Wavelet Linear Density Estimation for a GARCH Model under Various Dependence Structures

Christophe Chesneau¹, Hassan Doosti²

¹Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse-Normandie, Campus II, Science 3, France.
²Department of Mathematics, Tarbiat Moallem University, Tehran, Iran.

Abstract. We consider \( n \) observations from the GARCH-type model: \( S = \sigma^2 Z \), where \( \sigma^2 \) and \( Z \) are independent random variables. We develop a new wavelet linear estimator of the unknown density of \( \sigma^2 \) under four different dependence structures: the strong mixing case, the \( \beta \)-mixing case, the pairwise positive quadrant case and the \( \rho \)-mixing case. Its asymptotic mean integrated squared error properties are explored. In each case, we prove that it attains a fast rate of convergence.

Keywords. Dependent sequence, GARCH models, linear estimator, rate of convergence, wavelets.


1 Introduction

The following GARCH-type model is considered: let \( (S_i)_{i \in \mathbb{Z}} \) be a strictly stationary random sequence such that

\[
S_i = \sigma^2_i Z_i, \quad i \in \mathbb{Z},
\]

(1)

\( (Z_i)_{i \in \mathbb{Z}} \) is a sequence of identically distributed random variables with common known density \( f_Z : [0, 1] \rightarrow (0, \infty) \) and \( (\sigma^2_i)_{i \in \mathbb{Z}} \) is a sequence of
identically distributed random variables with common unknown density $f_{\sigma^2} : [0, 1] \to (0, \infty)$. For any $i \in \mathbb{Z}$, $Z_i$ and $\sigma_i^2$ are independent. We aim to estimate $f_{\sigma^2}$ when only $n$ random variables $S_1, \ldots, S_n$ are observed under various dependent structures. Financial applications related to (1) can be found in [5].

In most of the papers, (1) is re-written via a logarithmic transformation: $\ln S_i = \ln \sigma_i^2 + \ln Z_i$, $i \in \mathbb{Z}$. Since we have a sum of two independent random variables, the density of $\ln \sigma_i^2$ becomes a convolution product. The classical scheme consists in deconvolving and estimating this density by using Fourier transform and nonparametric methods. For dependent sequences, see e.g. [21], [10] and [31]. However, the estimation of $f_{\sigma^2}$ is not “direct” in the following sense: if we denote $f_{\ln \sigma^2}$ the density of $\ln \sigma_1^2$, we have $f_{\sigma^2}(x) = (1/x)f_{\ln \sigma^2}(\ln x)$, $x \in (0, 1)$ and, for any estimator $\hat{f}$ of $f_{\ln \sigma^2}$, the associated mean integrated squared error (MISE) is

$$
\mathbb{E} \left( \int_{-\infty}^0 \left( \hat{f}(x) - f_{\ln \sigma^2}(x) \right)^2 \, dx \right) = \mathbb{E} \left( \int_0^1 \left( \frac{1}{x} \hat{f}(\ln x) - f_{\sigma^2}(x) \right)^2 \, x \, dx \right).
$$

Thus we obtain the MISE for $(1/x)\hat{f}(\ln x)$ of $f_{\sigma^2}$ but with respect to the measure $x \, dx$, not $dx$. For this reason, the global estimation of $f_{\sigma^2}$ on $[0, 1]$ from $\ln S_1, \ldots, \ln S_n$ via the standard MISE (with respect to $dx$) is not obvious. This point is also underlined in [10, 3.5]. Considering the exponential strong mixing case, the “direct” estimation of $f_{\sigma^2}$ has been recently investigated by [8].

In this paper, we complete this last study by estimating “directly” $f_{\sigma^2}$ for other realistic and standard dependence conditions as the polynomial strong mixing dependence (introduced by [26]), the $\beta$-mixing dependence (introduced by [33]), the pairwise positive quadrant dependence (PPQD) (introduced by [20]) and the $\rho$-mixing dependence (introduced by [18]). For results, examples and references on the standard density estimation problem under such dependence conditions, see e.g. [19], [4], [32], [30], [11], [21], [25], [6] and [17].

