k)\big) \\ &= O\left(\frac{1}{n^4}\right) \sum_{k=2}^n \sum_{l \neq k}^n E[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l]^2 \\ &= O\left(\frac{1}{n^4}\right) \sum_{k=2}^n \sum_{l \neq k}^n E(h^2(Z_k, Z_l)\delta_l) \\ &= O\left(\frac{1}{n^4}\right) \sum_{k=2}^n \sum_{l \neq k}^n E\varphi_{ij}^2(Z_l)\gamma_0^2(Z_l)\delta_l \\ &= O\left(\frac{1}{n^2}\right) \int \varphi_{ij}^2 \, dF. \end{aligned}$$

As to the I_3, let $H(Z_k, Z_l) = [h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l] - \int[h(Z_k, w) - u(w)]\tilde{H}^1(dw)$, thus $I_3 = O(n^{-2})\sum_{k=2}^n \sum_{l\neq k}^n H(Z_k, Z_l)$. Noticing

$$EH(Z_k, Z_l) = E(E(H(Z_k, Z_l)|Z_k)) = 0,\ k \neq l. \tag{1.5.16}$$

$$EH(z_k, Z_l) = 0, \quad EH(Z_k, z_l) = 0,\ k \neq l.$$

Hence

$$\begin{aligned} EI_3^2 =& O\left(\frac{1}{n^4}\right)\sum_{k=2}^n \sum_{l\neq k}^n EH^2(Z_k, Z_l) \\ &+ O\left(\frac{1}{n^4}\right)\sum_{k\neq l_1}\sum_{k\neq l_2}\sum_{l_1\neq l_2} EH(Z_k, Z_{l_1})H(Z_k, Z_{l_2}) \\ &+ O\left(\frac{1}{n^4}\right)\sum_{k\neq l_1}\sum_{k\neq l_2}\sum_{l_1\neq l_2} EH(Z_{l_1}, Z_k)H(Z_{l_2}, Z_k) \\ &+ O\left(\frac{1}{n^4}\right)\sum_{k\neq l_1}\sum_{k\neq l_2}\sum_{l_1\neq l_2} EH(Z_k, Z_{l_1})H(Z_{l_2}, Z_k) \\ &+ O\left(\frac{1}{n^4}\right)\sum_{k_1\neq l}\sum_{k_2\neq l}\sum_{k_1\neq k_2} EH(Z_{k_1}, Z_l)H(Z_l, Z_{k_2}) \\ =& I_3(1) + I_3(2) + I_3(3) + I_3(4) + I_3(5),\ (say). \end{aligned}$$

The first term

$$\begin{aligned} I_3(1) &= O\left(\frac{1}{n^4}\right)\sum_{k=2}^n \sum_{l\neq k}^n E\big(E(H^2(Z_k, Z_l)|Z_k)\big) \\ &= O\left(\frac{1}{n^4}\right)\sum_{k=2}^n \sum_{l\neq k}^n E[h(Z_k, Z_l)\delta_l - u(Z_l)\delta_l]^2 \\ &= O\left(\frac{1}{n^4}\right)\sum_{k=2}^n \sum_{l\neq k}^n E(h^2(Z_k, Z_l)\delta_l) \\ &= O\left(\frac{1}{n^4}\right)\sum_{k=2}^n \sum_{l\neq k}^n E\varphi_{ij}^2(Z_l)\gamma_0^2(Z_l)\delta_l \\ &= O\left(\frac{1}{n^2}\right)\int \varphi_{ij}^2\, dF. \end{aligned}$$

As to the $I_3(2)$, for $k \neq l_1$, $k \neq l_2$ and $l_1 \neq l_2$, conditionally on $\{Z_k = z_k\}$ and $\{\delta_k = d_k\}$, we have

$$E\big(H(Z_k, Z_{l_1})H(Z_k, Z_{l_2})|z_k, d_k\big) = EH(z_k, Z_{l_1})\,EH(z_k, Z_{l_2}) = 0,$$

which is from (1.5.16). Thus $I_3(2) = 0$. By applying the same argument, we have $I_3(3) = I_3(4) = I_3(5) = 0$. Thus

$$EI_3^2 = O\left(\frac{1}{n^2}\right) \int \varphi_{ij}^2\, dF. \tag{1.5.17}$$

Together with (1.5.14), (1.5.15) and (1.5.17), we deduce $EK^2(z_1) = O(n^{-2}) \int \varphi_{ij}^2\, dF$. From (1.5.12), we finally obtain $ER_{n2,ij}^2(1) = O(n^{-2}) \int \varphi_{ij}^2\, dF$.

Proof of the Proposition 1.4.5. The proof is basically the same as the previous proposition.

Chapter 2

Nonlinear Wavelet-based Hazard Rate Estimator

2.1 Introduction

In this chapter, we consider the same setting of survival analysis with random censorship as that in Chapter 1, with the extra assumption that random variables X and Y are nonnegative. Our goal is to estimate the hazard rate function $\lambda(x)$ with censored data,

$$\lambda(x) = \lim_{\epsilon \to 0^+} \frac{P(x \leq X < x + \epsilon | X \geq x)}{\epsilon} = \frac{f(x)}{1 - F(x-)}, \qquad x \in (0, \infty).$$

There is an extensive literature avaliable on estimating $\lambda(x)$ from censored data, see e.g., the survey paper Singpurwalla and Wong (1983) and the review paper Padgett and McNichols (1984). Tanner and Wong (1983) and Lo, at al. (1983) studied a kernel estimation of density and hazard rate under random censorship and provided Mean Square Error (MSE) and asymptotic normality of hazard rate estimators.

The objective of this chapter, like that in the previous, is to provide a nonlinear wavelet-based hazard rate estimator for randomly censored data, its asymptotic formula for MISE and its asymptotic normality. We show this MISE formula, when the underlying survival density function and censoring distribution function are only piecewise smooth, has the analogous expansion for the kernel estimators. However, as to the kernel estimators, this MISE formula holds only under the smoothness assumption.

In the next section, we give the elements of wavelet transform and provide nonlinear wavelet-based hazard rate estimators. The main results are described in Section 3, while their proofs appear in Section 4.

2.2 Notations and Estimators

As that in Chapter 1, let $T < \tau_H$ be fixed and $\lambda_1(x) = \lambda(x)I(x \leq T)$. Since, in general, hazard rate function $\lambda(x)$ is not square integrable, we estimate $\lambda_1(x)$, i.e. hazard rate function $\lambda(x)$ for $x \in (0, T]$. Like in Section 1.2. the wavelet expansion of $\lambda_1(x)$ is

$$\begin{aligned} \lambda_1(x) &= \sum_{j=-\infty}^{\infty} b_j\phi_j(x) + \sum_{i=0}^{\infty}\sum_{j=-\infty}^{\infty} b_{ij}\psi_{ij}(x), \\ b_j &= \int \lambda_1\phi_j, \quad b_{ij} = \int \lambda_1\psi_{ij}. \end{aligned} \tag{2.2.1}$$

We propose a nonlinear wavelet estimator of $\lambda_1(x)$:

$$\hat{\lambda}_1(x) = \sum_{j=-\infty}^{\infty} \hat{b}_j \phi_j(x) + \sum_{i=0}^{q-1} \sum_{j=-\infty}^{\infty} \hat{b}_{ij} I(|\hat{b}_{ij}| > \delta) \psi_{ij}(x), \tag{2.2.2}$$

where now the wavelet coefficients $\hat{b}_j$ and $\hat{b}_{ij}$ are defined as follows:

$$\begin{aligned}
\hat{b}_j &= \int \phi_j(x) I(x \leq T) \frac{d\hat{F}_n(x)}{1 - \hat{F}_n(x-)} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k I(Z_k \leq T) \phi_j(Z_k)}{[1 - \hat{F}_n(Z_k-)][1 - \hat{G}_n(Z_k-)]}, \\
\hat{b}_{ij} &= \int \psi_{ij}(x) I(x \leq T) \frac{d\hat{F}_n(x)}{1 - \hat{F}_n(x-)} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k I(Z_k \leq T) \psi_{ij}(Z_k)}{[1 - \hat{F}_n(Z_k-)][1 - \hat{G}_n(Z_k-)]},
\end{aligned}$$

where $\hat{F}_n$ and $\hat{G}_n$ are the Kaplan-Meier estimators of distribution function F and G, respectively. For more details, see section 1.2.

2.3 Main results

Theorem 2.3.1 *In addition to the conditions on ϕ and ψ stated in Section 1.2, assume that the r-th derivative $\lambda_1^{(r)}$ is bounded and continuous on $(0, T]$, the censoring distribution G is continuous. Also assume that condition (A) holds with $C_0 \equiv 2\{r(2r+1)^{-1} \sup \lambda_1 (1-H)^{-1}\}^{1/2}$. Then*

$$\begin{aligned}
&E\left| \int (\hat{\lambda}_1 - \lambda_1)^2 - \left\{ n^{-1} p \int \frac{\lambda_1}{1-H} + p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int \lambda_1^{(r)2} \right\} \right| \\
&\qquad = o(n^{-1} p + p^{-2r}).
\end{aligned} \tag{2.3.1}$$

Remark 2.3.1 This result is stronger than traditional asymptotic formula for MISE. It implies a wavelet version of the MISE formula:

$$E \int (\hat{\lambda}_1 - \lambda_1)^2 \sim n^{-1} p \int \frac{\lambda_1}{1-H} + p^{-2r} \kappa^2 (1 - 2^{-2r})^{-1} \int \lambda_1^{(r)2}.$$

Remark 2.3.2 The truncation parameter q and threshold parameter δ are chosen to ensure that $\hat{\lambda}_1$ is very close to λ_1 in case where λ_1 is smooth, yet at the same time provide sufficient local adaptability to produce automatic incorporation of appropriate wavelet terms $\hat{b}_{ij}\psi_{ij}$ in place where λ_1 is not smooth. For the details about how to choose these smoothing parameters, see Hall and Patil (1995).

In the Theorem 2.3.1, we have assumed that hazard rate function λ_1 is r-times continuous differentiable and the censoring distribution function G is continous for simplicity and convenience of exposition. However, if λ_1 and G are only piecewise smooth, the following theorem gives a similar result.

Theorem 2.3.2 *In addition to the conditions on ϕ and ψ stated in Section 1.2, assume that $\lambda_1^{(r)}$ and G are only piecewise smooth, i.e. there exist points $x_0 = 0 < x_1 < x_2 < \cdots < x_N < T = x_{N+1}$ such that the first r derivatives of λ_1 exist and are bounded and continuous on (x_i, x_{i+1}) for $0 \le i \le N$, with left- and right-hand limits; In particular, λ_1 and G themselves may be only piecewise continuous. Also assume that condition (A) holds with C_0 as in Theorem 2.3.1 and $p_q^{2r+1}n^{-2r} \to \infty$. Then (2.3.1) continues to hold.*

While wavelet estimators allow us to obtain MISE and optimal convergence rates analogous to kernel estimators under weaker assumption, there is a fundamental instability in the asymptotic variance of wavelet estimator caused by the lack of translation invariance of the wavelet transform. For more details, see Antoniadis, et al. (1994). Because wavelet estimators are only dyadic translation invariant, we provide an asymptotic expansion of the variance and asymptotic normality result at dyadic point $x = l/2^k$, k and l are integers.

Theorem 2.3.3 *In addition to the conditions on ϕ and ψ stated in Section 1.2, assume $\lambda_1(x)$ is r-times continuously differentiable at $x = l/2^k$, Also assume that*

$$p = 2^N = O(n^{1/(2r+1)}), \quad q \to \infty, \quad p_q\delta^2 \to 0, \quad \delta \geq C\sqrt{n^{-1}\ln n},$$

where $C > C_1 \equiv \{(8r+2)(2r+1)^{-1} \sup \lambda_1 (1-H)^{-1}\}^{1/2}$. Then

$$\sqrt{np^{-1}}(\hat{\lambda}_1(x) - \lambda_1(x) + b(x)) \overset{d}{\Longrightarrow} N(0, \sigma^2(x)),$$

where

$$b(x) = (r!)^{-1}\lambda_1^{(r)}(x) \int u^r \sum_l \phi(u+l)\phi(l)\, du \cdot p^{-r},$$

$$\sigma^2(x) = \frac{\lambda_1(x)}{1 - H(x)} \int \Big[\sum_l \phi(u+l)\phi(l)\Big]^2 du.$$

Remark 2.3.3 This result is analogous to Theorem 4.2 in Lo, et al. (1989), a result of asymptotic normality of kernel estimator of hazard rate function with censored data.

