Journal Afrika Statistika

ISSN 0852-0305

Journal Afrika Statistika
Vol. 6, 2011, pages 391–411.
DOI: http://dx.doi.org/10.4314/afst.v7i1.2

Nonlinear wavelet regression function estimator for censored dependent data

Djabrane Yahia and Fateh Benatia

Laboratory of Applied Mathematics, Mohamed Khider University of Biskra, 07000, Algeria

Received 9 May 2012; Accepted 1 September 2012

Abstract. Let (Y, C, X) be a vector of random variables where Y, C and X are, respectively, the interest variable, a right censoring and a covariable (predictor). In this paper, we introduce a new nonlinear wavelet-based estimator of the regression function in the right censorship model. An asymptotic expression for the mean integrated squared error of the estimator is obtained to both continuous and discontinuous curves. It is assumed that the lifetime observations form a stationary $\alpha-$mixing sequence.

Résumé. Soit (Y, C, X) un vecteur de variables aléatoires où Y, C et X sont, respectivement, la variable d'intérêt, une censure à droite et une covariable (prédicteur). Dans cet article, nous introduisons un nouveau estimateur de la fonction de régression basé sur les ondelettes non linéaire dans le modèle de la censure à droite. Une expression asymptotique de l'erreur quadratique moyenne intégrée de l'estimateur est obtenue pour les deux courbes continues et discontinues. On suppose que les observations de la durée de vie forment une suite $\alpha-$mélangeante.

Key words: Censored data; Mean integrated squared error; Nonlinear wavelet-based estimator; Nonparametric regression; Strong mixing condition.
AMS 2010 Mathematics Subject Classification : 62G07; 62G20.

---

1. Introduction

Wavelets and their applications in several areas of both pure and applied mathematics has provided statisticians with powerful new techniques for nonparametric curve estimation by combining recent advances in approximation theory with insights gained from applied signal analysis. Because wavelets are localized in both time and frequency and have remarkable approximation properties, wavelet estimators automatically adapt to these varying degrees of

Djabrane Yahia: yahia_dj@yahoo.fr
Fateh Benatia: fatahbenatia@hotmail.com

regularity (discontinuities, cusps, sharp spikes, etc.) of the underlying curves to be estimated. This is a remarkable property of the wavelet method when compared to other common estimation techniques, such as the kernel method, which may fail in unsmooth situations. The recent monograph by Härdle *et al.* (1998) and the book by Vidakovic (1999) provide excellent systematic discussions on wavelets and their applications in statistics.

Under the assumption that the lifetime observations are mutually independent, the nonlinear wavelet estimator of the density function has first been considered for complete data; see, Hall and Patil (1995). These authors showed that the asymptotic mean integrated squared error ($AMISE$) formula is the same in both smooth and unsmooth density case, a fact that is not true for the kernel method. Similar results are available for the problem of estimating a regression function, see Hall and Patil (1996) for i.i.d. complete data and Truong and Patil (2001) for $\alpha-$mixing complete data. For right censorship model, Li (2003) consider a nonlinear wavelet estimator of a single density function with randomly censored data and derives its $MISE$. Li *et al.* (2008) considers the estimation of the regression function and they showed its convergence rate over a large function class in the i.i.d. setting.

In this paper we consider the right censorship model and we introduce a new nonlinear wavelet-based estimator of the regression function and we investigate the asymptotic expression for the $MISE$ of the estimator. It is assumed that the lifetime observations form a stationary $\alpha-$mixing sequence.

Let Y be a lifetime variable with continuous distribution function (df) F and X a continuous covariable (predictor) taking its values in $[0,1]$ with df L and corresponding density ℓ. In regression analysis one expects to identify, if any, the relationship between the Y_i's and X_i's. This means looking for a function $m^*(X)$ describing this relationship that realizes the minimum of the mean squared error criterion. It is well known that this minimum is achieved by the regression function of Y given $X = x$, that is $m(x) := E(Y|X = x) = h(x)/\ell(x)$, with $h(x) = \int yF(x, dy)$, where $F(\cdot,\cdot)$ being the joint df of the random vector (X, Y) with density $f(\cdot,\cdot)$.

In practice, the response lifetime variable $Y-$ a variable of interest may be subject to left truncation and/or right censoring. As in medical follow–up research, the observation of the time to an event may be prevented by a previous censoring occurrence. Examples of such events include the death of a patient and the relief from symptom. Examples of censoring occurrences include the end of the study and the loss of data caused by failure to follow up. In this case only part of the observations are true death time or real relief time. Wavelet procedures in conjunction with censoring have also been used for detecting change points in several biomedical applications. Typical examples are the detection of life-threatening cardiac arythmias in electrocardiographic signals recorded during the monitoring of patients, or the detection of venous air embolism in doppler heart sound signals recorded during surgery when the incision wounds lie above the heart. The recent work of Härdle *et al.* (1998) provide excellent selective review article on nonlinear wavelet methods in nonparametric curve estimation and their role on a variety of statistical applications.

Consider a real random variable (rv) Y and a strictly stationary rv's $(Y_i)_{i\geq 1}$ with common unknown absolutely continuous df F. In medical research, industrial life-testing, survival analysis and other studies, the rv's my be the lifetime of patient under study. Also let $(C_i)_{i\geq 1}$ be a sequence of censoring rv's with unknown df G. In contrast to statistics for complete data studies, right-censored model involves pairs (T_i, δ_i) where only $T_i := \min(Y_i, C_i) = Y_i \wedge C_i$

and $\delta_i = I(Y_i < C_i)$, $i = 1, 2, ..., n$ are observed, where $I(A)$ denotes the indicator function of the set A.

In this paper, we adapt the wavelet-based regression estimators to right censored data under strong mixing conditions whose definition is given below. First, let $\mathcal{F}_i^k(Z)$ denotes the $\sigma-$field of events generated by $\{Z_j,\ i \le j \le k\}$. For easy reference, let us recall the following definition.

Definition 1. *Let* $\{Z_i,\ i \ge 1\}$ *denotes a sequence of rv's. Given a positive integer* n, set:

$$\alpha(n) = \sup\left\{|\mathbf{P}(A\cap B) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathcal{F}_1^k(Z),\ B \in \mathcal{F}_{k+n}^{\infty}(Z),\ k \in \mathbb{N}^*\right\}.$$

The sequence is said to be $\alpha-$mixing (strongly mixing) if the mixing coefficient $\alpha(n) \to 0$ as $n \to \infty$.

Among various mixing conditions used in the literature, $\alpha-$mixing is reasonably weak and has many practical applications. Many processes do fulfill the strong mixing property. We quote, here, the usual ARMA processes which are geometrically strongly mixing, i.e., there exist $\rho \in (0,1)$ and $\gamma > 1$ such that, for any $n \ge 1$, $\alpha(n) \le \gamma\rho^n$ (see, e.g., Jones, 1978). The $ARCH$ models (see Engle, 1982), their $GARCH$ extension (see Bollerslev, 1986), the threshold models and the $EXPAR$ models (see Ozaki, 1979) are geometrically strongly mixing under some general conditions. We refer the reader to the recent Bradley's monograph Bradley (2007).

In the sequel, $\{T_i, \delta_i, X_i;\ i \ge 1\}$ is assumed to be a stationary $\alpha-$mixing sequence of random vectors with coefficient $\alpha(n)$. Moreover, we suppose that the sequences $\{Y_i,\ i \ge 1\}$ and $\{C_i,\ i \ge 1\}$ are $\alpha-$mixing with coefficient $\alpha_1(n)$ and $\alpha_2(n)$ respectively. Cai (2001, Lemma 2) showed that $\{T_i,\ i \ge 1\}$ is then strongly mixing; with coefficient $\alpha_3(n) = 4\max(\alpha_1(n), \alpha_2(n))$.

The rest of this paper is organized as follows. In Section 2, we give the necessary definitions and define the nonlinear wavelet-based estimators of $m(\cdot)$, $\ell(\cdot)$ and $h(\cdot)$. Assumptions and main results are given in Section 3. The proofs of the main results are postponed to Section 4, where some auxiliary results are also proved. In Appendix, we collect some preliminary lemmas, which are used in the proofs of our main results.