Combining the approaches of [8] and [17], we construct an estimator based on wavelet projections. We use wavelets because of their computational efficiency and asymptotic optimality properties. In particular, wavelet estimators enjoy interesting MISE for functions having possible complex singularities. We refer to [2] and [16] for a detailed coverage of wavelet theory in statistics. The asymptotic performance of our estimator is evaluated by determining an upper bound of the MISE over Besov balls. It is obtained as sharp as possible and coincides with the
one related to the standard \textit{i.i.d.} framework.

The organization of the paper is as follows. Assumptions on the model are presented in Section 2. Section 3 describes the wavelet basis and the Besov balls. Section 4 is devoted to our linear wavelet estimator and a general result. Applications are set in Section 5. Technical proofs are collected in Section 6.

2 Assumptions

Set $L^2([0,1]) = \left\{ h : [0,1] \rightarrow \mathbb{R}; \int_0^1 (h(x))^2 dx < \infty \right\}$. We assume that $f_{\sigma^2} \in L^2([0,1])$.

We suppose that there exists a positive integer $\nu$ such that, for any $i \in \{1, \ldots, n\}$,

$$Z_i = \prod_{r=1}^{\nu} U_{r,i},$$

(2)

where $U_{1,i}, \ldots, U_{\nu,i}$ are $\nu$ \textit{i.i.d.} random variables having the common uniform distribution on $[0,1]$. Assumption (2) excludes the Gaussian case but occurs in the study of some GARCH-type model as, for instance, the generalized multiplicative censoring model (see e.g. [1]).

It follows from (2) that

- the density of $Z_1$ is

$$f_{Z}(x) = \frac{1}{(\nu - 1)!} (-\ln x)^{\nu-1}, \quad x \in [0,1].$$

- the density of $S_1$ is

$$f_{S}(x) = \int_x^1 f_{Z}(\frac{x}{y}) f_{\sigma^2}(y) \frac{1}{y} dy$$

$$= \frac{1}{(\nu - 1)!} \sum_{u=0}^{\nu-1} \binom{\nu - 1}{u} (-\ln x)^{u} \int_x^1 (\ln y)^{\nu-1-u} f_{\sigma^2}(y) \frac{1}{y} dy, \quad x \in [0,1].$$

(3)
3 Wavelets and Besov Balls

For the purposes of this paper, we use the compactly supported wavelet bases on $[0, 1]$ described below.

Let $N$ be an integer such that $N > \nu + 1$ (where $\nu$ is the one in (2)), $\phi$ and $\psi$ be the initial wavelets of the Daubechies wavelets $\text{db}2N$. These functions have the particularity to be compactly supported and to belong to the class $C^\nu$.

Set

$$
\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k).
$$

With an appropriate treatment at the boundaries, there exists an integer $\tau$ satisfying $2^\tau \geq 2N$ such that the collection

$$
B = \{\phi_{\tau,k}(.), \ k \in \{0, \ldots, 2^\tau - 1\}; \psi_{j,k}(.); \ j \in \mathbb{N} - \{0, \ldots, \tau - 1\}, \ k \in \{0, \ldots, 2^j - 1\}\},
$$

is an orthonormal basis of $L^2([0, 1])$. We refer to [9].

For any integer $\ell \geq \tau$, any $h \in L^2([0, 1])$ can be expanded on $B$ as

$$
h(x) = \sum_{k=0}^{2^\ell - 1} \alpha_{\ell,k}\phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k=0}^{2^j - 1} \beta_{j,k}\psi_{j,k}(x), \quad x \in [0, 1],
$$

where $\alpha_{j,k}$ and $\beta_{j,k}$ are the wavelet coefficients of $h$ defined by

$$
\alpha_{j,k} = \int_0^1 h(x)\phi_{j,k}(x)dx, \quad \beta_{j,k} = \int_0^1 h(x)\psi_{j,k}(x)dx. \quad (4)
$$

As is traditional in the wavelet estimation literature, we shall investigate the performances of our estimator by assuming that $f_{\sigma^2}$ belongs to Besov balls. Their definitions in terms of wavelet coefficients are given below.