2.4 Proofs

The proofs of the Theorems 2.3.1 and 2.3.2 are very similar to those of Theorems 1.3.1 and 1.3.2. The difference is that we have one more term $1 - \hat{F}_n$ in the denominators of $\hat{b}_j$ and $\hat{b}_{ij}$, which create an extra technical difficulty. We begin with some lemmas. To state these, we need some additional notation. Let

$$\begin{aligned}
\bar{b}_j &= \frac{1}{n}\sum_{k=1}^{n} \frac{\delta_k \varphi_j(Z_k)}{[1-F(Z_k)][1-\hat{G}_n(Z_k-)]}, & \tilde{b}_j &= \frac{1}{n}\sum_{k=1}^{n} \frac{\delta_k \varphi_j(Z_k)}{[1-H(Z_k)]}, \\
\bar{b}_{ij} &= \frac{1}{n}\sum_{k=1}^{n} \frac{\delta_k \varphi_{ij}(Z_k)}{[1-F(Z_k)][1-\hat{G}_n(Z_k-)]}, & \tilde{b}_{ij} &= \frac{1}{n}\sum_{k=1}^{n} \frac{\delta_k \varphi_{ij}(Z_k)}{[1-H(Z_k)]}.
\end{aligned} \tag{2.4.1}$$

We also define

$$\overline{W}_j = \frac{1}{n}\sum_{k=1}^{n} W_j(Z_k), \qquad \overline{W}_{ij} = \frac{1}{n}\sum_{k=1}^{n} W_{ij}(Z_k),$$

$$W_j(Z_k) = U_j(Z_k) - V_j(Z_k), \qquad W_{ij}(Z_k) = U_{ij}(Z_k) - V_{ij}(Z_k)$$

and

$$\begin{aligned}
U_j(Z_k) &= \frac{1-\delta_k}{1-H(Z_k)} \int_{Z_k}^{\tau_H} \frac{\varphi_j(\omega)}{1-F(\omega)} F(d\omega), \\
U_{ij}(Z_k) &= \frac{1-\delta_k}{1-H(Z_k)} \int_{Z_k}^{\tau_H} \frac{\varphi_{ij}(\omega)}{1-F(\omega)} F(d\omega), \\
V_j(Z_k) &= \iint \frac{\varphi_j(\omega) I(v < Z_k \wedge \omega)}{[1-H(v)]^2} G(dv) F(d\omega), \\
V_{ij}(Z_k) &= \iint \frac{\varphi_{ij}(\omega) I(v < Z_k \wedge \omega)}{[1-H(v)]^2} G(dv) F(d\omega).
\end{aligned} \tag{2.4.2}$$

Remark 2.4.1 We use the same notation of the coefficents b_j, b_{ij}, their estimators $\hat{b}_j$, $\hat{b}_{ij}$ and $U_j(Z_k)$, $V_j(Z_k), W_j(Z_k)$, etc. as those in Chapter 1 for easy comparision. However here they all include one more term $1-\hat{F}_n$, or $1-F$ in their denominators.

Lemma 2.4.1 *Under the assumptions of Theorem 2.3.1, we have*

$$\begin{aligned}
\bar{b}_j &= \tilde{b}_j + \overline{W}_j + R_{n,j}, & E(R_{n,j}^2) &= O\left(\frac{1}{n^2}\right) \int \varphi_j^2\, dF, \\
\bar{b}_{ij} &= \tilde{b}_{ij} + \overline{W}_{ij} + R_{n,ij}, & E(R_{n,ij}^2) &= O\left(\frac{1}{n^2}\right) \int \varphi_{ij}^2\, dF.
\end{aligned} \tag{2.4.3}$$

Proof. The proof follows along the same lines as those in Lemma 1.4.1, use $\varphi_j/(1-F)$ and $\varphi_{ij}/(1-F)$ instead of φ_j and φ_{ij}. Because the denominators of $\bar{b}_j$ and $\bar{b}_{ij}$ are bounded away from zero, all needed conditions are satisfied.

Let

$$Q(t) = \frac{1}{n}\sum_{k=1}^{n} Q(Z_k, t), \qquad W(t) = \frac{1}{n}\sum_{k=1}^{n} W(Z_k, t)$$
$$Q(Z_k, t) = \frac{\delta_k I(Z_k \le t)}{1 - G(Z_k)}, \qquad W(Z_k, t) = U(Z_k, t) - V(Z_k, t) \tag{2.4.4}$$

and

$$U(Z_k, t) = \frac{1-\delta_k}{1 - H(Z_k)} \int I(Z_k < \omega \le t)\, F(d\omega),$$
$$V(Z_k, t) = \iint \frac{I(\omega \le t) I(v < Z_k \wedge \omega)}{[1 - H(v)][1 - G(v)]} G(dv) F(d\omega). \tag{2.4.5}$$

Lemma 2.4.2 *Under the assumptions of Theorem 2.3.1, for any $t \le T$, we have*

$$\hat{F}_n(t) = Q(t) + W(t) + R_n(t), \qquad \sup_{t \le T} E R_n^4(t) = O\left(n^{-4}\right). \tag{2.4.6}$$

Proof. From a result of Stute (1995, p.434) and details of the proof of the main theorem of Stute (1995), we have (2.4.6), an approximation of the Kaplan-Meier estimator when $\varphi(x) = I(x \le t)$. The rest of the proof is completely analogous to that of Lemma 1.4.1. Here we only consider the specific integrand $\varphi(x)$, instead of $\varphi_{ij}(x)$ there. Also we consider the fourth moment of $R_n(t)$, instead of the second moment there. However, from an argument similar to the one used in the proof of Lemma 1.4.1, we can obtain $ER_n^4(t) = O(n^{-4}) \int \varphi^4(x)\, dF$. Thus $\sup_{t \le T} ER_n^4(t) = O(n^{-4})$.

Remark 2.4.2 An argument similar to the one used in the proof of the above lemma, we can show that, for any integer $k \ge 1$, $\sup_{t \le T} ER_n^{2k}(t) = O(n^{-2k})$. Thus, from Hölder's inequality, $\sup_{t \le T} E|R_n^\alpha(t)| = O(n^{-\alpha})$, for any $\alpha \ge 1$. This fact is analogous to Lemma 2.1 in Lo, et al.(1989), where they show $\sup_{t \le T} E|R_n^\alpha(t)| = O([\ln n/n]^\alpha)$. Because of the technical reasons, they defined estimator $\hat{F}_n$ slightly different from here. In addition, by Theorem 1.1 of Stute (1995), where F and G are only piecewise continuous, Lemma 2.4.2 still holds, which will be used in proving Theorem 2.3.2.

Lemma 2.4.3 *Under the assumptions of Theorem 2.3.1, we have*

$$E\Big\{\sum_j (\hat{b}_j - \bar{b}_j)^2\Big\} = o(n^{-1}p).$$

Proof. In view of (2.2.3) and (2.4.1),

$$\hat{b}_j - \bar{b}_j = \frac{1}{n}\sum_{k=1}^{n} \frac{\delta_k \varphi_j(Z_k)}{1 - F(Z_k)} \frac{\hat{F}_n(Z_k-) - F(Z_k)}{[1 - \hat{F}_n(Z_k-)][1 - \hat{G}_n(Z_k-)]}.$$

From the definition of $\hat{F}_n$ and $\hat{G}_n$, we have

$$[1 - \hat{F}_n(Z_k-)][1 - \hat{G}_n(Z_k-)] = \frac{n + 1 - \text{Rank}Z_k}{n} = \frac{n - \sum_{l \neq k} I(Z_l \leq Z_k)}{n}.$$

Apply Lemma 2.4.2, and use the continuity of F to obtain

$$\begin{aligned} \hat{b}_j - \bar{b}_j &= \frac{1}{n}\sum_{k=1}^{n} A_j(Z_k)B(Z_k)P(Z_k) + \frac{1}{n}\sum_{k=1}^{n} A_j(Z_k)B(Z_k)W(Z_k) \\ &\quad + \frac{1}{n}\sum_{k=1}^{n} A_j(Z_k)B(Z_k)R_n(Z_k) \\ &= I_{1j} + I_{2j} + I_{3j}, \quad (say), \end{aligned}$$

where

$$A_j(Z_k) := \frac{\delta_k \varphi_j(Z_k)}{1 - F(Z_k)}, \qquad B(Z_k) := \frac{n}{n - \sum_{l \neq k} I(Z_l \leq Z_k)},$$

$$P(Z_k) := Q(Z_k) - F(Z_k),$$

$Q(Z_k)$, $W(Z_k)$ and $R_n(Z_k)$ are defined as in (2.4.4) and (2.4.5). Conditionally on $\{Z_1 = z_1\}$, $\{\delta_1 = d_1\}$, through direct calculations, we have

$$EP^4(z_1) = O(n^{-2}), \quad \text{uniformly in } d_1 \text{ and } z_1 \leq T. \tag{2.4.7}$$

Conditionally on $\{Z_1 = z_1\}$, $\{\delta_1 = d_1\}$, $B(z_1) = n/(n-V)$, where $V = \sum_{l=2}^{n} I(Z_l \leq z_1)$ is a binomial random variable with parameter $n-1$ and $p := H(z_1)$. Thus, through direct calculations, we have

$$EB^4(z_1) = O(1), \quad \text{uniformly in } d_1 \text{ and } z_1 \leq T. \tag{2.4.8}$$

Now,

$$\begin{aligned} I_{1j}^2 = {} & \frac{1}{n^2}\sum_{k=1}^{n} A_j^2(Z_k)B^2(Z_k)P^2(Z_k) \\ & + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1,\neq k}^{n} A_j(Z_k)B(Z_k)P(Z_k)A_j(Z_l)B(Z_l)P(Z_l) \\ = {} & I_{1j}^2(1) + I_{1j}^2(2), \ (say). \end{aligned}$$

The first term

$$\begin{aligned} EI_{1j}^2(1) &= \frac{1}{n}\,E\Big[A_j^2(Z_1)B^2(Z_1)P^2(Z_1)\Big] \\ &= O\Big(\frac{1}{n}\Big)E\Big[A_j^2(Z_1)E^{1/2}\Big(B^4(Z_1)\Big|Z_1,\delta_1\Big)\,E^{1/2}\Big(P^4(Z_1)\Big|Z_1,\delta_1\Big)\Big] \\ &= O\Big(\frac{1}{n^2}\Big)EA_j^2(Z_1) = O\Big(\frac{1}{n^2}\Big)\int \varphi_j^2 dF. \end{aligned} \tag{2.4.9}$$

Hence, from (1.4.19), we obtain

$$E\sum_j I_{1j}^2(1) = o(n^{-1}p). \tag{2.4.10}$$

The second term

$$\begin{aligned} & EI_{1j}^2(2) \\ & = O\left(\frac{1}{n^2}\right)\sum_{k=1}^{n}\sum_{l=1,\neq k}^{n} \\ & \quad E\Big[|A_j(Z_k)|\,|A_j(Z_l)|\,E\Big(|B(Z_k)|\,|B(Z_l)|\,|P(Z_k)|\,|P(Z_l)|\,\Big|Z_k, Z_l, \delta_k, \delta_l\Big)\Big]. \end{aligned}$$

Conditionally on $\{Z_k = z_k\}$, $\{\delta_k = d_k\}$, $\{Z_l = z_l\}$ and $\{\delta_l = d_l\}$, by the Cauchy-Schwarz inequality, we have

$$E\Big(|B(z_k)|\,|B(z_l)|\,|P(z_k)|\,|P(z_l)|\Big) \le \Big[EB^4(z_k)EB^4(z_l)EP^4(z_k)EP^4(z_l)\Big]^{1/4}.$$

Through direct calculations as that in (2.4.7) and (2.4.8), we have

$$\begin{aligned} EI_{1j}^2(2) &= O\left(\frac{1}{n^3}\right)\sum_{k=1}^{n}\sum_{l=1,\neq k}^{n} E\,|A_j(Z_k)|\,|A_j(Z_l)| \\ &= O\left(\frac{1}{n}\right)\left(\int |\varphi_j|dF\right)^2. \end{aligned} \tag{2.4.11}$$

Hence

$$E{\sum_j}' I_{1j}^2(2) = O\Big(\frac{1}{n}\Big)\sum_j\Big(\int |\varphi_j|dF\Big)^2 = O\Big(\frac{1}{n}\Big) = o(n^{-1}p) \tag{2.4.12}$$

by $\sum_j(\int|\varphi_j|dF)^2 < \infty$ and $p \to \infty$. This, together with (2.4.10), we have $E\sum_j I_{1j}^2 = o(n^{-1}p)$. Apply the previous same lines in I_{1j} to I_{2j}, we can show $E\sum_j I_{2j}^2 = o(n^{-1}p)$ too.