2. Notations and definition of estimators

Our aim is to estimate, $m(\cdot)$, $\ell(\cdot)$ and $h(\cdot)$ by non-linear empirical wavelet coefficients. Let $\phi(x)$ and $\psi(x)$ be father and mother wavelets, having the properties: ϕ and ψ are bounded and compactly supported; $\int \phi^2 = \int \psi^2 = 1$, $\mu_k = \int y^k \psi(y)\,dy = 0$, for $0 \le k \le r-1$ and $\mu_r = r!\kappa$, where $\kappa = (1/r!)\int y^r \psi(y)\,dy$. Therefore, the functions

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

for arbitrary $p > 0$, $i, j \in \mathbb{Z}$, $i \ge 0$ and $p_i = p2^i$ are orthonormal:

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

where δ_{ij} denotes the Kronecker delta, i.e., $\delta_{ij} = 1$, if $i = j$; and $\delta_{ij} = 0$, otherwise. Then, the collection $\{\phi_j(x), \psi_{ij}(x), \ i,j \in \mathbb{Z}, i \geq 0\}$ is an orthonormal basis of $L^2(\mathbb{R})$. For the existence and properties of such wavelet, we refer the reader to Cohen *et al.* (1993); Daubechies (1992) and Härdle *et al.* (1998).

Assume that ϕ and ψ are compactly supported on $[0,1]$. For all function $f \in L^2(\mathbb{R})$, we have the following wavelet expansion:

$$f(x) = \sum_{j=0}^{p-1} a_j \phi_j(x) + \sum_{i=0}^{\infty} \sum_{j=0}^{p_i-1} a_{ij} \psi_{ij}(x), \tag{2}$$

where $a_j = \int f\phi_j$ and $a_{ij} = \int f\psi_{ij}$ are the wavelet coefficients of the function $f(\cdot)$ and the series in (2) converges in $L^2([0,1])$.

As is usual in the wavelet literature we assume that the regression function $m(\cdot)$, density $\ell(\cdot)$ and function $h(\cdot)$ are supported on the unit interval $[0,1]$. In view of (2), the proposed nonlinear wavelet estimators of the covariate density $\ell(\cdot)$ is

$$\hat{\ell}(x) = \sum_{j=0}^{p-1} \hat{a}_j \phi_j(x) + \sum_{i=0}^{q-1} \sum_{j=0}^{p_i-1} \hat{a}_{ij} I(|\hat{a}_{ij}| > \delta) \psi_{ij}(x), \tag{3}$$

where $\delta > 0$ is the "threshold" and $q \geq 1$ is another smoothing parameter, and the empirical wavelet coefficients are defined as follows:

$$\hat{a}_j = \int \phi_j dL_n = n^{-1} \sum_{k=1}^{n} \phi_j(X_k), \qquad \hat{a}_{ij} = \int \psi_{ij} dL_n = n^{-1} \sum_{k=1}^{n} \psi_{ij}(X_k),$$

where $L_n = n^{-1} \sum_{k=1}^{n} I(X_k \leq x)$ is the empirical estimator of the covariat's cumulative df L.

Similarly, as for $\ell(\cdot)$, the wavelet expansion of the function $h(\cdot)$ is given by

$$h(x) = \sum_{j=0}^{p-1} b_j \phi_j(x) + \sum_{i=0}^{\infty} \sum_{j=0}^{p_i-1} b_{ij} \psi_{ij}(x),$$

where $b_j = \int h\phi_j$ and $b_{ij} = \int h\psi_{ij}$. Note that the estimator of the joint df $F(x,y) = P(X \leq x, Y \leq y)$ of (X,Y) under censorship model (see Stute, 1993) is given by

$$F_n(x,y) = n^{-1} \sum_{k=1}^{n} \frac{\delta_k}{\bar{G}_n(T_k)} I\left(X_{(k)} \leq x, T_{(k)} \leq y\right),$$

where $T_{(k)}$ is the $k-th$ ordered $T-$value and $X_{(k)}$ is the concomitant variable associated with the $k-th$ order statistic T_k, i.e., $X_{(k)} = X_j$ if $T_{(k)} = T_j$. Here $\bar{G}_n$ denote the Kaplan and Meier (1958) estimator of the df $\bar{G} := 1 - G$, i.e.,

$$\bar{G}_n(t) = \prod_{i=1}^{n} \left(1 - \frac{1-\delta_i}{n-i+1}\right)^{I(Y_i \leq t)} I\left(Y_{(n)} > t\right).$$

Hence the proposed nonlinear wavelet estimators of $h(x) = \int yF(x,dy)$ is

$$\hat{h}(x) = \sum_{j=0}^{p-1} \hat{b}_j \phi_j(x) + \sum_{i=0}^{q-1} \sum_{j=0}^{p_i-1} \hat{b}_{ij} I\left(\left|\hat{b}_{ij}\right| > \delta\right) \psi_{ij}(x), \tag{4}$$

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

$$\hat{b}_j = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k T_k}{\bar{G}_n(T_k)} \phi_j(X_k), \qquad \hat{b}_{ij} = \frac{1}{n} \sum_{k=1}^{n} \frac{\delta_k T_k}{\bar{G}_n(T_k)} \psi_{ij}(X_k). \tag{5}$$

Further, from (3) and (4) a wavelet estimator of $m(x)$ is given by $\hat{m}(x) = \hat{h}(x)/\hat{\ell}(x)$, with the convention $0/0 = 0$.

3. Assumptions and main results

Throughout this paper, c denotes a positive constant which might take different values at different place. We define the endpoints of F and G by $\tau_F = \sup\{y,\ F(y) < 1\}$, $\tau_G = \sup\{y,\ G(y) < 1\}$ and we assume that $\tau_F < \infty$ and $\bar{G}(\tau_F) > 0$ (this implies $\tau_F < \tau_G$). We point out that since Y can be the lifetime we can suppose it bounded. In addition, we assume that $\{C_i, i \geq 1\}$ and $\{(X_i, Y_i), i \geq 1\}$ are independent. We introduce our assumptions, gathered below for easy reference:

A.1 The joint density $\ell_j(\cdot,\cdot)$ of (X_1, X_{j+1}) exists and satisfies

$$|\ell_j(s,t) - \ell(s)\ell(t)| \leq c, \quad \forall s,t \in [0,1],$$

for some constant c not depending on j.

A.2 The marginal density $\ell(\cdot)$ satisfies $\ell(x) \leq c, \quad \forall x \in [0,1]$.

A.3 The smoothing parameters p, q and δ are functions of n. Suppose that $p \to \infty$, $q \to \infty$ as $n \to \infty$ in such a manner that $p_q \delta^2 = O(n^{-\varepsilon})$ for some $0 < \varepsilon < 1$, $p^{2r+1}\delta^2 \to \infty$, $\delta \geq c\left(n^{-1}\log n\right)^{1/2}$.

A.4 The mixing condition satisfies $\alpha(n) = O\left(n^{-\lambda}\right)$ for some

$$\lambda \geq \max\{(2-\varepsilon)/\varepsilon, 3+4r, 1+(2r+1)/\varepsilon, (\nu-1)(2\nu+1)(2-\varepsilon)/2\varepsilon(\nu-2)\},$$

where $\nu > 2$, and

$$\varepsilon(\lambda + 1 + 2b) + 2b/(2r+1) \geq 2(b+1), \quad \text{for } b > 1.$$

Proposition 1. *Under assumptions (A.1)–(A.4) and the conditions on ϕ and ψ stated in section 2. Assume that the $r-th$ derivatives $\ell^{(r)}$ and $h^{(r)}$ are continuous and bounded. Then*

$$E\left|\int \left(\hat{\ell} - \ell\right)^2 - \left\{n^{-1}p + \kappa^2\left(1 - 2^{-2r}\right)^{-1} p^{-2r} \int \ell^{(r)2}\right\}\right| = o\left(n^{-1}p + p^{-2r}\right). \tag{6}$$

$$E\left|\int \left(\hat{h} - h\right)^2 - \left\{n^{-1}p \iint \frac{y^2 f(x,y)}{\bar{G}(y)} dxdy + \kappa^2\left(1 - 2^{-2r}\right)^{-1} p^{-2r} \int h^{(r)2}\right\}\right| = o\left(n^{-1}p + p^{-2r}\right). \tag{7}$$

Theorem 1. *Under the assumptions of Proposition 1. Let* $r > 1$, *suppose that* $1 - r/(2r+1) < \varepsilon < 2r/(2r+1)$ *and* $p^{2r+1} = O(n)$, *then*

$$\int (\hat{m} - m)^2 = O_p\left(n^{-1}p + p^{-2r}\right). \tag{8}$$

Moreover, if p *is chosen of size* $n^{1/(2r+1)}$, *then* $\int (\hat{m} - m)^2 = O_p\left(n^{-2r/(2r+1)}\right)$.

In Proposition 1 and Theorem 1 we described the performance of wavelet methods for functions ℓ and h with r derivatives. Clearly that smoothness assumption does have a bearing on our results. Nevertheless, the failure of the smoothness condition at a finite number of points does not affect Proposition 1 and also Theorem 1, as our next result shows.