Let $s > 0$ and $M > 0$. A function $h$ belongs to $B^s_{2,\infty}(M)$ if and only if there exists a constant $M^* > 0$ (depending on $M$) such that the associated wavelet coefficients (4) satisfy

$$
\sup_{j \geq \tau} 2^{2js} \sum_{k=0}^{2^j - 1} \beta_{j,k}^2 \leq M^*.
$$

The Besov balls capture a wide variety of smoothness features in a function. Further details can be found in [22].
4 Estimator and result

4.1 Linear wavelet estimator

We follow the methodology of [8, Lemma 1]. For any positive integer $\ell$ and any $h \in C^\ell([0, 1])$, set

$$T(h)(x) = (xh(x))', \quad T_\ell(h)(x) = T(T_{\ell-1}(h))(x), \quad x \in [0, 1]. \quad (5)$$

The definition of this operator is such that, for any $h \in C^\nu([0, 1])$, we have

$$\int_0^1 f_{\sigma^2}(x)h(x)dx = \int_0^1 f_S(x)T_\nu(h)(x)dx. \quad (6)$$

The proof of (6) is based on the two following steps:

**Step 1.** For any positive integer $\ell$, set $G(h)(x) = -xh'(x)$ and $G_\ell(h)(x) = G(G_{\ell-1}(h))(x)$. It follows from (3), derivations and the binomial theorem that

$$f_{\sigma^2}(x) = -x(G_{\nu-1}(f_S)(x))' = G_\nu(f_S)(x), \quad x \in [0, 1]. \quad (7)$$

**Step 2.** By (7) and $\nu$ integrations by parts, we have

$$\int_0^1 f_{\sigma^2}(x)h(x)dx = \int_0^1 G_\nu(f_S)(x)h(x)dx$$

$$= \int_0^1 G_{\nu-1}(f_S)(x)T(h)(x)dx$$

$$= \ldots = \int_0^1 f_S(x)T_\nu(h)(x)dx.$$

Note that, in the simplest case $\nu = 1$, since $f_{\sigma^2}(x) = -xf'_S(x)$, $x \in [0, 1]$, the proof is reduced to

$$\int_0^1 f_{\sigma^2}(x)h(x)dx = -\int_0^1 f'_S(x)xh(x)dx = \int_0^1 f_S(x)(xh(x))'dx$$

$$= \int_0^1 f_S(x)T(h)(x)dx.$$

Using the method of moments, for any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$, we estimate the unknown wavelet coefficient $\alpha_{j,k} = \int_0^1 f_{\sigma^2}(x)\phi_{j,k}(x)dx$ by

$$\widehat{\alpha}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} T_\nu(\phi_{j,k})(S_i),$$
Assuming that $f_\sigma^2 \in B_{2,\infty}^s(M)$, we define the linear estimator $\hat{f}$ by

$$\hat{f}(x) = \sum_{k=0}^{2^{j_0} - 1} \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x), \quad x \in [0,1],$$  \hspace{1cm} (8)

where $j_0$ is an integer which will be chosen later (see Theorem 4.1 below).

Such an estimator is standard in nonparametric estimation via wavelets. See e.g. [16, Section 10]. For a survey on wavelet linear estimators in various density models, we refer to [7].

### 4.2 Result

**Theorem 4.1.** Consider (1) under the assumptions of Section 2. We suppose that

- there exist three constants, $\gamma \geq 2\nu$, $\delta \in [0,1)$ and $C > 0$, such that, for any integer $j \geq \tau$,

$$\sum_{k=0}^{2^j - 1} \sum_{m=1}^{n} |\mathbb{C}_{ov}(T_\nu(\phi_{j,k})(S_m), T_\nu(\phi_{j,k})(S_0))| \leq C 2^j (\gamma+1)^{n\delta},$$  \hspace{1cm} (9)

where $\mathbb{C}_{ov}(\ldots)$ denotes the covariance function and $T_\nu$ is (5).

- there exists a constant $C_* > 0$ such that

$$\sup_{x \in [0,1]} f_S(x) \leq C_*,$$  \hspace{1cm} (10)

Suppose that $f_\sigma^2 \in B_{2,\infty}^s(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^{j_0} = \left\lceil n^{1-\delta}/(2s+\gamma+1) \right\rceil$ (where $\lceil a \rceil$ denotes the integer part of $a$). Then there exists a constant $C > 0$ such that

$$\mathbb{E} \left( \int_0^1 \left( \hat{f}(x) - f_\sigma^2(x) \right)^2 \, dx \right) \leq C n^{-2s(1-\delta)/(2s+\gamma+1)}.$$

Naturally, the rate of convergence in Theorem 4.1 is obtained to be as sharp as possible.