To complete the proof the lemma, it remains to show that $E\sum_j I_{3j}^2 = o(n^{-1}p)$. Apply the moment inequality to I_{3j}, we have $EI_{3j}^2 \le E[A_j^2(Z_1)B^2(Z_1)R_n^2(Z_1)]$. Conditionally on $\{Z_1 = z_1\}$, $\{\delta_1 = d_1\}$, by the Cauchy-Schwarz inequality, we have

$$E[B^2(z_1)R_n^2(z_1)] \le [EB^4(z_1)ER_n^4(z_1)]^{1/2} = O(1/n^2)$$

by (2.4.6) and (2.4.8). Thus

$$\begin{aligned} E\sum_j I_{3j}^2 &= O\left(\frac{1}{n^2}\right)\sum_j\int\varphi_j^2dF \\ &= O\left(\frac{p}{n^2}\right)\sum_j p^{-1}\int\varphi_j^2dF \\ &= o(n^{-1}p). \end{aligned} \tag{2.4.13}$$

Lemma 2.4.4 *Under the assumptions of Theorem 2.3.1, we have*

$$s_1 \equiv E\left|\sum_j (\hat{b}_j - b_j)^2 - n^{-1}p\int \frac{\lambda_1}{1-H}\right| = o(n^{-1}p).$$

Proof. The proof is similar to that in Lemma 1.4.2, but here we need one more step to approximate $\hat{b}_j$. In view of (2.4.3) and

$$\sum_j (\hat{b}_j - b_j)^2 = \sum_j (\bar{b}_j - b_j)^2 + \sum_j (\hat{b}_j - \bar{b}_j)^2 + 2\sum_j (\bar{b}_j - b_j)(\hat{b}_j - \bar{b}_j),$$

we have

$$\begin{aligned}
s_1 \le\ & E\left|\sum_j (\tilde{b}_j - b_j)^2 - n^{-1}p\int \frac{\lambda_1}{1-H}\right| + E\sum_j \overline{W}_j^2 + E\sum_j R_{n,j}^2 \\
& + 2E\sum_j |\tilde{b}_j - b_j||R_{n,j}| + 2E\sum_j |\tilde{b}_j - b_j||\overline{W}_j| + 2E\sum_j |R_{n,j}||\overline{W}_j| \\
& + E\sum_j (\hat{b}_j - \bar{b}_j)^2 + 2E\sum_j |\bar{b}_j - b_j||\hat{b}_j - \bar{b}_j| \\
=\ & I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8, \quad (say).
\end{aligned}$$

Notice that

$$nE(\tilde{b}_j - b_j)^2 = \int \phi_j^2(x)\frac{\lambda_1(x)}{1-H(x)}\,dx - b_j^2,$$

we have

$$\sum_j E(\tilde{b}_j - b_j)^2 = \frac{p}{n}\int \phi^2(y)\sum_j p^{-1}\frac{\lambda_1((y+j)/p)}{1-H((y+j)/p)}\,dy - \frac{1}{n}\sum_j b_j^2,$$

since $\int \phi^2 = 1$,

$$\sum_j p^{-1}\lambda_1((y+j)/p)/(1-H((y+j)/p)) \to \int \lambda_1/(1-H),$$

$\sum_j b_j^2 = O(\int \lambda_1^2)$, it follows that $E\sum_j(\tilde{b}_j - b_j)^2 = n^{-1}p\int \lambda_1/(1-H) + o(n^{-1}p)$. Because the denominator appearing in $\tilde{b}_j$ is bounded away from below. Thus these

$\tilde{b}_j$ may be handled along the the same lines as those of Hall and Patil (1995, p.922) to show that $\mathrm{Var}\{\sum_j(\tilde{b}_j - b_j)^2\} = o(n^{-2}p^2)$. So we obtain $I_1 = o(n^{-1}p)$. By Lemma 2.4.1,

$$I_2 = n^{-1}\sum_j EW_j^2(Z_1) \le 2n^{-1}\sum_j (EU_j^2(Z_1) + EV_j^2(Z_1)).$$

By direct calculation, notice all denominators are bounded away from below, we have

$$EU_j^2(Z_1) = EV_j^2(Z_1) = O\left(p^{-1}\int \phi^2(u)\lambda_1^2\left(\frac{u+j}{p}\right)du\right).$$

Thus, $I_2 = O(n^{-1})\int \phi^2(u)\sum_j p^{-1}\lambda_1^2((u+j)/p)du = o(n^{-1}p)$ by $\sum_j p^{-1}\lambda_1^2((u+j)/p) \to \int \lambda_1^2 < \infty$ and $p \to \infty$.

By Lemma 2.4.1 or (2.4.3), we have $I_3 = o(n^{-1}p)$. From Lemma 2.4.3, $I_7 = o(n^{-1}p)$. Applying the Cauchy-Schwarz inequality to the rest terms, we complete the proof.

Lemma 2.4.5 *Under the assumptions of Theorem 2.3.1, we have*

$$s_2 \equiv \sum_{i=0}^{q-1}\sum_j E\Big\{(\hat{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta)\Big\} = o(n^{-2r/(2r+1)}).$$

Proof. Let α and β denote positive numbers satisfying $\alpha + \beta = 1$, we have

$$\begin{aligned}
s_2 &\le 2\sum_{i=0}^{q-1}\sum_j E\{(\hat{b}_{ij} - \bar{b}_{ij})^2 I(|\hat{b}_{ij}| > \delta)\} + 2\sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij}| > \delta)\} \\
&\le 2\sum_{i=0}^{q-1}\sum_j E\{(\hat{b}_{ij} - \bar{b}_{ij})^2\} + 2\sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\bar{b}_{ij}| > \alpha\delta)\} \\
&\quad + 2\sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij} - \bar{b}_{ij}| > \beta\delta)\} \\
&= 2(s_{21} + s_{22} + s_{23}), \quad (say).
\end{aligned}$$

Apply the argument analogous to (2.4.9), (2.4.11) and (2.4.13) appearing in the proof of Lemma 2.4.3 to s_{21}, we conclude that

$$\begin{aligned}
s_{21} &= O\left(\frac{1}{n^2}\right)\sum_{i=0}^{q-1}\sum_j \int \varphi_{ij}^2 dF + O\left(\frac{1}{n}\right)\sum_{i=0}^{q-1}\sum_j \left(\int |\varphi_{ij}| dF\right)^2 \\
&= O\left(\frac{1}{n^2}\right)\sum_{i=0}^{q-1} p_i \sum_j p_i^{-1} \int \varphi_{ij}^2 dF + O\left(\frac{1}{n}\right)\sum_{i=0}^{q-1}\sum_j \left(\int |\varphi_{ij}| dF\right)^2 \\
&= O\left(\frac{p_q}{n^2}\right) + O\left(\frac{\ln n}{n}\right) \\
&= o(n^{-2r/(2r+1)}).
\end{aligned}$$

The third equality follows from $\sum_j p_i^{-1} \int \varphi_{ij}^2 dF < \infty$, $\sum_j (\int |\varphi_{ij}| dF)^2 < \infty$ and $q = O(\ln n)$, while the last equality follows from $n^{-1}p_q \to 0$. Apply the same argument as that in the proof of Lemma 1.4.3 to s_{22}, use $\bar{b}_{ij}$ instead of $\hat{b}_{ij}$ in there, we conclude that $s_{22} = o(n^{-2r/(2r+1)})$. Now let $A = \{|\bar{b}_{ij} - b_{ij}| > \delta\}$, then

$$\begin{aligned}
s_{23} &= \sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij} - \bar{b}_{ij}| > \beta\delta) I(A)\} \\
&\quad + \sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\hat{b}_{ij} - \bar{b}_{ij}| > \beta\delta) I(A^c)\} \\
&\le \sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\bar{b}_{ij} - b_{ij}| > \delta)\} + \sum_{i=0}^{q-1}\sum_j \delta^2 P(|\hat{b}_{ij} - \bar{b}_{ij}| > \beta\delta) \\
&\le \sum_{i=0}^{q-1}\sum_j E\{(\bar{b}_{ij} - b_{ij})^2 I(|\bar{b}_{ij} - b_{ij}| > \delta)\} + \sum_{i=0}^{q-1}\sum_j \beta^{-2} E(\hat{b}_{ij} - \bar{b}_{ij})^2 \\
&= s_{23}(1) + s_{23}(2), \quad (say),
\end{aligned}$$

where $s_{23}(1)$ is analogous to s_{22} in section 1.4, which is $o(n^{-2r/(2r+1)})$. While $s_{23}(2)$ is $O(s_{21})$, which is $o(n^{-2r/(2r+1)})$ too. Together with s_{21} and s_{22}, we prove the lemma.

Lemma 2.4.6 *Under the assumptions of Theorem 2.3.1, we have*

$$s_3 \equiv E\left|\sum_{i=0}^{q-1}\sum_j b_{ij}^2 I(|\hat{b}_{ij}| \le \delta) - p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int \lambda_1^{(r)^2}\right| = o(p^{-2r}).$$

Proof. The proof follows the same lines as that of Lemma 1.4.4.

Lemma 2.4.7 *Under the assumptions of Theorem 2.3.1, we have*

$$s_4 \equiv \sum_{i=q}^{\infty}\sum_j b_{ij}^2 = o(p^{-2r}).$$

Proof. The proof follows from the step 3 of Theorem 2.1 of Hall and Patil (1995).

We are now in the position to give the proof of the Theorem 2.3.1 and 2.3.2.

Proof of the Theorem 2.3.1. Observe that

$$E\left|\int(\hat{\lambda_1}-\lambda_1)^2 - \left\{n^{-1}p\int\frac{\lambda_1}{1-H} + p^{-2r}\kappa^2(1-2^{-2r})^{-1}\int\lambda_1^{(r)^2}\right\}\right|$$
$$\le s_1 + s_2 + s_3 + s_4.$$

Thus Lemma 2.4.4, 2.4.5, 2.4.6 and 2.4.7 together prove the Theorem 2.3.1.

Proof of the Theorem 2.3.2. The basic idea of the proof is similar to that of Theorem 1.3.2. We omit the details.

In the sequel we prove the Theorem 2.3.3. This will involve the following two lemmas.