Theorem 2. *Under assumptions (A.1)–(A.4) and the conditions on* ϕ *and* ψ *stated in section 2. Also assume that* $p_q^{2r+1} n^{-2r} \to \infty$, *and impose the condition of* $r-$*times differentiability of* ℓ *and* h *only in a piecewise continuous sense; that is, we ask that there exist points* $0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1$ *such that the first* r *derivatives of* ℓ *and* h *exist and are bounded and continuous on* (x_j, x_{j+1}) *for* $0 \leq j \leq N$, *with left- and right-hand limits. In particular,* ℓ *and* h *themselves my be only piecewise continuous. Then the conclusions (6) and (7) in Proposition 1 still hold, and also (8) in Theorem 1 holds when* $\ell^{(r)}$ *is continuous and bounded.*

Remark 1. Condition *(A.1)* is needed for covariance calculus and takes similar forms to those used in complete data under dependence. Note also that, it is satisfied in the i.i.d. case. Hypothesis *(A.2)* and *(A.3)* are used in de Uña-Álvarez *et al.* (2010) and is needed to establish Lemmas 5–4. Assumptions *(A.4)* concern the mixing processes structure which is standard in such situation. Furthermore, if we replace $\alpha(n) = O\left(n^{-\lambda}\right)$ by $\alpha(n) = O(\rho^n)$ for some $0 < \rho < 1$, then (A.4) *is automatically satisfied.*

Remark 2. The error rates in our Theorems are same as that in Hall and Patil (1996) for i.i.d. complete data, Truong and Patil (2001) for dependent complete data and de Uña-Álvarez *et al.* (2010) for truncated dependent data.

Remark 3. Compared with the kernel estimator, the wavelet analogue of the bandwidth h_n of the kernel estimator is p^{-1}. As point out by Hall and Patil (1996), the $n^{-1}p$ term derives from variance (compare $(nh_n)^{-1}$ in the the kernel estimator case) and the p^{-2r} term from squared bias (compare h_n^{2r} for an $rth-$order kernel estimator), the optimal size of p is $cn^{1/(2r+1)}$. Moreover, by choosing $p \sim n^{1/(2r+1)}$ it can be shown that the MISE satisfy

$$E\int \left(\hat{\ell} - \ell\right)^2 \sim n^{-1}p + \kappa^2\left(1 - 2^{-2r}\right)^{-1} p^{-2r} \int \ell^{(r)2} \sim n^{-2r/(2r+1)},$$

$$E\int \left(\hat{h} - h\right)^2 \sim n^{-1}p \iint \frac{y^2 f(x,y)}{\bar{G}(y)} dxdy + \kappa^2\left(1 - 2^{-2r}\right)^{-1} p^{-2r} \int h^{(r)2} \sim n^{-2r/(2r+1)}.$$

4. Proofs

Observing that the orthogonality of ϕ and ψ implies

$$\begin{aligned}\int \left(\hat{h}-h\right)^2 &= \sum_{j=0}^{p-1}\left(\hat{b}_j-b_j\right)^2+\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 I\left(\left|\hat{b}_{ij}\right|\leq\delta\right)\\ &+\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}\left(\hat{b}_{ij}-b_{ij}\right)^2 I\left(\left|\hat{b}_{ij}\right|>\delta\right)+\sum_{i=q}^{\infty}\sum_{j=0}^{p_i-1} b_{ij}^2\\ &=: I_1+I_2+I_3+I_4.\end{aligned}$$

In order to prove (7), it suffices to bound each term I_1, I_2, I_3 and I_4 separately, which is done in Lemmas 1–4 respectively.

Lemma 1. *Under the assumptions of Proposition 1,*

$$E\left|I_1-n^{-1}p\iint\frac{y^2f(x,y)}{\bar{G}(y)}dxdy\right|=o\left(n^{-1}p\right). \tag{9}$$

Proof. Using Lemma 5 it follows that

$$\begin{aligned}E\left|I_1-n^{-1}p\iint\frac{y^2f(x,y)}{\bar{G}(y)}dxdy\right| &\leq E\left|\sum_{j=0}^{p-1}\left(\tilde{b}_j-b_j\right)^2-n^{-1}p\iint\frac{y^2f(x,y)}{\bar{G}(y)}dxdy\right|\\ &+\sum_{j=0}^{p-1}EB_j^2+2E\sum_{j=0}^{p-1}\left|\tilde{b}_j-b_j\right||B_j|\\ &=: I_{11}+I_{12}+I_{13}.\end{aligned}$$

Firstly, using the conditional expectation property we get

$$\begin{aligned}E\left[\frac{\delta_1T_1}{\bar{G}(T_1)}\middle|X_1=x\right] &= E\left[E\left[\frac{I(Y_1\leq C_1)Y_1}{\bar{G}(Y_1)}\middle|Y_1\right]\middle|X_1=x\right]\\ &=E\left[\frac{Y_1}{\bar{G}(Y_1)}E\Big[I(Y_1\leq C_1)\Big|Y_1\Big]\middle|X_1=x\right]\\ &=E\Big[Y_1\Big|X_1=x\Big]=m(x).\end{aligned}$$

Then we have

$$\begin{aligned}E\tilde{b}_j &= E\left[\frac{\delta_1T_1}{\bar{G}(T_1)}\phi_j(X_1)\middle|X_1\right]=E\left[\phi_j(X_1)E\left[\frac{\delta_1T_1}{\bar{G}(T_1)}\middle|X_1\right]\right]\\ &=\int\phi_j(x)m(x)\ell(x)dx=\int\phi_j(x)h(x)dx=b_j.\end{aligned}$$

Now, for simplicity we set $U_{j,k} := \delta_k T_k \phi_j(X_k) \Big/ \bar{G}(T_k)$. Note that

$$nE\left(\tilde{b}_j - b_j\right)^2 = \frac{1}{n} Var \sum_{k=1}^{n} U_{j,k}$$
$$= Var(U_{j,1}) + 2\sum_{l=1}^{n-1}\left(1 - \frac{l}{n}\right) Cov(U_{j,1}, U_{j,l+1}). \tag{10}$$

We have

$$Var(U_{j,1}) = E\left[U_{j,1}^2\right] - E^2[U_{j,1}] := \mathcal{V}_1 - \mathcal{V}_2.$$

Using again the conditional expectation property, (1) and a change of variable, we get

$$\mathcal{V}_1 = E\left[\phi_j^2(X_1) E\left[\frac{I(Y_1 \le C_1) Y_1^2}{\bar{G}^2(Y_1)}\Big| Y_1\right]\right]$$
$$= \iint \frac{y^2}{\bar{G}(y)} \phi_j^2(x) f(x,y)\, dxdy$$
$$= \iint \frac{y^2}{\bar{G}(y)} \phi^2(t) f((t+j)/p, y)\, dtdy. \tag{11}$$

In other hand

$$\mathcal{V}_2 = \left(\iint y\phi_j(x) f(x,y)\, dxdy\right)^2 = \left(\int \phi_j(x) h(x)\, dx\right)^2$$
$$= \left(p^{-1/2} \int \phi(t) h((t+j)/p)\, dt\right)^2. \tag{12}$$

By $\int \phi^2 = 1$ and the compactness of the support of ϕ we get

$$\sum_{j=0}^{p-1} \mathcal{V}_2 \le C \sum_{j=0}^{p-1} \int p^{-1}\phi^2(t) h^2((t+j)/p)\, dt \to C \int h^2(u)\, du.$$

Hence, from (11), (12) and $\sum_{j=0}^{p-1} p^{-1} f((t+j)/p, y) \to \int f(x,y)\, dx$, it follows

$$\sum_{j=0}^{p-1} Var(U_{j,1}) = p \iint \frac{f(x,y)}{\bar{G}(y)} dxdy + o(p). \tag{13}$$

Now, from $(\tau_F < \tau_G)$ and $\ell(x) \le c$, we have

$$|Cov(U_{j,1}, U_{j,l+1})| = |E[U_{j,1}U_{j,l+1}] - E[U_{j,1}]E[U_{j,l+1}]|$$
$$\le \frac{C}{\bar{G}^2(\tau_F)} \iint |\phi_j(x)\phi_j(y)| \, |\ell_l(x,y) - \ell(x)\ell(y)|\, dxdy$$
$$\le \frac{Cp^{-1}}{\bar{G}^2(\tau_F)} \iint |\phi(s)\phi(t)| \, |\ell_l(s,t) - \ell(s)\ell(t)|\, dsdt.$$