Lemma 2.4.8 *Under the assumptions of Theorem 2.3.3, we have*

$$\sqrt{np^{-1}}\left(\sum_j(\tilde{b}_j - b_j)\phi_j(x)\right) \stackrel{d}{\Longrightarrow} N(0,\sigma^2(x)),$$

where

$$\sigma^2(x) = \frac{\lambda_1(x)}{1-H(x)}\int\Big[\sum_l \phi(u+l)\phi(l)\Big]^2 du.$$

Proof. In view of (2.4.1),

$$
\begin{aligned}
&\sqrt{np^{-1}}\Big(\sum_j(\tilde{b}_j - b_j)\phi_j(x)\Big)\\
&\quad= \sum_{k=1}^{n}\Big[\sqrt{\frac{p}{n}}\frac{\delta_k I(Z_k \le T)K(pZ_k, px)}{1-H(Z_k)} - \sqrt{\frac{p}{n}}\Big(\int \lambda_1(t)K(pt,px)dt\Big)\Big]\\
&\quad:= \sum_{k=1}^{n} V_{n,k}\,,
\end{aligned}
$$

where $K(t,x) = \sum_j \phi(t-j)\phi(x-j)$. For the wavelets in Section 1.2, the kernel $K(t,x)$ satisfies the moment condition (See Theorem 8.3 of Härdle, et al. 1998, p.95), i.e. $\int (t-x)^k K(t,x)\,dt = \delta_{0k}$, for $k = 0, 1, \cdots, r-1$. Notice (Z_k, δ_k) are *i.i.d.* for $k = 1, 2, \cdots, n$, $EV_{n,k} = 0$, and

$$
\begin{aligned}
EV_{n,k}^2 &= \frac{p}{n}\int \frac{\lambda_1(t)}{1-H(t)}K^2(pt,px)\,dt - \frac{p}{n}\Big(\int \lambda_1(t)K(pt,px)\,dt\Big)^2\\
&= \frac{1}{n}\int \frac{\lambda_1(x+u/p)}{1-H(x+u/p)}\Big[\sum_j \phi(u+px-j)\phi(px-j)\Big]^2 du\\
&\qquad - \frac{1}{np}\Big(\int \lambda_1(x+u/p)\sum_j \phi(u+px-j)\phi(px-j)\,du\Big)^2\\
&= \frac{1}{n}\int \frac{\lambda_1(x)}{1-H(x)}\Big[\sum_l \phi(u+l)\phi(l)\Big]^2 du + O(n^{-1}p^{-1}).
\end{aligned}
$$

The second equality follows by the change of variable, while the third equality by $p = 2^N$, $x = l/2^k$, $N \to \infty$ and the Taylor expansion. Thus $\sum_{k=1}^{n} EV_{n,k}^2 = \sigma^2(x) + O(p^{-1}) \to \sigma^2(x)$. In addition, $K(t,x)$ being uniformly bounded, we have $|V_{n,k}| \le c\sqrt{n^{-1}p} \to 0$, c is a positive constant. So for all $\epsilon > 0$, $\lim_{n\to\infty}\sum_{k=1}^{n} E(|V_{n,k}|^2; |V_{n,k}| > \epsilon) = 0$. Thus by Lindeberg-Feller CL's Theorem, the lemma follows.

Let

$$
J_6 \equiv \sum_{i=0}^{q-1}\sum_j \hat{b}_{ij}\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta).
$$

Lemma 2.4.9 *Under the assumptions of Theorem 2.3.3, $EJ_6^2 = o(n^{-1}p)$.*

Proof. In view of (2.4.3), write $\hat{b}_{ij}$ as following

$$\hat{b}_{ij} = \tilde{b}_{ij} + (\hat{b}_{ij} - \bar{b}_{ij}) + \overline{W}_{ij} + R_{n,ij}\,. \tag{2.4.14}$$

Then

$$\begin{aligned} J_6 = &\sum_{i=0}^{q-1}\sum_j \tilde{b}_{ij}\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) + \sum_{i=0}^{q-1}\sum_j (\hat{b}_{ij} - \bar{b}_{ij})\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) \\ &+ \sum_{i=0}^{q-1}\sum_j \overline{W}_{ij}\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) + \sum_{i=0}^{q-1}\sum_j R_{n,ij}\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) \end{aligned} \tag{2.4.15}$$

$$= I_1 + I_2 + I_3 + I_4,\ (say).$$

Because of the compact support of $\psi(x)$, for each i, there are only finite number of j such that $\psi_{ij}(x)$ are nonzero. So

$$\begin{aligned} EI_2^2 &= O(q)\sum_{i=0}^{q-1}\sum_j E(\hat{b}_{ij} - \bar{b}_{ij})^2\psi_{ij}^2(x) \\ &= O(q)\sum_{i=0}^{q-1}\sum_j \left[\frac{1}{n^2}\int \varphi_{ij}^2 dF + \frac{1}{n}\left(\int |\varphi_{ij}| dF\right)^2\right] p_i \\ &= O(q)\left[\frac{p_q}{n^2} + \frac{q}{n}\right] \\ &= o(n^{-1}p). \end{aligned}$$

In the above, the second equality follows from the argument similar to (2.4.9), (2.4.11) and (2.4.13) in Lemma 2.4.3, while the last equality is from $p_q n^{-1} \to 0$ and $q = O(\ln n)$. Similarly, we can show $EI_3^2 = EI_4^2 = o(n^{-1}p)$.

As to the first term of J_6,

$$I_1 = \sum_{i=0}^{q-1}\sum_j (\tilde{b}_{ij} - b_{ij})\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) + \sum_{i=0}^{q-1}\sum_j b_{ij}\psi_{ij}(x)I(|\hat{b}_{ij}| > \delta) = I_{11} + I_{12}.$$

Let α and β are positive numbers such that $\alpha + \beta = 1$, so

$$|I_{11}| \leq \sum_{i=0}^{q-1} \sum_{j} |\tilde{b}_{ij} - b_{ij}||\psi_{ij}(x)| I(|b_{ij}| > \alpha\delta)$$
$$+ \sum_{i=0}^{q-1} \sum_{j} |\tilde{b}_{ij} - b_{ij}||\psi_{ij}(x)| I(|\tilde{b}_{ij} - b_{ij}| > \beta\delta).$$

Because $\lambda_1(x)$ is r-times continuously differentiable at x, so $|b_{ij}| \leq cp_i^{-(r+1/2)}$, or $b_{ij}^2 \leq c^2 p_i^{-(2r+1)}$, c is a constant (see Hall and Patil, 1995, p.917). Notice $\delta^2 = O(\ln n/n)$, $p_i = p2^i$, $p = O(n^{1/(2r+1)})$, thus $I(|b_{ij}| > \delta) = 0$ for large n, hence the first term in the bound of I_{11} actually is zero for all sufficient large n. In view of (2.4.14), the leading term to approximate $\hat{b}_{ij}$ is $\tilde{b}_{ij}$. Apply a similar argument as in (2.4.15) to I_{11}, all the rest of the terms are of smaller order, we have

$$\begin{aligned}
EI_{11}^2 &= O(q) \sum_{i=0}^{q-1} \sum_{j} E(\tilde{b}_{ij} - b_{ij})^2 \psi_{ij}^2(x) I(|\tilde{b}_{ij} - b_{ij}| > \beta\delta) \\
&= O(q) \sum_{i=0}^{q-1} \sum_{j} E^{1/a}(\tilde{b}_{ij} - b_{ij})^{2a} \psi_{ij}^2(x) P^{1/b}(|\tilde{b}_{ij} - b_{ij}| > \beta\delta) \\
&= O(q) \sum_{i=0}^{q-1} \sum_{j} \frac{p_i}{n} p_i n^{-d} = O(q) p_q^2 n^{-d-1}, \quad \text{where} \quad d > \frac{4r+1}{2r+1} \\
&= o(n^{-2r/(2r+1)}) = o(n^{-1}p).
\end{aligned}$$

The second equality follows by Hölder's inequality, while the third equality by Rosenthal's and Bernstein's inequality and let $a \to \infty$, $b \to 1$ (see the details in Hall and Patil, 1995, p.917-918). The fifth equality follows by $n^{-1}p_q \to 0$. Apply the same argument to I_{12}, using $b_{ij}^2 \leq c^2 p_i^{-(2r+1)}$, we can show that $EI_{12}^2 = o(n^{-1}p)$ too, which proves the lemma.

Proof of Theorem 2.3.3 In view of (2.2.2), by analogous equality of (2.4.14) to $\hat{b}_j$ and the definition of J_6 in Lemma 2.4.9, we have

$$\begin{aligned}
\hat{\lambda_1}(x) - \lambda_1(x) - b(x) &= \sum_j (\hat{b}_j - b_j)\phi_j(x) + \Big[\sum_j b_j\phi_j(x) - \lambda_1(x) - b(x)\Big] + J_6 \\
&= \sum_j (\tilde{b}_j - b_j)\phi_j(x) + \sum_j (\hat{b}_j - \bar{b}_j)\phi_j(x) + \sum_j \overline{W}_j\phi_j(x) \\
&\quad + \sum_j R_{n,j}\phi_j(x) + \Big[\sum_j b_j\phi_j(x) - \lambda_1(x) - b(x)\Big] + J_6 \\
&= J_1 + J_2 + J_3 + J_4 + J_5 + J_6, \ (say).
\end{aligned}$$

By Lemma 2.4.8, $\sqrt{np^{-1}}J_1 \stackrel{d}{\Longrightarrow} N(0, \sigma^2(x))$. By Lemma 2.4.9, we have $\sqrt{np^{-1}}J_6 \stackrel{p}{\longrightarrow} 0$. The terms J_2, J_3 and J_4 are analogous to I_2, I_3 and I_4 in Lemma 2.4.9, so applying the same argument, we can show that $EJ_2^2 = EJ_3^2 = EJ_4^2 = o(n^{-1}p)$. Thus $\sqrt{np^{-1}}J_2 \stackrel{p}{\longrightarrow} 0$, same as J_3 and J_4. Hence, in order to prove the theorem, it suffices to show that $J_5 = o(p^{-r})$. Apply the same argument as in Lemma 2.4.8, using the moment condition of $K(t,x)$, it is easy to see

$$\begin{aligned}
J_5 &= \int [\lambda_1(t) - \lambda_1(x)]\, p\, K(pt, px)\, dt - b(x) \\
&= \int [\lambda_1(x + u/p) - \lambda_1(x)] K(px + u, px)\, du - b(x) \\
&= \int \sum_{k=1}^{r} \frac{\lambda_1^{(k)}(x)}{k!} \frac{u^k}{p^k} K(px + u, px)\, du + o(p^{-r}) - b(x) \\
&= \frac{\lambda_1^{(r)}(x)}{r!} \int u^r \sum_l \phi(u + l)\phi(l)\, du\, p^{-r} - b(x) + o(p^{-r}) \\
&= o(p^{-r}),
\end{aligned}$$

the last equality follows from the moment condition of $K(t,x)$, which proves the theorem.

Chapter 3

Minimum Distance Estimators in Regression Models under Long Memory

3.1 Introduction

The practice of obtaining estimators of parameters by minimizing a certain distance between some functions of observations and parameters has long been present in statistics. These estimators have many desirable properties, including consistency, asymptotic normality under weak assumptions and robustness against outlier in the errors. Koul and DeWet (1983) and Koul (1985a, b; 1986) pointed out the importance of this methodology in linear regression models, using certain weighted empirical processes that arise naturally in these models. For more details and references on this methodology, see the monograph by Koul (1992b).

Koul and Mukherjee (1993) extended the above results to linear regression models with long range dependent errors that are either Gaussian or subordinate to Gasussian. More specifically, they considered the multiple linear regression model

$$Y_{ni} = x'_{ni}\beta + \varepsilon_i, \quad \varepsilon_i = G(\eta_i), \quad i = 1, 2, \cdots, n,$$

where $\{x_{ni}, i \geq 1\}$ are known fixed constants, G is a measurable function from $\mathbb{R}$ to $\mathbb{R}$, $\{\eta_i, i \geq 1\}$ is a stationary, mean zero, unit variance Gaussion process with correlation $\rho(k) := E\eta_1\eta_{1+k} \sim k^{-\theta}L(k)$, $k \geq 1$, $0 < \theta < 1$, where L is a function of positive integers, slowly varying at infinity, and $L(k)$ is positive for large k. Thus $\sum_{k=1}^{\infty} \rho(k) = \infty$, implying the errors have long memory. For motivation and arguments in support of this Gaussian and/or Gaussian subordinated long memory error process, see Taqqu (1975), Dehling and Taqqu (1989) and a review paper by Beran (1992).

The other class of long memory process is of the moving average type. For more on their importance in economics and other sciences, see Robinson (1994), Beran (1994), and Baillie (1996). These processes include an important class of fractional ARIMA processes. For various theoretical results pertaining to the empirical processes of long memory moving averages, see Ho and Hsing (1996, 1997), Giraitis et al. (1996), Koul and Surgailis (1997, 2001b), Giraitis and Surgailis (1999), among others.