Assumption (A.1) give

$$|Cov\,(U_{j,1}, U_{j,l+1})| = O\left(p^{-1}\right). \tag{14}$$

On the other hand, since $|U_{j,k}| \leq p^{1/2} \|\phi\|_\infty \bar{G}^{-1}(\tau_F) = O\left(p^{1/2}\right)$, from a result in Hall and Heyde (1980, Corollary A.1), we have

$$|Cov\,(U_{j,1}, U_{j,l+1})| = O\,(p\alpha\,(l)). \tag{15}$$

Then to evaluate this covariance term, the idea is to introduce a sequence of integers w_n which we precise below. Then we use (14) for the close 1 and l and (15) otherwise. That is

$$2|\sum_{l=1}^{n-1}(1-l/n)\,Cov\,(U_{j,1}, U_{j,l+1})| \leq c(\sum_{l\leq w_n} + \sum_{l>w_n})\,|Cov\,(U_{j,1}, U_{j,l+1})|.$$

Note that $p_q\delta^2 = O\left(n^{-\varepsilon}\right)$ and $p^{2r+1}\delta^2 \to \infty$, implies $p > cn^{\varepsilon/2r}$. Choosing $w_n = p/\ln\ln(n)$, we have

$$2|\sum_{l=1}^{n-1}(1-l/n)\,Cov\,(U_{j,1}, U_{j,l+1})| \leq c(\sum_{l\leq w_n} + \sum_{l>w_n})\min\left(p^{-1}, p\alpha\,(l)\right) = o\,(1). \tag{16}$$

In the same way as for the term $Var\sum_{j=0}^{p-1}(\tilde{a}_j - a_j)^2$ of de Uña-Álvarez *et al.* (2010) (see the proof of (4.11) in their Appendix, pp 341–344), it can be shown that

$$Var\sum_{j=0}^{p-1}\left(\tilde{b}_j - b_j\right)^2 = o\left(n^{-2}p^{-2}\right). \tag{17}$$

Then (10), (13), (16) and (17) yield that $I_{11} = o\left(n^{-1}p\right)$.

For I_{12}, following the line as for I_{11}, it is easy to see that

$$\begin{aligned}
I_{12} &= O\left(\frac{\ln\ln(n)}{n}\right)\sum_{j=0}^{p-1}E\left[\frac{1}{n}\sum_{k=1}^{n}|\phi_j\,(X_k)|\right]^2 \\
&\leq O\left(\frac{\ln\ln(n)}{n}\right)\sum_{j=0}^{p-1}\left\{E\left[\frac{1}{n}\sum_{k=1}^{n}|\phi_j\,(X_k)| - E\,|\phi_j\,(X_k)|\right]^2 + \left(E\left[|\phi_j\,(X_k)|\right]\right)^2\right\} \\
&= O\left(\frac{\ln\ln(n)}{n}\right)O\left(n^{-1}p\right) + O\left(\frac{\ln\ln(n)}{n}\right) = o\left(n^{-1}p\right).
\end{aligned}$$

Finally, as to I_{13}, we have

$$I_{13} \leq 2\left(\sum_{j=0}^{p-1}E\left[\tilde{b}_j - b_j\right]^2\right)^{1/2}\left(\sum_{j=0}^{p-1}EB_j^2\right)^{1/2} = o\left(n^{-1}p\right).$$

Combining the estimates on I_{11}, I_{12} and I_{13} together, we obtain (9). □

Lemma 2. *Under the assumptions of Proposition 1,*

$$E\left|I_2 - p^{-2r}\kappa^2\left(1-2^{-2r}\right)^{-1}\int h^{(r)^2}\right| = o\left(p^{-2r}\right).$$

Proof. Let $\zeta > 0$, and define

$$I_{21} = \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 I\left(|b_{ij}| \le (1-\zeta)\,\delta\right), \quad I_{22} = \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 I\left(|b_{ij}| \le (1+\zeta)\,\delta\right),$$

$\Delta = \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 I\left(\left|\hat{b}_{ij}-b_{ij}\right| \le \zeta\delta\right)$. Then $I_{21}-\Delta \le I_2 \le I_{21}+\Delta$. By using Taylor expansion, we have

$$\begin{aligned}
b_{ij} &= \int h(x)\,\psi_{ij}(x)\,dx = p_i^{1/2}\int \psi(u)\,h\left(\frac{u+j}{p_i}\right)du \\
&= p_i^{1/2}\int \psi(u)\left\{\sum_{s=0}^{r-1}\frac{1}{s!}(u/p_i)^s h^{(s)}(j/p_i) + \frac{1}{(r-1)!}(u/p_i)^r\int_0^1 (1-t)^{r-1} h^{(s)}\left(\frac{j+tu}{p_i}\right)dt\right\}du \\
&= p_i^{-(r+1/2)}\frac{1}{(r-1)!}\int u^r\psi(u)\int_0^1 (1-t)^{r-1} h^{(r)}\left(\frac{j+tu}{p_i}\right)dtdu \\
&= \kappa p_i^{-(r+1/2)}\left(\mathcal{G}_{ij}+\mathcal{H}_{ij}\right), \qquad (18)
\end{aligned}$$

where $\mathcal{G}_{ij} = h^{(s)}(j/p_i)$ and $\sup_{0\le i\le q-1, 0\le j\le p_i-1}|\mathcal{H}_{ij}| \to 0$.

Note that $\sup_j |b_{ij}| \le cp_i^{-(r+1/2)} \le cp^{-(r+1/2)}$ and $p^{r+1/2}\delta \to \infty$. Hence, for n large enough we have

$$\begin{aligned}
I_{21} = I_{22} &= \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}\kappa^2 p_i^{-(2r+1)}\left(\mathcal{G}_{ij}+\mathcal{H}_{ij}\right)^2 \\
&= p^{-2r}\kappa^2\left(1-2^{-2r}\right)^{-1}\int h^{(r)^2} + o\left(p^{-2r}\right).
\end{aligned}$$

Hence, to prove Lemma 2, it suffice to show that $E\Delta = o(I_{22})$. According to Lemma 5 we have

$$E\Delta \le \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 P\left(\left|\tilde{b}_{ij}-b_{ij}\right| > c_1\zeta\delta\right) + \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} b_{ij}^2 P\left(|B_{ij}| > c_2\zeta\delta\right), \qquad (19)$$

where c_1 and c_2 are positive constants such that $c_1+c_2=1$.
In order to evaluate $E\Delta$, we first use Lemma 7 to bound $P\left(\left|\tilde{b}_{ij}-b_{ij}\right| > c_1\zeta\delta\right)$. Set $\Psi_{ijk} = \frac{\delta_k T_k}{G(T_k)}\psi_{ij}(X_k)$. Then $E\Psi_{ijk} = b_{ij}$, $|\Psi_{ijk} - E\Psi_{ijk}| \le cp_i^{1/2} := s$,

$$E\left(\Psi_{ijk}-E\Psi_{ijk}\right)^2\leq E\Psi_{ijk}^2\leq c \quad \text{and} \quad \left|Cov\left(\Psi_{ijt},\Psi_{iju}\right)\right|=O\left(p_i^{-1}\right) \quad \text{for } t\neq u.$$

Therefore, by Lemma 8, taking $\nu=\infty$, for $N\in\mathbb{N}$, $0<N\leq n/2$ we have

$$D_N=\max_{1\leq l\leq 2N} Var\left(\sum_{k=1}^{l}\Psi_{ijk}\right)\leq cN\left\{\left(p_i^{1/2}\right)^{2/r}\left(p_i^{-1}\right)^{1-1/r}+c\right\}\leq cN. \qquad (20)$$

Assumption (A.3) and $\lambda\geq(2-\epsilon)/\epsilon$ imply $p_i^{\lambda+1}\delta^{2(\lambda-1)}<p_q^{\lambda+1}\delta^{2(\lambda-1)}\to 0$. So, according to Lemma 7, taking $N=\left\lceil\left(p_i\delta^2\right)\right\rceil$, it follows that

$$\begin{aligned}P\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_1\zeta\delta\right)&=P\left(\sum_{k=1}^{n}\left(\Psi_{ijk}-E\Psi_{ijk}\right)>nc_1\zeta\delta\right)\\&\leq 4\exp\left\{-\frac{\left(nc_1\zeta\delta\right)^2/16}{nN^{-1}D_N+Cnc_1\zeta\delta Ns}\right\}+\frac{32s}{nc_1\zeta\delta}n\alpha\left(N\right)\\&\leq 4\exp\left\{-C\delta^2n\right\}+C\left(p_i^{(\lambda+1)/2}\delta^{\lambda-1}\right)\to 0. \qquad (21)\end{aligned}$$

By using arguments similar to those behind (20), it follows that

$$Var\left(\sum_{k=1}^{n}\frac{\delta_kT_k}{\bar{G}(T_k)}\psi_{ij}\left(X_k\right)\right)\leq cn.$$