Because of the importance of multiple linear models with long memory moving average errors, and the desirable properties of the above mentioned minimum distance estimators, it is natural to investigate their properties under the long memory moving average errors. The objective of this paper is to obtain the asymptotic distribution of the m.d. estimators of regression parameter in multiple linear model with long memory moving average symmetric errors when the design variables are either known constants or i.i.d. random variables, independent of the errors. These results thus extend those of Koul (1985a,b) and Koul and Mukherjee (1993) to these models.

The rest of this chapter is organized as follows. Section 2 provides the m.d. estimators and their asymptotic normality under both fixed and i.i.d. random design cases, while their proofs appear in Section 3 and Section 4, respectively.

3.2 Main results

3.2.1 The case of non-random designs

Consider the linear regression model where one observes the response variable $\{Y_{ni}\}$, $1 \leq i \leq n$, satisfying

$$Y_{ni} = x'_{ni}\beta + \varepsilon_i, \quad 1 \leq i \leq n, \quad \beta \in \mathbb{R}^p. \tag{3.2.1}$$

Let X denote the $n \times p$ design matrix of known constants whose i^{th} row is x'_{ni}, $1 \leq i \leq n$. Here $\mathbb{R}^p$ denotes p-dimensional Euclidean space, $\mathbb{R} = \mathbb{R}^1$. In the sequel, for the sake of convenience, the dependence of various entities on n will not be exhibited. We assume the errors $\{\varepsilon_i, 1 \leq i \leq n\}$ to form a stationary moving average sequence,

$$\varepsilon_i = \sum_{k=1}^{\infty} b_k \zeta_{i-k}, \quad b_k \sim L_1(k)\, k^{-(1+\theta)/2}, \quad 0 < \theta < 1, \quad 1 \le i \le n, \tag{3.2.2}$$

with the common distribution function F, where ζ_s, $s \in \mathbb{Z}$ are i.i.d. standard random variables, symmetric around zero and L_1 is a slowly varying function at infinity. This implies that $\rho(k) = \mathrm{Cov}(\varepsilon_1, \varepsilon_{1+k}) = L(k)k^{-\theta}$, where $L(k) = C_\theta L_1^2(k)$, $C_\theta = 2(2-\theta)^{-1}(1-\theta)^{-1}\int_0^\infty (u+u^2)^{-(1+\theta)/2}\, du$, and hence the errors have long memory. We assume that ζ_0 in (3.2.2) satisfy the following conditions:

A.1 $\quad \left|Ee^{iu\zeta_0}\right| \le C\,(1+|u|)^{-\delta}, \quad$ for some $\quad C < \infty,\ \delta > 0,\ \forall u \in \mathbb{R}$.

A.2 $\quad E\left|\zeta_0\right|^3 < \infty$.

Giraitis et al. (1996, Lemma 1) proved that under the Condition A.1, the error distribution function F is infinitely differentiable. The assumption A.2 is a condition on the decreasing rate of its density function in the tails.

Now, let $\tau_n = L^{1/2}(n)\, n^{(1-\theta)/2}$ and define, following Koul and Mukherjee (1993),

$$M(\Delta) := \tau_n^{-2} \int \left\| (X'X)^{-1/2} \Big\{ \sum_i x_i \big[I(Y_i - x_i'\Delta \le y) - I(-Y_i + x_i'\Delta < y) \big] \Big\} \right\|^2 dH(y),$$

$$Q(\Delta) := \tau_n^{-2} \int \Big\| (X'X)^{-1/2} \Big\{ \sum_i x_i \big[I(\epsilon_i \le y) - I(\epsilon_i > -y) \big] \Big\}$$
$$+ (X'X)^{1/2}(\Delta - \beta)\big[f(y) + f(-y)\big] \Big\|^2 dH(y), \quad \Delta \in \mathbb{R}^p,$$

where $I(A)$ is the indicator function of set A, $\|u\|$ denotes Euclidean norm and H is a nondecreasing right continuous function from $\mathbb{R}$ to $\mathbb{R}$. The m.d. estimator of the regression parameter β is defined by

$$\hat{\beta} := \text{argmin}\,\{M(\Delta),\ \Delta \in \mathbb{R}^p\}.$$

Note that $\hat{\beta}$ is the estimator β^+ defined in Koul (1985b) for the independent errors case and is the estimator β_K^+ defined in Koul and Mukherjee (1993) for the Gaussian subordinated process errors. The motivation for considering these m.d. estimators and its finite sample properties are discussed in Koul (1985b, 1992b). In particular, for $p = 1$, $x_i \equiv 1$, $H(x) = x\,[H(x) = I(x \geq 0)]$, $\hat{\beta}$ is the Hodge-Lehmann [Median estimator] estimator of the one sample location parameter.

Before we state the asymptotic normality of $\hat{\beta}$, we need the following additional assumptions on the model (3.2.1) and (3.2.2):

A.3 $\quad (X'X)^{-1}$ exists for all $n \geq p$.

A.4 $\quad n \max\limits_{1 \leq i \leq n} |x_i'(X'X)^{-1}x_i| = O(1)$.

A.5 $\quad \int (1 + y^2)^{-1}\, dH(y) < \infty$.

Conditions A.3 and A.4 are the same as those in Koul and Mukherjee (1993), while A.5 replaces the conditions $\int f^r dH < \infty$, $r = 1, 2$ and $\int_0^\infty (1 - F)dH < \infty$ of the above paper.

Let $A = \tau_n^{-1}(X'X)^{1/2}$, $B = \tau_n(X'X)^{1/2}$, $c_i = A^{-1}x_i$, $d_i = B^{-1}x_i$. We now state the main result:

Theorem 3.2.1 *In addition to (3.2.1) and (3.2.2), assume that A.1-A.5 hold.*

Then,

$$A(\hat{\beta}-\beta)$$
$$= \left(2\int f^2 dH\right)^{-1} \int \sum_i d_i \left[I(\varepsilon_i \le y) - I(\varepsilon_i > -y)\right] f(y)\, dH(y) + o_p(1). \quad (3.2.3)$$

The next result gives the asymptotic equivalence of the m.d. estimator in the first order to the least square estimator and its asymptotic normality.

Corollary 3.2.1 *Under the assumptions of Theorem 3.2.1,*

$$\tau_n^{-1}(X'X)^{1/2}(\hat{\beta}-\beta) = -\tau_n^{-1}(X'X)^{-1/2}\sum_{i=1}^{n} x_i\,\varepsilon_i + o_p(1). \quad (3.2.4)$$

Moreover,

$$G_n^{-1/2}\tau_n^{-1}(X'X)^{1/2}(\hat{\beta}-\beta) \Longrightarrow N_p(0, I_{p\times p}), \quad (3.2.5)$$

where $I_{p\times p}$ is $p \times p$ identity matrix, and

$$G_n = \tau_n^{-2}(X'X)^{-1/2}X'R_nX(X'X)^{-1/2}, \quad R_n = \left(\rho(i-j)\right)_{n\times n}, \quad i,j = 1,2,\cdots,n.$$

3.2.2 The case of random designs

In this subsection, we consider the following multiple linear regression model

$$Y_i = X_i'\beta + \varepsilon_i, \quad 1 \le i \le n, \quad \beta \in \mathbb{R}^p, \quad (3.2.6)$$

under the same assumptions as those in the previous section, except that here $\{X_i, i \ge 1\}$ are i.i.d. random variables, independent of the errors and $EX_1 \ne 0$.

Similarly, define

$$M_1(\Delta) := \tau_n^{-2} \int \Big\| n^{-1/2} \Big\{ \sum_i X_i \big[I(Y_i - X_i'\Delta \leq y) - I(-Y_i + X_i'\Delta < y) \big] \Big\} \Big\|^2 dH(y),$$

$$Q_1(\Delta) := \tau_n^{-2} \int \Big\| n^{-1/2} \Big\{ \sum_i X_i \big[I(\varepsilon_i \leq y) - I(\varepsilon_i > -y) \big] \Big\}$$

$$+ n^{-1/2} (X'X)(\Delta - \beta) \big[f(y) + f(-y) \big] \Big\|^2 dH(y).$$

The m.d. estimator of the parameter β in (3.2.6) is defined by

$$\hat{\beta}_1 := \text{argmin} \, \{ M_1(\Delta), \Delta \in \mathbb{R}^p \}.$$

Before we present the asymptotic normality of m.d. estimators, we need the following assumptions on the model (3.2.6).

A.6 $\quad E\|X_1\|^5 < \infty$.

Let $a_n = \tau_n^{-1} n^{1/2}$, $b_n = \tau_n n^{1/2}$, $C_i = a_n^{-1} X_i$, $D_i = b_n^{-1} X_i$, then we have the following analogous result of Theorem 3.2.1 under the i.i.d. random design case.

Theorem 3.2.2 *In addition to (3.2.6) and (3.2.2), assume that A.1, A.2, A.5 and A.6 hold, then*

$$a_n(\hat{\beta}_1 - \beta)$$
$$= \Big(2 \int f^2 dH \Big)^{-1} \int \sum_i D_i \big[I(\varepsilon_i \leq y) - I(\varepsilon_i > -y) \big] f(y) \, dH(y) + o_p(1). \tag{3.2.7}$$

Corollary 3.2.2 *Under the assumptions of Theorem 3.2.2,*

$$\tau_n^{-1} n^{1/2} (\hat{\beta}_1 - \beta) = -\tau_n^{-1} n^{-1/2} \sum_{i=1}^{n} X_i \varepsilon_i + o_p(1). \tag{3.2.8}$$

Moreover, let $EX_1 = \mu \neq 0$, *then*

$$\tau_n^{-1} n^{1/2}(\hat{\beta}_1 - \beta) = -\mu\, \tau_n^{-1} n^{-1/2} \sum_{i=1}^{n} \varepsilon_i + o_p(1), \tag{3.2.9}$$

and

$$\tau_n^{-1} n^{-1/2} \sum_{i=1}^{n} \varepsilon_i \Longrightarrow N(0, 1). \tag{3.2.10}$$

3.3 Proofs of the theorems

The method of proof is similar to that of Koul (1985a or 1992b; Ch5) which requires that $M(\Delta)$ is uniformly locally asymptotically approximated by quadratic form $Q(\Delta)$ and shows $\|A(\hat{\beta} - \beta)\| = O(1)$. This approximation in turn is used to obtain the asymptotic normality of m.d. estimators $\hat{\beta}$. For more details, see Koul (1992b; Ch5) and Koul (1985a).

In order to provide the details, we need some notations and several lemmas. Let C stand for a generic constant which may change from line to line. As in Ho and Hsing (1996, 1997) and Koul and Surgailis (1997, 2001a, b), put

$$\begin{aligned} \varepsilon_{il} &:= \sum_{k=1}^{l} b_k \zeta_{i-k}, & \tilde{\varepsilon}_{il} &:= \sum_{k=l+1}^{\infty} b_k \zeta_{i-k}, \\ F_l(x) &:= P(\varepsilon_{il} \le x), & f_l(x) &:= F_l'(x). \end{aligned} \tag{3.3.1}$$

The following two lemmas are analogous to Lemmas 5.1 and 5.2 of Koul and Surgailis (2001b), thus their proofs can be deduced from there.

Lemma 3.3.1 *Under the assumptions A.1 and A.2, there exist* $l_0 \geq 1$ *and a con-*

stant C such that for any $l \geq l_0$, $x \in \mathbb{R}$,

$$\left|f^{(p)}(x)\right| + \left|f_l^{(p)}(x)\right| \leq C\left(1 + |x|^3\right)^{-1}, \quad p = 0, 1, 2, \tag{3.3.2}$$

$$\left|f_l(x) - f_{l-1}(x)\right| \leq C b_l^2 \left(1 + |x|^3\right)^{-1}. \tag{3.3.3}$$

Lemma 3.3.2 *Let $g_\gamma(x) := (1 + |x|^\gamma)^{-1}$ and $h(x)$, $x \in \mathbb{R}$ be a real valued function such that, for some $C < \infty$,*

$$|h(x)| \leq C g_\gamma(x), \quad \gamma = 2, 3. \tag{3.3.4}$$

Then there exists a constant C_γ, depending only on C in (3.3.4), such that for any x, $y \in \mathbb{R}$,

$$|h(x + y)| \leq C_\gamma g_\gamma(x)\left(1 \vee |y|^\gamma\right), \tag{3.3.5}$$

where $a \vee b = max\{a,\ b\}$.