Hence

$$\begin{aligned}EB_{ij}^2&=O\left(\frac{\ln\ln(n)}{n}\right)E\left(\frac{1}{n}\sum_{k=1}^{n}\frac{\delta_kT_k\left|\psi_{ij}\left(X_k\right)\right|}{\bar{G}(T_k)}\right)^2\\&=O\left(\frac{\ln\ln(n)}{n}\right)\left\{Var\left(\frac{1}{n}\sum_{k=1}^{n}\frac{\delta_kT_k\left|\psi_{ij}\left(X_k\right)\right|}{\bar{G}(T_k)}\right)+E\left(\frac{\delta_1T_1\left|\psi_{ij}\left(X_1\right)\right|}{\bar{G}(T_1)}\right)^2\right\}\\&=O\left(\frac{\ln\ln(n)}{n}\right)\left\{1/n+1/p_i\right\}=o\left(1/n\right). \qquad (22)\end{aligned}$$

From (19), (21) and (22), and the fact that $n\delta^2\to\infty$, it yields that

$$E\Delta\leq o\left(\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2\right)+\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}b_{ij}^2\frac{EB_{ij}^2}{c_2^2\zeta^2\delta^2}=o\left(I_{22}\right).$$

This finishes the proof of Lemma 2. □

Lemma 3. *Under the assumptions of Proposition 1,*

$$E\left(I_3\right)=o\left(n^{-2r/(2r+1)}\right).$$

Proof. Let c_3, c_4 denote positive numbers satisfying $c_3 + c_4 = 1$. Then, from

$$I\left(\left|\hat{b}_{ij}\right| > \delta\right) \le I\left(|b_{ij}| > c_3\delta\right) + I\left(\left|\hat{b}_{ij} - b_{ij}\right| > c_4\delta\right),$$

we have

$$\begin{aligned}E\left(I_3\right) &\le \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\hat{b}_{ij} - b_{ij}\right)^2 I\left(|b_{ij}| > c_3\delta\right)\right] + \sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\hat{b}_{ij} - b_{ij}\right)^2 I\left(\left|\hat{b}_{ij} - b_{ij}\right| > c_4\delta\right)\right]\\ &= I_{31} + I_{32}.\end{aligned}$$

According to Lemma 5, it follows that

$$\begin{aligned}I_{31} &\le 2\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij} - b_{ij}\right)^2 I\left(|b_{ij}| > c_3\delta\right)\right] + 2\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[B_{ij}^2 I\left(|b_{ij}| > c_3\delta\right)\right]\\ &= I_{311} + I_{312}.\end{aligned}$$

The proof of (20) shows that $E\left(\tilde{b}_{ij} - b_{ij}\right)^2 \le C/n$, and (18) implies $\sup_j |b_{ij}| \le c p_i^{-(r+1/2)}$. Therefore, from $n^{1/2}\delta \to \infty$, we find

$$\begin{aligned}I_{311} &= O\left(n^{-1}\right)\sum_{i=0}^{q-1} p_i I\left(p_i \le \left(c/c_3\delta\right)^{2/(2r+1)}\right)\\ &= O\left(n^{-1}\delta^{-2/(2r+1)}\right) = o\left(n^{-2r/(2r+1)}\right). \qquad (23)\end{aligned}$$

Note that $p_q\delta^2 = O\left(n^{-\epsilon}\right)$ and $\delta \ge c\left(\ln(n)/n\right)^{1/2}$ implies that $q = O\left(\ln(n)\right)$ and $p_q \ln(n)/n \to 0$. Then by using (22) we have

$$\begin{aligned}I_{312} &= O\left(\frac{\ln\ln(n)}{n}\right)\sum_{i=0}^{q-1}\left(\frac{p_i}{n} + 1\right) = O\left(\frac{\ln\ln(n)}{n}\right)\left(\frac{p_q}{n} + q\right)\\ &= o\left(n^{-2r/(2r+1)}\right). \qquad (24)\end{aligned}$$

Equations (23) and (24) yield that $I_{31} = o\left(n^{-2r/(2r+1)}\right)$.

As to I_{32}, let c_5, c_6 denote positive numbers satisfying $c_5 + c_6 = 1$. On applying Lemma 5 again, we have

$$\begin{aligned}I_{32} &\le 2\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij} - b_{ij}\right)^2 I\left(|B_{ij}| > c_4c_6\delta\right)\right]\\ &+ 2\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} EB_{ij}^2 + 2\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij} - b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij} - b_{ij}\right| > c_4c_5\delta\right)\right]. \qquad (25)\end{aligned}$$

Observe that

$$
\begin{aligned}
&\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(|B_{ij}|>c_4c_6\delta\right)\right]\\
&=\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(|B_{ij}|>c_4c_6\delta,\ \left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]\\
&+\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(|B_{ij}|>c_4c_6\delta,\ \left|\tilde{b}_{ij}-b_{ij}\right|\le c_4c_5\delta\right)\right]\\
&\le\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]+C\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} EB_{ij}^2,
\end{aligned}
$$

which, together with (25) and the proof of (24), leads to

$$
\begin{aligned}
I_{32}&\le 3\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]+C\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} EB_{ij}^2\\
&\le 3\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]+o\left(n^{-2r/(2r+1)}\right).
\end{aligned}
$$

Therefore, it suffice to show that

$$
\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]=o\left(n^{-2r/(2r+1)}\right). \tag{26}
$$

Let a denote a positive number such that $1/a+1/b=1$. By using Lemma 9 and (21), according to Hölder's inequality, we get

$$
\begin{aligned}
&\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1} E\left[\left(\tilde{b}_{ij}-b_{ij}\right)^2 I\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right]\\
&\le\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}\left(E\left|\tilde{b}_{ij}-b_{ij}\right|^{2a}\right)^{1/a}\left(P\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c_4c_5\delta\right)\right)^{1/b}\\
&\le c\sum_{i=0}^{q-1}\sum_{j=0}^{p_i-1}\frac{1}{n}\left\{\exp\left(-c_1\delta^2 n\right)+\left(p_i^{\lambda+1}\delta^{2(\lambda-1)}\right)^{1/(2b)}\right\}\\
&\le c\frac{p_q}{n}\exp\left(-c_1\delta^2 n\right)+cn^{-1}p_q^{(\lambda+1)/(2b)+1}\delta^{(\lambda-1)/b}.
\end{aligned}
$$

By choosing $\delta\ge c_2\left(\ln(n)/n\right)^{1/2}$ with c_2 is such that $c_1c_2=2r/(2r+1)$, and by noticing that $p_q\delta^2=O\left(n^{-\epsilon}\right)$, $\delta\ge c_3\left(\ln(n)/n\right)^{1/2}$ and $\epsilon\left(\lambda+1+2b\right)+2b/(2r+1)\ge 2\left(b+1\right)$ imply $n^{-1}p_q^{(\lambda+1)/(2b)+1}\delta^{(\lambda-1)/b}=o\left(n^{-2r/(2r+1)}\right)$. Combining (23), (24) and (26), we have proved the lemma. □

Lemma 4. *Under the assumptions of Proposition 1,*

$$I_4 = \sum_{i=q}^{\infty} \sum_{j=0}^{p_i-1} b_{ij}^2 = o\left(p^{-2r}\right).$$

Proof. From (18), it follows that

$$\begin{aligned} I_4 &= \sum_{i=q}^{\infty} \sum_{j=0}^{p_i-1} \kappa^2 p_i^{-(2r+1)} \left(\mathcal{G}_{ij} + \mathcal{H}_{ij}\right)^2 \leq 2\kappa^2 \sum_{i=q}^{\infty} p_i^{-(2r+1)} \sum_{j=0}^{p_i-1} \mathcal{G}_{ij}^2 \\ &= O\left(p_q^{-2r}\right) = o\left(p^{-2r}\right). \end{aligned}$$

We achieve the proof. □

Proof of Proposition 1. The proof of (7) follows from Lemmas 1–4. The proof of (6) concerning the covariate density is a particular case of (7) (it suffice to take $\delta_i = 1, T_i = 1, \bar{G}(T_i) = 1$; $i \geq 1$, i.e., without censoring).
We are now in a position to give the proof of Theorem 1. □
Proof of Theorem 1. We use the following classical decomposition

$$\hat{m}(x) - m(x) = \frac{\hat{h}(x) - h(x)}{\hat{\ell}(x)} + m(x) \cdot \frac{\ell(x) - \hat{\ell}(x)}{\hat{\ell}(x)}.$$

Then

$$\begin{aligned} \int \left(\hat{m}(x) - m(x)\right)^2 dx \leq & \frac{2}{\beta - \sup_{x\in[0,1]} \left|\hat{\ell}^2(x) - \ell^2(x)\right|} \\ & \times \left\{ \int \left(\hat{h}(x) - h(x)\right)^2 dx + \beta^{-1} \sup_{x\in[0,1]} \left(m(x)\right)^2 \int \left(\hat{\ell}(x) - \ell(x)\right)^2 dx \right\}, \end{aligned}$$

where $\beta = \inf_{x\in[0,1]} \ell^2(x)$. Recall that, by using the fact that $|\omega| = O_p\left(E|\omega|\right)$ for any rv ω, Proposition **1** yield that

$$\int \left(\hat{\ell}(x) - \ell(x)\right)^2 dx = O_p\left(n^{-1}p + p^{-2r}\right), \quad \int \left(\hat{h}(x) - h(x)\right)^2 dx = O_p\left(n^{-1}p + p^{-2r}\right).$$

From assumption (A.2) it suffices to show that

$$\sup_{x\in[0,1]} \left|\hat{\ell}(x) - \ell(x)\right| = o_p(1). \tag{27}$$

The proof of (27) is analogous to that given in the proof of Theorem 3.1 of de Uña-Álvarez *et al.* (2010, pages: 331–332) concerning the covariate density under truncation, therefore, it is omitted. □

Proof of Theorem 2. The proof is analogous to Proposition 1 and Theorem 1, we prove only (7), the proof of (6) is similar, and (8) is a consequence of (6) and (7).

Observe that, by the orthogonality properties of ϕ and ψ, $\int\left(\hat{h}-h\right)^2=\mathcal{I}_q\left(\mathbb{Z},\mathbb{Z},...\right)$, where $\mathbb{Z}$ denotes the set of all integers and

$$\begin{aligned}\mathcal{I}_q\left(\mathcal{J},\mathcal{J}_0,\mathcal{J}_1,...\right)&=\sum_{j\in\mathcal{J}}\left(\hat{b}_j-b_j\right)^2+\sum_{i=0}^{q-1}\sum_{j\in\mathcal{J}_i}b_{ij}^2I\left(\left|\hat{b}_{ij}\right|\le\delta\right)\\&+\sum_{i=0}^{q-1}\sum_{j\in\mathcal{J}_i}\left(\hat{b}_{ij}-b_{ij}\right)^2I\left(\left|\hat{b}_{ij}\right|>\delta\right)+\sum_{i=q}^{\infty}\sum_{j\in\mathcal{J}_i}b_{ij}^2\\&:=\mathcal{I}_1\left(\mathcal{J}\right)+\mathcal{I}_2\left(\mathcal{J}_0,\mathcal{J}_1,...\right)+\mathcal{I}_3\left(\mathcal{J}_0,\mathcal{J}_1,...\right)+\mathcal{I}_4\left(\mathcal{J}_0,\mathcal{J}_1,...\right),\end{aligned}$$

where $\mathcal{J}_i=\{0,1,...,p_i-1\}$. When $h(\cdot)$ is only piecewise continuous, let $\mathcal{X}$ denote the finite set of points where $h^{(s)}$ has point of discontinuities for some $0\le s\le r$. If supp$\phi\subseteq(-u,u)$, then, unless

$$j\in\mathcal{K}=\{k:k\in(py-u,py-u)\ \text{ for some } y\in\mathcal{X}\},$$

both b_j and $\hat{b}_j$ are constructed entirely from an integral over or an average of data values from an interval where $h^{(r)}$ exists and is bounded and continuous. Likewise, if supp$\phi\subseteq(-u,u)$, then, unless

$$j\in\mathcal{K}_i=\{k:k\in(p_iy-u,p_iy-u)\ \text{ for some } y\in\mathcal{X}\},$$

b_j and $\hat{b}_j$ are also constructed solely from such regions. Then we may write

$$\int\left(\hat{h}-h\right)^2=\mathcal{I}_q\left(\mathcal{K},\mathcal{K}_0,\mathcal{K}_1,...\right)+\mathcal{I}_q\left(\tilde{\mathcal{K}},\tilde{\mathcal{K}}_0,\tilde{\mathcal{K}}_1,...\right),$$

where $\tilde{\mathcal{K}}$ and $\tilde{\mathcal{K}}_i$ denotes the complements of $\mathcal{K}$ and $\mathcal{K}_i$ in $\mathbb{Z}$. The proof of (10) shows $E\left(\tilde{b}_j-b_j\right)^2=O\left(n^{-1}\right)$, the evaluation for I_{12} shows $EB_j^2=O\left(\ln\ln(n)/n\right)\left(n^{-1}+p^{-1}\right)$. Furthermore, noting that both $\mathcal{K}$ and $\mathcal{K}_i$ have no more than $(2u+1)(\#\mathcal{X})$ elements for each i. Then by Lemma 5 it follows that

$$\begin{aligned}E\mathcal{I}_1\left(\mathcal{K}\right)&\le2\sum_{j\in\mathcal{K}}EB_j^2+2\sum_{j\in\mathcal{K}}E\left(\tilde{b}_j-b_j\right)^2\\&=O\left(\ln\ln(n)/n\right)\left(n^{-1}+p^{-1}\right)+O\left(n^{-1}\right)=o\left(n^{-2r/(2r+1)}\right).\end{aligned}$$

Note that

$$\begin{aligned}E\mathcal{I}_2\left(\mathcal{K}_0,\mathcal{K}_1,...\right)&\le\sum_{i=0}^{q-1}\sum_{j\in\mathcal{K}_i}b_{ij}^2I\left(|b_{ij}|\le(1-\zeta)\delta\right)+\sum_{i=0}^{q-1}\sum_{j\in\mathcal{K}_i}b_{ij}^2I\left(\left|\tilde{b}_{ij}\right|>\zeta\delta\right)\\&=O\left(p\delta^2\right)+\sum_{i=0}^{q-1}\sum_{j\in\mathcal{K}_i}p_i^{-1}\left\{P\left(\left|\tilde{b}_{ij}-b_{ij}\right|>c\delta\right)+b_{ij}^2P\left(|B_{ij}|>c\delta\right)\right\}.\quad(28)\end{aligned}$$

From $\delta \geq c\left(n^{-1}\log n\right)^{1/2}$ we have

$$E\left|B_{ij}\right|\delta^{-1} \leq c\left(\frac{\ln\ln(n)}{n\delta^2}\right)^{1/2} E\left(\frac{\delta_1 T_1\left|\psi_{ij}(X_1)\right|}{\bar{G}(T_1)}\right) \leq c\left(\frac{\ln\ln(n)}{p\ln(n)}\right)^{1/2} \to 0,$$

hence similarly to the proof as for (21) one can verify that

$$P\left(\left|B_{ij}\right| > c\delta\right) \leq P\left(\left|B_{ij} - EB_{ij}\right| > c\delta\right) \leq 4\exp\left\{-c\delta^2 n\right\} + c\left(p_i^{(\lambda+1)/2}\delta^{\lambda-1}\right).$$

Therefore, in view of $p_q^{2r+1}n^{-2r} \to \infty$ and (A.3), from (28) and

$$n^{2r/(2r+1)}p_q^{(\lambda+1)/2}\delta^{\lambda-1} \leq cn^{-(\varepsilon(\lambda-1)/2-2r/(2r+1))} \to 0$$

since $\varepsilon(\lambda-1)/2 - 2r/(2r+1) > b$ by $\lambda \geq 1 + (2r+1)/\varepsilon$ we have

$$\begin{aligned} E\mathcal{I}_2\left(\mathcal{K}_0, \mathcal{K}_1, ...\right) &\leq O\left(q\left(p_q\delta^2\right)p_q^{-1}\right) + C\sum_{i=0}^{q-1} p_i^{-1}\left(\exp\left\{-C\delta^2 n\right\} + p_i^{(\lambda+1)/2}\delta^{\lambda-1}\right) \\ &\leq o\left(n^{-2r/(2r+1)}\right) + cp^{-1}\exp\left\{-C\delta^2 n\right\} + Cp_q^{(\lambda+1)/2}\delta^{\lambda-1} = o\left(n^{-2r/(2r+1)}\right). \end{aligned}$$

By Lemma 5 and Lemma 6, from (22) it follows that

$$E\mathcal{I}_3\left(\mathcal{K}_0, \mathcal{K}_1, ...\right) \leq 2\sum_{i=0}^{q-1}\sum_{j\in\mathcal{K}_i}\left\{EB_{ij}^2 + E\left(\tilde{b}_{ij} - b_{ij}\right)^2\right\} = O(q/n) = o\left(n^{-2r/(2r+1)}\right).$$

Thus $\mathcal{I}_1(\mathcal{K}) + \mathcal{I}_2(\mathcal{K}_0, \mathcal{K}_1, ...) + \mathcal{I}_3(\mathcal{K}_0, \mathcal{K}_1, ...)$ is negligible compared to the main terms of $MISE$. In view of $b_{ij} = O\left(p_i^{-1/2}\right)$ and $p_q^{2r+1}n^{-2r} \to \infty$ we have $\mathcal{I}_3(\mathcal{K}_0, \mathcal{K}_1, ...) = O\left(p_q^{-1}\right) = o\left(n^{-2r/(2r+1)}\right)$. The methods in the proof of Proposition 1 may be employed to prove that $\mathcal{I}_q\left(\tilde{\mathcal{K}}, \tilde{\mathcal{K}}_0, \tilde{\mathcal{K}}_1, ...\right)$ has precisely the asymptotic properties claimed for $\int\left(\hat{h} - h\right)^2$ in Proposition 1. □

Acknowledgements The authors thank the referee for his/her comments and valuable suggestions, which have greatly improved the quality of the paper.