Remark 3.3.1 From (3.3.2) in the Lemma 3.3.1, $f(x)$ and $f'(x)$ satisfy conditions of $h(x)$ in Lemma 3.3.2, thus, $|f(x + y)| \leq C(1 + x^2)^{-1}(1 + y^2)$, $\quad |f'(x + y)| \leq C(1 + x^2)^{-1}(1 + y^2)$.

Lemma 3.3.3 (Surgailis). *Under the assumptions A.1 and A.2, there exists a constant $C < \infty$ such that*

$$\left|Cov\left(I(\varepsilon_0 \leq x),\ I(\varepsilon_i \leq x)\right)\right| \leq C\left(1 + x^2\right)^{-1} i^{-\theta},$$

for all $i \in \mathbb{Z}$, $x \in \mathbb{R}$.

Proof. The proof is in Appendix.

Lemma 3.3.4 *Under assumptions of A.1 and A.2, there exists a constant C such that*

$$\begin{aligned}&\left|Cov\big(I(x < \varepsilon_0 \le x + \alpha_0),\ I(x < \varepsilon_i \le x + \alpha_i)\big)\right| \\ &\qquad\qquad \le C\big(1 + x^2\big)^{-1} i^{-\theta} \big[|\alpha_0| \vee |\alpha_0|^3\big]^{1/2} \big[|\alpha_i| \vee |\alpha_i|^3\big]^{1/2},\end{aligned}$$

for all $i \in \mathbb{Z},\ x \in \mathbb{R},\ \alpha_0 \in \mathbb{R},\ \alpha_i \in \mathbb{R}$.

Proof. The proof is very similar to that of Lemma 3.3.3, so here we only give the outline. Let $\mathcal{F}_i$ be the σ-field generated by $\zeta_k,\ k \le i$ and $F(x, y) := F(y) - F(x)$. Write the telescoping identity:

$$I(x < \varepsilon_i \le x + \alpha_i) - F(x, x + \alpha_i) = \sum_{l=1}^{\infty} U_{i,l}(x, x + \alpha_i), \tag{3.3.6}$$

where

$$\begin{aligned}U_{i,l}(x, x + \alpha_i) &= F_{l-1}(x - \tilde{\varepsilon}_{i,l-1}, x + \alpha_i - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l}, x + \alpha_i - \tilde{\varepsilon}_{i,l}) \\ &= U_{i,l}^{(1)}(x, x + \alpha_i) + U_{i,l}^{(2)}(x, x + \alpha_i),\end{aligned}$$

where

$$\begin{aligned}U_{i,l}^{(1)}(x, x + \alpha_i) &= F_l(x - \tilde{\varepsilon}_{i,l-1}, x + \alpha_i - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l}, x + \alpha_i - \tilde{\varepsilon}_{i,l}), \\ U_{i,l}^{(2)}(x, x + \alpha_i) &= F_{l-1}(x - \tilde{\varepsilon}_{i,l-1}, x + \alpha_i - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l-1}, x + \alpha_i - \tilde{\varepsilon}_{i,l-1}).\end{aligned}$$

In order to prove the lemma, as that of Lemma 3.3.3, we only need to show the following (3.3.7) and (3.3.8).

$$E\Big[U_{i,l}(x, x + \alpha_i)\Big]^2 \le C(1 + x^2)^{-1}\big(|\alpha_i| \vee |\alpha_i|^3\big), \quad l = 1, 2, \cdots, l_0, \tag{3.3.7}$$

$$E\Big[U_{i,l}^{(q)}(x, x + \alpha_i)\Big]^2 \le C(1 + x^2)^{-1} l^{-1-\theta}\big(|\alpha_i| \vee |\alpha_i|^3\big), \quad l > l_0, \quad q = 1, 2. \tag{3.3.8}$$

Proof of (3.3.7). According to the definition, we have

$$E\Big[U_{i,l}(x, x+\alpha_i)\Big]^2 \le 2\big[EF_{l-1}^2(x - |\alpha_i| - \tilde{\varepsilon}_{i,l-1}, x + |\alpha_i| - \tilde{\varepsilon}_{i,l-1})$$

$$+ EF_l^2(x - |\alpha_i| - \tilde{\varepsilon}_{i,l}, x + |\alpha_i| - \tilde{\varepsilon}_{i,l})\big]$$

$$\le 4F(x - |\alpha_i|, x + |\alpha_i|)$$

$$= 4\int_{-|\alpha_i|}^{|\alpha_i|} f(x+v)\, dv$$

$$\le C(1+x^2)^{-1}\big(|\alpha_i| \vee |\alpha_i|^3\big),$$

the last inequality follows from lemma 3.3.2 with $h(x)$ replaced by $f(x)$ and $\gamma = 2$.

Proof of (3.3.8). For $q = 1$,

$$U_{i,l}^{(1)}(x, x+\alpha_i) = \int_x^{x+\alpha_i} \big[f_l(u - b_l\zeta_{i-l} - \tilde{\varepsilon}_{i,l}) - f_l(u - \tilde{\varepsilon}_{i,l})\big]\, du. \tag{3.3.9}$$

Follows the argument of Lemma 3.3.3, apply Lemma 3.3.1 and 3.3.2 with $\gamma = 2$, we can obtain the following analogous inequality

$$\big|U_{i,l}^{(1)}(x, x+\alpha_i)\big| \le C\big(|b_l\zeta_{i-l}| \vee |b_l\zeta_{i-l}|^2\big)(1+\tilde{\varepsilon}_{i,l}^2)(1+x^2)^{-1}\big(|\alpha_i| \vee |\alpha_i|^3\big).$$

From (3.3.9) and (3.3.2), we have $\big|U_{i,l}^{(1)}(x, x+\alpha_i)\big| \le C(|b_l\zeta_{i-l}| \wedge 1)$, thus we obtain

$$E\Big[U_{i,l}^{(1)}(x, x+\alpha_i)\Big]^2 \le CE|b_l\zeta_{i-l}|^2(1 + E\tilde{\varepsilon}_{i,l}^2)(1+x^2)^{-1}\big(|\alpha_i| \vee |\alpha_i|^3\big),$$

which is (3.3.8) for $q = 1$.

For $q = 2$, apply Lemma 3.3.1 and 3.3.2, we have

$$\begin{aligned}
\left|U_{i,l}^{(2)}(x, x+\alpha_i)\right| &\le \left|\int_x^{x+\alpha_i} \left[f_l(u - \tilde{\varepsilon}_{i,l-1}) - f_{l-1}(u - \tilde{\varepsilon}_{i,l-1})\right] du\right| \\
&\le C \int_{x-|\alpha_i|}^{x+|\alpha_i|} b_l^2 \left(1 + |u - \tilde{\varepsilon}_{i,l-1}|^2\right)^{-1} du \\
&\le C b_l^2 \int_{x-|\alpha_i|}^{x+|\alpha_i|} (1+u^2)^{-1}(1+\tilde{\varepsilon}_{i,l-1}^2)\, du \\
&\le C b_l^2 (1+\tilde{\varepsilon}_{i,l-1}^2)(1+x^2)^{-1}\left(|\alpha_i| \vee |\alpha_i|^3\right).
\end{aligned}$$

Again, as $\left|U_{i,l}^{(2)}(x, x+\alpha_i)\right| \le 2$, we obtain (3.3.8) for $q = 2$. Hence, we proved the lemma.

We are now ready to state and prove the asymptotic uniform quadraticity of $M(\Delta)$.

Lemma 3.3.5 *Under the assumptions of Theorem 3.2.1, for all $b \in (0, \infty)$,*

$$E \sup_{s \in N(b)} \left|M(\beta + A^{-1}s) - Q(\beta + A^{-1}s)\right| = o(1), \tag{3.3.10}$$

where $N(b) = \{s \in \mathbb{R}^p : \|s\| \le b\}$.

Proof. The proof basically is similar to that of Theorem 2.1 of Koul (1985a). As there, use the symmetry assumption of $f(y)$, it is enough to show that $\forall\, b \in (0, \infty)$,

$$\sup_{s \in N(b)} \int \left\|\sum_i d_i \left[F(y, y + c_i's) - c_i's f(y)\right]\right\|^2 dH(y) = o(1), \tag{3.3.11}$$

$$E \sup_{s \in N(b)} \int \left\|\sum_i d_i \left[I(y < \varepsilon_i \le y + c_i's) - F(y, y + c_i's)\right]\right\|^2 dH(y) = o(1), \tag{3.3.12}$$

$$E \sup_{s \in N(b)} \int \left\|\sum_i d_i \left[I(\varepsilon_i \le y) - F(y) + c_i's f(y)\right]\right\|^2 dH(y) = O(1), \tag{3.3.13}$$

The first equality (3.3.11) follows from the infinite differentiablity of F, $\sum_i d_i c_i' = I_{p \times p}$, $\max_i |c_i's| \to 0$, (3.3.2) and assumption A.5.

As to the (3.3.12), here we only give the proof for fixed $s \in N(b)$. The uniform convergence can be obtained by the compactness of $N(b)$, similar to that of Theorem 2.1 (Koul,1985a). Let d_{ij}:=the j-th entry of the vector d_i. Thus the integrand of the j-th summand of the left hand side (LHS) of (3.3.12) does not exceed

$$\sum_i \sum_r |d_{ij} d_{rj}| \left|\mathrm{Cov}\big(I(y < \varepsilon_i \leq y + c_i' s),\ I(y < \varepsilon_r \leq y + c_r' s)\big)\right|.$$

Apply Lemma 3.3.4, notice $\|c_i' s\| \leq \|c_i\| \|s\| = O(n^{-\theta/2})\|s\| \to 0$, so, for any $0 < h < \theta$, the above bound does not exceed

$$\begin{aligned} & C \sum_i \sum_r |d_{ij} d_{rj}| \, (1+y^2)^{-1} \big(1 + |i-r|\big)^{-\theta} n^{-h/2} \|s\| \\ & \leq C \|s\| n^{-(2-\theta)} n^{2-\theta} n^{-h/2} (1+y^2)^{-1}, \end{aligned}$$

where the last inequality follows from $\max_i \|d_i\| = O(n^{-(2-\theta)/2})$. Thus the j-th entry of the LHS of (3.3.12) does not exceed

$$C\|s\| n^{-h/2} \int (1+y^2)^{-1} \, dH(y) \to 0, \quad n \to \infty,$$

which proves the (3.3.12).

As to (3.3.13), we need to prove

$$\int \Big\| \sum_i d_i c_i' s f(y) \Big\|^2 dH(y) = O(1), \tag{3.3.14}$$

$$\int E \Big\| \sum_i d_i \big[I(\varepsilon_i \leq y) - F(y)\big] \Big\|^2 dH(y) = O(1). \tag{3.3.15}$$

But $\sum_i d_i c_i' = I_{p \times p}$, $\|s f(y)\|^2 \leq \|s\|^2 f^2(y) \leq C b^2 (1+y^2)^{-1}$. Thus (3.3.14) follows from (3.3.2) and assumption $A.5$. As to (3.3.15), like (3.3.12), apply Lemma 3.2.1, the j-th entry of the LHS of (3.3.15) does not exceed $c n^{-2+\theta} n^{2-\theta} \int (1+y^2)^{-1} \, dH(y) < \infty$, hence the lemma is proved.

Lemma 3.3.6 *Let* $\hat{\Delta} = argmin\{Q(\Delta), \Delta \in \mathbb{R}^p\}$, *assume the conditions of Theorem 3.2.1 hold, then*

$$||A(\hat{\beta} - \beta)|| = O_p(1). \tag{3.3.16}$$

$$||A(\hat{\Delta} - \beta)|| = O_p(1). \tag{3.3.17}$$

Proof. The proof of (3.3.16) follows from the following Lemma, while proof of (3.3.17) basically is the same as that of (3.3.16).