Appendix A:

In this section we give some preliminary lemmas which have been used in the proofs of our main results. Let $\{Z_i,\ i \geq 1\}$ be a stationary $\alpha-$mixing sequence of real rv's with the mixing coefficients $\alpha(n)$ (see, Definition 1).

Lemma 5. *Set*

$$\tilde{b}_j = \frac{1}{n}\sum_{k=1}^{n}\frac{\delta_k T_k}{\bar{G}(T_k)}\phi_j(X_k), \qquad \tilde{b}_{ij} = \frac{1}{n}\sum_{k=1}^{n}\frac{\delta_k T_k}{\bar{G}(T_k)}\psi_{ij}(X_k). \tag{A1}$$

Then under the assumption $\alpha(n) = O(n^{-r})$ for some $r > 3$, we have $\hat{b}_j = \tilde{b}_j + B_j$ and $\hat{b}_{ij} = \tilde{b}_{ij} + B_{ij}$, where

$$|B_j| = O\left(\sqrt{\ln\ln(n)/n}\right) \frac{1}{n} \sum_{k=1}^{n} |\phi_j(X_k)| \, |Y_k| \quad a.s, \tag{A2}$$

$$|B_{ij}| = O\left(\sqrt{\ln\ln(n)/n}\right) \frac{1}{n} \sum_{k=1}^{n} |\psi_{ij}(X_k)| \, |Y_k| \quad a.s. \tag{A3}$$

Proof. We have from (5) and (A1)

$$\begin{aligned}
\left|\hat{b}_j - \tilde{b}_j\right| &= \frac{1}{n}\left|\sum_{k=1}^{n} \frac{I(Y_k < C_k) Y_k}{\bar{G}_n(Y_k)} \phi_j(X_k) - \sum_{k=1}^{n} \frac{I(Y_k < C_k) Y_k}{\bar{G}(Y_k)} \phi_j(X_k)\right| \\
&\le \frac{1}{n} \sum_{k=1}^{n} |Y_k| \, |\phi_j(X_k)| \left|\frac{\bar{G}(Y_k) - \bar{G}_n(Y_k)}{\bar{G}_n(Y_k)\bar{G}(Y_k)}\right| \\
&\le \frac{1}{\bar{G}_n(\tau_F)\bar{G}(\tau_F)} \sup_{t\le\tau_F} \left(\left|\bar{G}_n(t) - \bar{G}(t)\right|\right) \frac{1}{n} \sum_{k=1}^{n} |Y_k| \, |\phi_j(X_k)| \, .
\end{aligned}$$

In the same way as for Theorem 2 of Cai (2001), it can be shown that

$$\sup_{t\le\tau_F} \left(\left|\bar{G}_n(t) - \bar{G}(t)\right|\right) = O\left(\sqrt{\ln\ln(n)/n}\right) \quad a.s.$$

Similarly, we get (A3). □

Lemma 6 (Hall and Heyde (1980, Corollary A.1)). *Suppose that X and Y are rv's such that $|X| < c_1$ and $|Y| < c_2$. Then*

$$|EXY - EXEY| \le 4c_1c_2 \left\{ \sup_{A\in\sigma(X), B\in\sigma(Y)} |P(XY) - P(X)P(Y)| \right\}.$$

Lemma 7 (Liebscher (2001, Proposition 5.1)). *Assume that $EZ_1 = 0$, $EZ_1^2 < \infty$ and $|Z_j| \le s < \infty$ a.s. $(j = 1, ..., n)$. Then, for $n, N \in \mathbb{N}$, $0 < N \le n/2$, for $\epsilon > 0$,*

$$P\left(\left|\sum_{j=1}^{n} Z_j\right| > \epsilon\right) \le 4\exp\left\{-\frac{\epsilon^2/16}{nN^{-1}D_N + \epsilon s N/3}\right\} + 32\frac{s}{\epsilon} n\alpha(N),$$

where $D_N = \max_{1\le l\le 2N} Var\left(\sum_{j=1}^{l} Z_j\right)$.

Lemma 8 (Liebscher (1996, Lemma 2.3)). *Assume* $\alpha(n) \le c_1 n^{-r}$, *for some* $c_1 > 0$, $r > 1$. *Let* $\sup_{1\le i;j\le n, i\neq j} |Cov(Z_i, Z_j)| := R^*(n) < \infty$ *be satisfied. Moreover, let* $R_\nu(n) < \infty$ *for some* ν, $2r/(r-1) < \nu \le \infty$, *where* $R_\nu(n) = \sup_{1\le j\le n} (E|Z_j|^\nu)^{1/\nu}$ *for* $1 \le \nu < \infty$ *and* $R_\infty(n) = \sup_{1\le j\le n} ess\ \sup_{\omega\in\Omega} |Z_j|$. *Then*

$$Var\left(\sum_{j=1}^{n} Z_j\right) \le n\left\{c_{(r,\nu)} (R_\nu(n))^{2\nu/(\nu-2)r} (R^*(n))^{1-\nu/(\nu-2)r} + R_2^2(n)\right\}$$

holds with $c_{(r,\nu)} := \frac{20r-40r/\nu}{r-1-2r/\nu} c_1^{1/r}$ *is a constant depending on* r,ν *only.*

In order to obtain the bounds on the term I_{32} in Lemma 3, we also need the following Lemma.

Lemma 9. *Under the assumptions of Lemma 3. Let* $\nu > 2$, *if* $\lambda \ge (\nu-1)(2\nu+1)(2-\varepsilon)/2\varepsilon(\nu-2)$, *then* $E|\tilde{a}_{ij} - a_{ij}|^\nu = O(n^{-\nu/2})$, $\quad E\left|\tilde{b}_{ij} - b_{ij}\right|^\nu = O(n^{-\nu/2})$.

Proof. Following the lines of Lemma 4.5 in Liang *et al.* (2005) or that of de Uña-Álvarez *et al.* (2010, Lemma A.7), we can verify the Lemma. We only prove the second equation, the proof of first equation is analogous. Choosing $r(n) = [(n/p_q)^{(\nu-2)/2(\nu-1)}]$, and positive integers $k(n)$ and $\gamma(n)$ such that $n = r(n)k(n) + \gamma(n)$, with $0 \le \gamma(n) \le r(n)$. Set $\mathcal{M}_k := \frac{1}{n}\left(\delta_k T_k \psi_{ij}(X_k)/\bar{G}(T_k) - b_{ij}\right)$. Then

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

The contribution of the remainder last term is negligible (and is subsequently ignored). So, without loss of generality, we assume $\gamma(n) = 0$, and further $k(n) = 2s(n)$. Then

$$\begin{aligned}\tilde{b}_{ij} - b_{ij} &= \sum_{l=1}^{2s(n)} \sum_{j=(l-1)r(n)+1}^{lr(n)} \mathcal{M}_j =: \sum_{l=1}^{2s(n)} \zeta_n(l) \\ &= \sum_{l=1}^{2s(n)} \zeta_n(2l) + \sum_{l=1}^{2s(n)} \zeta_n(2l-1) =: S(n) + T(n).\end{aligned} \tag{A4}$$

Hence $E\left|\tilde{b}_{ij} - b_{ij}\right|^\nu \le C\{E|S(n)|^\nu + E|T(n)|^\nu\}$. Next, we evaluate only $E|T(n)|^\nu$, since the evaluation of $E|S(n)|^\nu$ is similar. In view of Theorem 3 of Bradley (1983), there exist i.i.d. rv's $\zeta_n^*(2l-1)$, $l = 1, 2, ..., s(n)$ such that $\zeta_n^*(2l-1)$ has the same distribution as $\zeta_n(2l-1)$ for each l, and satisfies

$$P(|\zeta_n^*(2l-1) - \zeta_n(2l-1)| \ge \varepsilon_l) \le 18\left(\frac{\|\zeta_n(2l-1)\|_\infty}{\varepsilon_l}\right)^{1/2} \alpha(r(n)), \tag{A5}$$

where $0 < \varepsilon_l \leq \|\zeta_n(2l-1)\|_\infty$, if $\|\zeta_n(2l-1)\|_\infty > 0$ and $\varepsilon_l > 0$, if $\|\zeta_n(2l-1)\|_\infty = 0$. Then

$$E|T(n)|^\nu \leq c\{E|\sum_{l=1}^{2s(n)}(\zeta_n^*(2l-1)-\zeta_n(2l-1))|^\nu + E|\sum_{l=1}^{2s(n)}\zeta_n^*(2l-1)|^\nu\}$$
$$= c\{T_1(n)+T_2(n)\}.$$

Let us take $N_n > 0$ such that $s(n)N_n \asymp n^{-1/2}$, where $\alpha_n \asymp \beta_n$ means $0 < \liminf \alpha_n/\beta_n \leq \limsup \alpha_n/\beta_n < \infty$, and assume $\|\zeta_n(2l-1)\|_\infty \geq N_n$, for $l = 1,2,...,s(n)$. Otherwise, by rearranging the terms appropriately, we may assume that $\|\zeta_n(2l-1)\|_\infty \geq N_n$, for $l = 1,2,...,s(n)$, and $\|\zeta_n(2l-1)\|_\infty < N_n$, for $l = s_1(n)+1,...,s(n)$, where $s_1(n)$ is a positive integer with $s_1(n) \leq s(n)$, then

$$T_1(n) \leq c\{(N_n s(n))^\nu + E(\sum_{l=1}^{2s(n)}|\zeta_n^*(2l-1)-\zeta_n(2l-1)|)^\nu\}.$$

Therefore,

$$T_1(n) \leq c\{(N_n s(n))^\nu + E(\sum_{l=1}^{2s(n)}|\zeta_n^*(2l-1)-\zeta_n(2l-1)|I(|\zeta_n^*(2l-1)-\zeta_n(2l-1)| \geq N_n))^\nu\},$$

where $\|\zeta_n(2l-1)\|_\infty \geq N_n$. Observe that

$$|\zeta_n^*(2l-1)-\zeta_n(2l-1)| \leq 2r(n)\left(\frac{cp_i^{1/2}\|\psi\|_\infty}{\bar{G}_n(\tau_F)} + |b_{ij}|\right)\frac{1}{n} \leq cn^{-1}2r(n)p_q^{1/2}.$$

Note that, assumptions (A.3) and (A.4) imply $n^{-\frac{\lambda(\nu-2)}{2(\nu-1)}+\frac{1}{4}}p_q^{\frac{\lambda(\nu-2)}{2(\nu-1)}+\frac{\nu}{2}+\frac{1}{4}} = o\left(n^{-\nu/2}\right)$. Then, according to (A5) and $N_n s(n) = O\left(n^{-1/2}\right)$, it follows that

$$T_1(n) \leq O\left(n^{-\nu/2}\right) + c\left(n^{-1}r(n)p_q^{1/2}\right)^\nu (s(n))^{\nu-1}\sum_{l=1}^{s(n)}P(|\zeta_n^*(2l-1)-\zeta_n(2l-1)| \geq N_n)$$
$$\leq O\left(n^{-\nu/2}\right) + c\left(n^{-1}r(n)p_q^{1/2}\right)^\nu (s(n))^\nu \left((nN_n)^{-1}r(n)p_q^{1/2}\right)^{1/2}(r(n))^{-\lambda}$$
$$\leq cn^{-\frac{\lambda(\nu-2)}{2(\nu-1)}+\frac{1}{4}}p_q^{\frac{\lambda(\nu-2)}{2(\nu-1)}+\frac{\nu}{2}+\frac{1}{4}} + O\left(n^{-\nu/2}\right) = O\left(n^{-\nu/2}\right).$$

Next, we estimate $T_2(n)$. Applying the Rosenthal inequality for sums of independent rv's (see, Petrov, 1995, Theorem 2.9, page 59), we get

$$T_2(n) \leq c\{\sum_{l=1}^{s(n)}E|\zeta_n^*(2l-1)|^\nu + (\sum_{l=1}^{s(n)}E(\zeta_n^*(2l-1))^2)^{\nu/2}\}$$
$$\leq c\left\{s(n)E|\zeta_n(1)|^\nu + \left(s(n)E(\zeta_n(1))^2\right)^{\nu/2}\right\}. \quad \text{(A6)}$$

From $(\tau_F < \tau_G)$, we have

$$\begin{aligned} E|\zeta_n(1)|^\nu &= E|\sum_{k=1}^{r(n)} \mathcal{M}_k|^\nu \le (r(n))^\nu E|\mathcal{M}_1|^\nu \\ &\le c(r(n))^\nu n^{-\nu} E|\psi_{ij}(X_1)|^\nu \\ &\le c(r(n))^\nu n^{-\nu} p_i^{\nu/2-1} \int |\psi(t)|^\nu \ell\left(\frac{t+j}{p_i}\right) dt \\ &\le c(r(n))^\nu n^{-\nu} p_q^{\nu/2-1}. \end{aligned}$$

Then

$$s(n) E|\zeta_n(1)|^\nu = O\left(n^{-\nu/2}\right). \tag{A7}$$

As to $E(\zeta_n(1))^2$, by using Lemma 8, it follows that

$$E(\zeta_n(1))^2 = E|\sum_{k=1}^{r(n)} \mathcal{M}_k|^2 \le r(n)\left\{c(R_\infty(r(n))^{2/\lambda}(R^*(r(n)))^{1-1/\lambda} + R_2^2(r(n))\right\},$$

where

$$\begin{aligned} R_\infty(r(n)) &:= \sup_{1\le j\le n} ess\sup_{\omega\in\Omega} |\mathcal{M}_k| \le c\left(\frac{Cp_i^{1/2}\|\psi\|_\infty}{\bar{G}_n(\tau_F)} + |b_{ij}|\right)\frac{1}{n} = O\left(n^{-1}p_q^{1/2}\right), \\ R_2^2(r(n)) &:= E|\mathcal{M}_1|^2 \le cn^{-2}\int |\psi(t)|^2 \ell\left(\frac{t+j}{p_i}\right) dt = O\left(n^{-2}\right), \\ R^*(r(n)) &:= \sup_{1\le i;j\le r(n), i\ne j} |Cov(\mathcal{M}_i, \mathcal{M}_j)| \le cn^{-2}p_q^{-1}. \end{aligned}$$

Hence, $E(\zeta_n(1))^2 \le cr(n)n^{-2}$ and $s(n)E(\zeta_n(1))^2 \le cs(n)r(n)n^{-2} = O\left(n^{-1}\right)$, which, together with (A6) and (A7), yields $T_2(n) = O\left(n^{-\nu/2}\right)$. This finishes the proof. □

References

Antoniadis, A., 2007. Wavelet methods in statistics: Some recent developments and their applications. *Statistics Surveys.* Vol. **1**, 16-55.

Bollerslev, T., 1986. General autoregressive conditional heteroskedasticity. *J. Economt.* **31**, 307–327.

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

Bradley, R.C., 2007. *Introduction to strong mixing conditions.* Vol I-III, Kendrick Press, Utah.

Cai, Z., 2001. Estimating a distribution function for censored time series data. *J. Multivariate Anal.* **78**, 299-318.

Cohen, A., Daubechies, I. and Vial, P., 1993. Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.* **1**, 54-82.

Daubechies, I., 1992. *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica.* **50**, 987-1007.

Hall, P. and Heyde, C.C., 1980. *Martingale Limit Theory and its Application.* Academic Press, New York.

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

Hall, P. and Patil, P., 1996. On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods. *J. of the Royal Statistical Society. Series B.* Methodological **58**, 361-377.

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

Jones, D.A., 1978. Nonlinear autoregressive processes. *Proc. Roy. Soc. London* A **360**, 71-95.

Kaplan, E.L. and Meier, P., 1958. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.

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

Li, L., MacGibbon, B. and Valenta, C., 2008. On the optimality of wavelet-based nonparametric regression with censored data. *J. of Appl. Prob. and Statist.* **3**, 243-261.

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

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

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

Ozaki, T., 1979. Nonlinear time series models for nonlinear random vibrations. *Technical report. Univ. of Manchester.*

Petrov, V.V., 1995. *Limit Theorems of Probability Theory.* Oxford Univ. Press Inc., New York.

Stute, W., 1993. Consistent estimation under random censorship when covariables are present. *J. Multiv. Analysis* **45**, 89-103.

Truong, Y.K. and Patil, P.N., 2001. Asymptotics for wavelet based estimates of piecewise smooth regression for stationary time series. *The Ann. of Instit. of Statist. Math.* **53**,159–178.

de Uña-Álvarez, J., Liang, H.Y. and Rodríguez-Casal, A., 2010. Nonlinear wavelet estimator of the regression function under left-truncated dependent data. *J. of Nonparametric Statist.*, **22**(3), 319-344.

Vidakovic, B., 1999. *Statistical Modeling by Wavelets.* John Wiley & Sons, Inc.