Lemma 3.3.7 *Assume the conditions of Theorem 3.2.1 hold, then*

(a). for any $\varepsilon > 0$, *there exists a* $0 < z_\varepsilon < \infty$ *and* $N_{1\varepsilon}$ *such that*

$$P\Big(|M(\beta)| \le z_\varepsilon\Big) \ge 1 - \varepsilon, \quad \textit{for all} \quad n \ge N_{1\varepsilon}.$$

(b). for any $\varepsilon > 0,\ 0 < z < \infty$, *there exists* $N_{2\varepsilon}$ *and a positive* $b > 0$ *such that*

$$P\Big(\inf_{||s|| \ge b} M(\beta + A^{-1}s) \ge z\Big) \ge 1 - \varepsilon \quad \textit{for all} \quad n \ge N_{2\varepsilon}.$$

Proof. The proof of part (a) is from finite moment $EM(\beta) < \infty$, which is from (3.3.15). The part (b) is very similar to that of Lemma 3.1 of Koul (1985a) which we omit here.

Finally, we are in the position to provide the proof of main theorem.

Proof of Theorem 3.2.1. The proof follows that of Theorem 5.41 of Koul (1992b) and Theorem 3.1 of Koul (1985a). We only give the sketch here. From Lemma 3.3.6, we have

$$\begin{aligned}\Big|M(\hat{\beta}) - Q(\hat{\Delta})\Big| &= \Big|\inf_{||s|| \le b} M(\beta + A^{-1}s) - \inf_{||s|| \le b} Q(\beta + A^{-1}s)\Big| \\ &\le \sup_{||s|| \le b} \Big|M(\beta + A^{-1}s) - Q(\beta + A^{-1}s)\Big|.\end{aligned}$$

From above inequality and Lemma 3.3.5, we get

$$M(\hat{\beta}) = Q(\hat{\Delta}) + o_p(1). \tag{3.3.18}$$

The last equality (3.3.18) together with $M(\hat{\beta}) = Q(\hat{\beta}) + o_p(1)$, yield $Q(\hat{\Delta}) = Q(\hat{\beta}) + o_p(1)$, which is precisely $\|A(\hat{\beta} - \hat{\Delta})\| = o_p(1)$. Thus

$$A(\hat{\beta} - \beta) = A(\hat{\Delta} - \beta) + o_p(1). \tag{3.3.19}$$

Now, from the defintion of $Q(\Delta)$ and $\hat{\Delta}$, we readily get the (3.2.3) of Theorem 3.2.1 from (3.3.19).

In order to prove the Corollary 3.2.1, we need the following lemma.

Lemma 3.3.8 *Let* $S_n(x) = \sum_{i=1}^{n} d_i\big[I(\varepsilon_i \le x) - F(x) + f(x)\varepsilon_i\big]$, *under the assumptions of Theorem 3.2.1, then*

$$\sup_x \big|S_n(x)\big| = o_p(1).$$

Proof. The proof of the lemma can be deduced from Theorem 3.1 of Koul and Surgailis (2001c), where they proved more general case, i.e. the uniform reduction priciple for weighted residuals empirical processes.

Proof of Corollary 3.2.1. From the Theorem 3.2.1 and notation of $S_n(x)$, we obtain

$$\begin{aligned} A(\hat{\beta} - \beta) &= \Big(2\int f^2 dH\Big)^{-1} \int [S_n(y) + S_n(-y) - 2\sum_{i=1}^{n} d_i\varepsilon_i f(y)] f(y) dH(y) + o_p(1) \\ &= -\sum_{i=1}^{n} d_i\varepsilon_i + o_p(1), \end{aligned}$$

the last equality , which is (3.2.4), follows from lemma 3.3.8, while (3.2.5) follows from Theorem 2 of Giraitis et al. (1996).

The following lemma is the asymptotic uniform quadraticity of $M_1(\Delta)$ under i.i.d. random case.

Lemma 3.3.9 *Assume the conditions of Theorem 3.2.2 hold. Then, for all* $b \in (0, \infty)$,

$$\sup_{s \in N(b)} \left| M_1(\beta + a_n^{-1} s) - Q_1(\beta + a_n^{-1} s) \right| = o_p(1).$$

Proof. The proof is similar to that of Lemma 3.3.5 except here $\{X_i, i \geq 1\}$ are i.i.d. r.v's, instead of fixed known constants. To prove the theorem, it is enough to show that $\forall\, b \in (0, \infty)$,

$$\sup_{s \in N(b)} \int \Big\| \sum_i D_i \big[F(y, y + C_i' s) - C_i' s f(y) \big] \Big\|^2 dH(y) = o_p(1), \tag{3.3.20}$$

$$E \sup_{s \in N(b)} \int \Big\| \sum_i D_i \big[I(y < \varepsilon_i \leq y + C_i' s) - F(y, y + C_i' s) \big] \Big\|^2 dH(y) = o(1), \tag{3.3.21}$$

$$E \sup_{s \in N(b)} \int \Big\| \sum_i D_i \big[I(\varepsilon_i \leq y) - F(y) + C_i' s f(y) \big] \Big\|^2 dH(y) = O(1). \tag{3.3.22}$$

Proof of (3.3.20). Write $F(x, y) = \int_x^y f(u) du$, use the differentiability of f, we can obtain

$$\text{LHS of (3.3.20)} \leq b^4 \int \Big\| n^{-1} \sum_i X_i X_i' \|X_i\| \int_{-a_n^{-1}}^{a_n^{-1}} \left| f'\left(y + b\|X_i\| z\right) \right| dz \Big\|^2 dH(y). \tag{3.3.23}$$

Now, from Remark 3.3.1, we have

$$\int_{-a_n^{-1}}^{a_n^{-1}} \left| f'\left(y + b\|X_i\| z\right) \right| dz \leq C a_n^{-1} \left(1 + b^2 \|X_i\|^2\right) \left(1 + y^2\right)^{-1}.$$

Thus, from (3.3.23), we have

$$\text{LHS of (3.3.20)} \leq Ca_n^{-2}b^4 \left\| n^{-1} \sum_i X_i X_i' \|X_i\| (1 + b^2 \|X_i\|^2) \right\|^2 \int (1+y^2)^{-1} dH(y),$$

which is $o_p(1)$ from A.6, A.5 and $a_n^{-2} \to 0$, hence (3.3.20) is proved.

Proof of (3.3.21). Similar to the proof of (3.3.12), we here only give the proof for the fixed $s \in N(b)$. Let D_{ij}:=the j-th entry of the vector D_i, which are analogous to d_{ij} in the fixed design case and

$$K_i(z) := I(y < z \leq y + C_i's) - F(y, y + C_i's).$$

Thus the integrand of the j-th summand of (3.3.21) does not exceed

$$\sum_i \sum_r E\left[|D_{ij} D_{rj}| \left| E\left[K_i(\varepsilon_i) K_r(\varepsilon_r) \middle| X_i, X_r \right] \right| \right]. \tag{3.3.24}$$

Apply Lemma 3.3.4, as $|C_i's| \leq a_n^{-1} \|X_i\| \|s\|$, we have

$$\begin{aligned} &\left| E\left[K_i(\varepsilon_i) K_r(\varepsilon_r) \middle| X_i, X_r \right] \right| \\ &\qquad \leq C(1+y^2)^{-1} (1 + |i-r|)^{-\theta} \left[|C_i's| \vee |C_i's|^3 \right]^{1/2} \left[|C_r's| \vee |C_r's|^3 \right]^{1/2}. \end{aligned} \tag{3.3.25}$$

Combining (3.3.24) and (3.3.25), use A.6, we obtain, for any $0 < h < \theta$, the j-th entry of

$$\begin{aligned} \text{LHS of (3.21)} &\leq C(\|s\| \vee \|s\|^3) n^{-h/2} n^{-2+\theta} \sum_i \sum_r (1 + |i-r|)^{-\theta} \int (1+y^2)^{-1}\, dH(y) \\ &\leq C(\|s\| \vee \|s\|^3) n^{-h/2} \int (1+y^2)^{-1}\, dH(y) \to 0, \quad n \to \infty, \end{aligned}$$

which proves the (3.3.21).

Proof of (3.3.22). It suffices to prove

$$\int E\|\sum_i D_i C_i' s f(y)\|^2 \, dH(y) = O(1), \tag{3.3.26}$$

$$\int E\left\|\sum_i D_i [I(\varepsilon_i \le y) - F(y)]\right\|^2 dH(y) = O(1). \tag{3.3.27}$$

The first equality (3.3.26) follows from the following inequality and assumptions A.6, (3.3.2) and A.5.

$$\text{LHS of (3.3.26)} \le \|s\|^2 E\left\|n^{-1}\sum_i X_i X_i'\right\|^2 \int f^2(y)\, dH(y) < \infty.$$

As to the (3.3.27), like that of (3.3.15), we have the integrand of the j-th summand of the LHS of (3.3.27) does not exceed

$$\sum_i \sum_r E\left[|D_{ij} D_{rj}| \left| E\left\{[I(\varepsilon_i \le y) - F(y)][I(\varepsilon_r \le y) - F(y)] \middle| X_i, X_r\right\}\right|\right].$$

Thus, from Lemma 3.3.3 and similar argument as (3.3.24), we obtain the j-th entry of

$$\text{LHS of (3.3.27)} \le C b_n^{-2} \sum_i \sum_r (1 + |i - r|)^{-\theta} \int (1 + y^2)^{-1} \, dH(y).$$

Thus, LHS of (3.3.27) $\le C n^{-(2-\theta)} n^{2-\theta} \int (1 + y^2)^{-1} dH(y) < \infty$. Hence, lemma is proved.

Proof of Theorem 3.2.2. The proof is completely analogous to that of Theorem 3.2.1.

Proof of Corollary 3.2.2. Proof of the (3.2.8) is completely analogous to (3.2.4) of Corollary 3.2.1. From (3.2.8), we obtain

$$\tau_n^{-1} n^{1/2} (\hat{\beta}_1 - \beta) = -\tau_n^{-1} n^{-1/2} \sum_{i=1}^{n} (X_i - \mu)\varepsilon_i - \mu \tau_n^{-1} n^{-1/2} \sum_{i=1}^{n} \varepsilon_i + o_p(1). \tag{3.3.28}$$

But, the variance of the first term of the RHS of (3.3.28) goes to zero, thus the first term is $o_p(1)$, which proves (3.2.9). The last equality (3.2.10) follows from the Lemma 5.1 of Surgailis (1982).

3.4 Appendix

Before we give the proof of Lemma 3.3.3, we need a following lemma.

Lemma 3.4.1 *Let $g(x) = (1+|x|^3)^{-1}$ and $h(x)$, $x \in \mathbb{R}$ be a real valued function such that*

$$|h(x)| \le Cg(x), \tag{3.4.1}$$

hold for any $x \in \mathbb{R}$. Then, for any $x \le 0$ and any $v, w \in \mathbb{R}$

$$\left|\int_{-\infty}^{x} \big[h(u+v+w) - h(u+w)\big]du\right| \le C\big(|v| \vee |v|^3\big)\big(1 \vee |w|^3\big)(1+x^2)^{-1}. \tag{3.4.2}$$

Proof. First consider $|v| \le 1$, then by (3.4.1) and (3.3.5) with $\gamma = 3$, the LHS of (3.4.2) does not exceed

$$\begin{aligned} C|v|\int_{-\infty}^{x} \big(1+|u+w|^3\big)^{-1}du &\le C|v|(1 \vee |w|^3)\int_{-\infty}^{x}(1+|u|^3)^{-1}du \\ &\le C|v|(1 \vee |w|^3)(1+x^2)^{-1}. \end{aligned}$$

Next, consider $|v| > 1$. Then the LHS of (3.4.2) does not exceed

$$C\int_{\infty}^{x}\big(1+|u+v+w|^3\big)^{-1}du + C\int_{-\infty}^{x}\big(1+|u+w|^3\big)^{-1}du. \tag{3.4.3}$$

By (3.3.5), the first term of (3.4.3) does not exceed

$$C(1 \vee |v+w|^3)\int_{-\infty}^{x}(1+|u|^3)^{-1}du \le C|v|^3(1 \vee |w|^3)(1+x^2)^{-1}.$$

The second term of (3.4.3) follows similarly. This proves the lemma.

Proof of Lemma 3.3.3. Let $\mathcal{F}_i$ be the σ-field generated by $\zeta_k,\ k \le i$. Write the telescoping identity:

$$I(\varepsilon_i \le x) - F(x) = \sum_{l=1}^{\infty} U_{i,l}(x), \tag{3.4.4}$$

where

$$\begin{aligned} U_{i,l}(x) &= F_{l-1}(x - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l}) \\ &= U_{i,l}^{(1)}(x) + U_{i,l}^{(2)}(x), \end{aligned} \tag{3.4.5}$$

where

$$U_{i,l}^{(1)}(x) = F_l(x - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l}),$$

$$U_{i,l}^{(2)}(x) = F_{l-1}(x - \tilde{\varepsilon}_{i,l-1}) - F_l(x - \tilde{\varepsilon}_{i,l-1}).$$

The lemma 3.3.3 follows from the following (3.4.6) and (3.4.7).

$$E\Big[U_{i,l}(x)\Big]^2 \le C(1+x^2)^{-1}, \quad l = 1, 2, \cdots, l_0, \tag{3.4.6}$$

$$E\Big[U_{i,l}^{(q)}(x)\Big]^2 \le C(1+x^2)^{-1}\, l^{-1-\theta}, \quad l > l_0, \quad q = 1, 2, \tag{3.4.7}$$

where l_0 will be chosen sufficiently large in order that the bounds of Lemma 3.3.1 hold. Indeed, by orthogonality of (3.4.4) and (3.4.6), (3.4.7).

$$\begin{aligned} \Big|\mathrm{Cov}\big(I(\varepsilon_0 \le x), I(\varepsilon_i \le x)\big)\Big| &= \Big|\sum_{l=1}^{\infty} EU_{i,i+l}(x) U_{0,l}(x)\Big| \\ &\le \sum_{l=1}^{\infty} E^{1/2}\big[U_{i,i+l}(x)\big]^2 E^{1/2}\big[U_{0,l}(x)\big]^2 \\ &\le C(1+x^2)^{-1} \sum_{l=1}^{\infty} (i+l)^{-(1+\theta)/2}\, l^{-(1+\theta)/2} \\ &\le C(1+x^2)^{-1} i^{-\theta}. \end{aligned}$$

Now, it suffices to show (3.4.6) and (3.4.7) for $x \le 0$ only. As to (3.4.6),

$$\begin{aligned} E\big[U_{i,l}(x)\big]^2 &\le 2\big[EF_{l-1}^2(x-\tilde{\varepsilon}_{i,l-1}) + EF_l^2(x-\tilde{\varepsilon}_{i,l})\big] \\ &\le 2\big[EF_{l-1}(x-\tilde{\varepsilon}_{i,l-1}) + EF_l(x-\tilde{\varepsilon}_{i,l})\big] \\ &= 4F(x). \end{aligned}$$

Notice $F(x) = \int_{-\infty}^{x} f(u)du$ and by Lemma 3.3.1 (3.3.2), we have $F(x) \le C(1+x^2)^{-1}$, this proves (3.4.6).

Consider (3.4.7) for $q=1$. In view of (3.4.5), as $\tilde{\varepsilon}_{i,l-1} = b_l\zeta_{i-l} + \tilde{\varepsilon}_{i,l}$, we have

$$U_{i,l}^{(1)}(x) = \int_{-\infty}^{x} \big[f_l(u - b_l\zeta_{i-l} - \tilde{\varepsilon}_{i,l}) - f_l(u - \tilde{\varepsilon}_{i,l})\big]du.$$

Here, f_l satisfies Lemma 3.4.1's $h(x)$ by Lemma 3.3.1. Thus from (3.4.2), we obtain

$$\begin{aligned} \big|U_{i,l}^{(1)}(x)\big| &\le C\big(|b_l\zeta_{i-l}| \vee |b_l\zeta_{i-l}|^3\big)\big(1 \vee |\tilde{\varepsilon}_{i,l}|^3\big)(1+x^2)^{-1} \\ &\le C\big(|b_l\zeta_{i-l}| \vee |b_l\zeta_{i-l}|^3\big)\big(1 + |\tilde{\varepsilon}_{i,l}|^3\big)(1+x^2)^{-1}. \end{aligned} \tag{3.4.8}$$

Combining (3.4.8) with the estimate $|U_{i,l}^{(1)}(x)| \le C(|b_l\zeta_{i-l}| \wedge 1)$, which is an easy consequence of (3.3.2), we obtain

$$\begin{aligned} E\big[U_{i,l}^{(1)}(x)\big]^2 &\le C\big(E|b_l\zeta_{i-l}|^2 + E|b_l\zeta_{i-l}|^3\big)\big(1 + E|\tilde{\varepsilon}_{i,l}|^3\big)(1+x^2)^{-1} \\ &\le Cb_l^2(1+x^2)^{-1} \\ &\le C(1+x^2)^{-1}l^{-1-\theta}, \end{aligned} \tag{3.4.9}$$

the second inequality follows from $E|\tilde{\varepsilon}_{i,l}|^3 < \infty$, which follows from the Rosenthal inequality.

$$E\Big|\sum_{l=1}^{\infty} b_l\zeta_l\Big|^3 \le C\sum_{l=1}^{\infty} E|b_l\zeta_l|^3 + C\Big(\sum_{l=1}^{\infty} E|b_l\zeta_l|^2\Big)^{3/2},$$

this proves (3.4.7) for $q = 1$.

As to (3.4.7) for $q = 2$. From Lemma 3.3.1 (3.3.3) and Lemma 3.3.2 (3.3.5) with $\gamma = 3$, we obtain

$$\begin{aligned}
|U_{i,l}^{(2)}(x)| &= \left| \int_{-\infty}^{x} [f_l(u - \tilde{\varepsilon}_{i,l-1}) - f_{l-1}(u - \tilde{\varepsilon}_{i,l-1})] du \right| \\
&\leq C b_l^2 \int_{-\infty}^{x} \left(1 + |u - \tilde{\varepsilon}_{i,l-1}|^3\right)^{-1} du \\
&\leq C b_l^2 \left(1 \vee |\tilde{\varepsilon}_{i,l-1}|^3\right) \int_{-\infty}^{x} (1 + |u|^3)^{-1} du \\
&\leq C b_l^2 \left(1 \vee |\tilde{\varepsilon}_{i,l-1}|^3\right)(1 + x^2)^{-1}.
\end{aligned}$$

Hence, as $|U_{i,l}^{(2)}(x)| \leq 2$, similarly as $q = 1$, we obtain

$$E\left[U_{i,l}^{(2)}(x)\right]^2 \leq C b_l^2 (1 + x^2)^{-1} \leq C(1 + x^2)^{-1} l^{-1-\theta},$$

this, together with (3.4.9), proves (3.4.7). Hence the lemma is proved.

Bibliography

[1] Antoniadis, A., Grĕgoire, G. and McKeague, I. W. (1994) Wavelet methods for curve estimation. *J. Amer. Statist. Assoc.* 89, 1340-1353.

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

[3] Baillie, R.T. (1996) Long memory processes and fractional integration in econometrics. *J. Econometrics*, 73, 5-59.

[4] Beran, J. (1992) Statistical methods for data with long range dependence. *Statist. Science* 7, 404-427.

[5] Beran, J. (1994) Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probab., 61. Chapman and Hall. N.Y.

[6] Chui, K. (1992) *Wavelets: A tutorial in theory and applications* , Boston: Academic Press.

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

[8] Dehling, H. and M.S. Taqqu. (1989) The empirical processes of some long rang dependent sequences with application to U-statistics. *Ann. Statist.*, 17, 1767-1783.

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

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

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

[12] Feller, W. (1957) An Introduction to Probability Theory and its Application. Volume I, *Second edition*, John Wiley & Sons, Inc.

[13] Giraitis, L., Koul, H.L. and Surgailis, D. (1996) Asymptotic normality of regression estimators with long memory errors *Statist. Probab. Lett.* 29, 317-335.

[14] Giraitis, L. and Surgailis, D. (1999) Central limit theorem for the empirical process of a linear sequence with long memory. *J. Statist. Plann. and Inference*, 80, 81-93.

[15] Hall, P. and Patil, P. (1993) On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods. Re-

search Report SMS-72-93, Center for Mathematics and Statistics. Australian National University, Canberra.

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

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

[18] Ho, H.C. and Hsing, T. (1996) On the asymptotic expansion of the empirical process of long memory moving averages. *Ann. Statist.* 24, 992-1024.

[19] Ho, H.C. and Hsing, T. (1997) Limit theorems for functionals of moving averages, *Ann. Probab.* 25, 1636-1669.

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

[21] Kerkyacharian, G. and Picard, D. (1993) Density estimation by kernel and wavelet methods, optimality in Besov space. *Statist. Probab. Lett.* , 18, 327-336.

[22] Koul, H.L. (1985a) Minimum distance estimation in multiple linear regression, *Sankhya Ser. A* 47, 57-74.

[23] Koul, H.L. (1985b) Minimum distance estimation in multiple linear regression with unknown error distributions *Statist. Probab. Lett.* 3, 1-8.

[24] Koul, H.L. (1986) Minimum distance estimation and goodness-of-fit tests in first order autoregression. *Ann. Statist.* 14, 1194-1213.

[25] Koul, H.L. (1992a) M-estimators in linear models with long range dependent errors. *Statist. Probab. Lett.* 14, 153-164.

[26] Koul, H.L. (1992b) Weighted Empiricals and Linear Models. IMS Lecture Notes, 21.

[27] Koul, H.L. and DeWet, T. (1983) Minimum distance estimation in linear regression models. *Ann. Statist.* 11, 921-932.

[28] Koul, H.L. and K. Mukherjee. (1993) Asymptotics of R-, MD- and LAD-estimators in linear regression models with long range dependent errors. *Probab. Theory Related Fields*, 95, 535-553.

[29] Koul, H.L. and D. Surgailis. (1997) Asymptotic expansion of M-estimators with long memory errors. *Ann. Statist.* 25, 818-850.

[30] Koul, H.L. and D. Surgailis. (2001a) Asymptotics of the empirical process of long memory moving averages with infinite variance *J. Stochastic Processes & Appl.*, 91(2), 309-336.

[31] Koul, H.L. and D. Surgailis. (2001b) Asymptotic expansion of the empirical process of long memory moving averages. *preprint.*

[32] Koul, H.L. and D. Surgailis. (2001c) Robust estimators in regression models with long memory errors. *preprint.*

[33] Lo, S.H., Mack, Y.P. and Wang, J.L. (1989) Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator. *Probab. Theory and Related Fields*, 80, 461-473.

[34] Mallat, S. (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674-693.

[35] Marron, J.S. and Padgett, W.J. (1987) Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples. *Ann. Statist.*, 15, 1520-1535.

[36] Meyer, Y. (1990) *Ondelettes et operateurs*, Hermann, Paris.

[37] Padgett, W.J. and McNichols. D.T. (1984) Nonparametric density estimation from censored data. *Comm. Statist. Theory and Methods*, 13, 1581-1611.

[38] Patil, P. (1997) Nonparametric hazard rate estimation by orthogonal wavelet methods. *Journal of Statistical Planning and Inference*, 60, 153-168.

[39] Robinson, P.M. (1994) Semiparametric analysis of long-memory time series. *Ann. Statist.*, 22, 515-539.

[40] Singpurwalla, N.D. and Wong, M.Y. (1983) Estimation of the failure rate, a survey of the nonparametric methods. Part I: Non Bayesian Methods. *Comm. Statist. Theory and Methods*, 12, 559-588.

[41] Stute, W. (1995) The central limit theorem under random censorship. *Ann. Statist.*, 23, 422-439.

[42] Stute, W. and Wang, J. -L. (1993) The strong law under random censorship. *Ann. Statist.* 21, 1591-1607.

[43] Surgailis, D. (1982) Zones of attraction of self-similar multiple integrals. *Lithuanian Mathematical J.* 22, 327-340.

[44] Tanner, M. A. and Wong, W.H. (1983) The estimation of the hazard function from randomly censored data by the kernel method. *Ann. Statist.*, 11, 989-993.

[45] Taqqu, M.S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahhrsch.* 31, 287-302.

[46] Wu, S. and Wells, M. (1999) Estimating hazard rate with truncated and censored data by wavelet methods. *preprint*

[47] Zhang, B. (1996) Some asymptotic results for kernel density estimation under random censorship. *Bernoulli* 2, 183-198.