The assumption (9) measures the dependence between $S_m$ and $S_0$. It is close to the $\Xi$-weak dependence introduced by [13, 14] but with the operator $T_\nu$. 
5 Applications

The four following subsections investigate separately the strong mixing case, the $\beta$-mixing case, the PPQD case and the $\rho$-dependence case which occur in a large variety of applications. We refer to [35], [27], [24] and [3].

5.1 Application to the strong mixing dependence

**Definition 5.1.** Let $(Y_i)_{i \in \mathbb{Z}}$ be a strictly stationary random sequence. For any $m \in \mathbb{Z}$, we define the $m$-th strong mixing coefficient of $(Y_i)_{i \in \mathbb{Z}}$ by

$$\alpha_m = \sup_{(A,B) \in \mathcal{F}_{-\infty,0} \times \mathcal{F}_{m,\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

where $\mathcal{F}_{-\infty,0}$ is the $\sigma$-algebra generated by $\ldots, Y_{-1}, Y_0$ and $\mathcal{F}_{m,\infty}$ is the $\sigma$-algebra generated by $Y_m, Y_{m+1}, \ldots$.

We say that $(Y_i)_{i \in \mathbb{Z}}$ is strong mixing if and only if $\lim_{m \to \infty} \alpha_m = 0$.

Applications on strong mixing can be found in e.g. [34], [12] and [5]. In the context of GARCH-type models as (1), see [15].

**Proposition 5.1.** Consider (1) under the assumptions of Section 2. Suppose that

- $(S_i)_{i \in \mathbb{Z}}$ is strong mixing,
- there exist three constants, $q \in (0, 1)$, $\delta \in [0, 1)$ and $C > 0$, such that
  $$\sum_{m=1}^{n} m^q \alpha_m^q \leq C n^\delta,$$
  (11)
- there exists a constant $C_s > 0$ such that
  $$\sup_{m \in \{1, \ldots, n\}} \sup_{(x,y) \in [0,1]^2} f_{(S_m,S_0)}(x,y) \leq C_s, \quad \sup_{x \in [0,1]} f_S(x) \leq C_s,$$
  (12)
- $f_{(S_m,S_0)}$ is the density of $(S_m, S_0)$ and $f_S$ is (3).

Suppose that $f_{\sigma^2} \in B^s_{2,\infty}(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^j = [n^{(1-\delta)/(2s+2\nu+1)}]$. Then there exists a constant $C > 0$ such that

$$\mathbb{E} \left( \int_0^1 \left( \hat{f}(x) - f_{\sigma^2}(x) \right)^2 dx \right) \leq C n^{-2s(1-\delta)/(2s+2\nu+1)}.$$
Note that (11) includes strong mixing coefficients with a polynomial rate of decay. For instance, if \( \alpha_m = 1/m^p \) with \( p > 1 \), then (11) holds with \( \delta = 0 \) (by taking \( q \in (1/(p-1), 1) \)).

The first inequality in (12) can be viewed as a “Castellana-Leadbetter”-type condition. It is standard in nonparametric estimation via dependent observations. Remark that, thanks to this assumption, Proposition 5.1 improves [8, Theorem 1]: the obtained rate of convergence is faster; the parameter \( q \) does not deteriorate the rate of convergence. However, this condition seems difficult to check in some situations. An alternative is explored in the next subsection.

5.2 Application to the \( \beta \)-mixing dependence

**Definition 5.2.** Let \((Y_i)_{i \in \mathbb{Z}}\) be a strictly stationary random sequence. For any \( m \in \mathbb{Z} \), we define the \( m \)-th \( \beta \)-mixing coefficient of \((Y_i)_{i \in \mathbb{Z}}\) by

\[
\beta_m = \frac{1}{2} \sup \sum_{i \in I} \sum_{j \in J} \left| \mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j) \right|,
\]

where the supremum is taken over all finite partitions \((A_i)_{i \in I}\) and \((B_j)_{j \in J}\) of \( \Omega \), which are respectively \( \mathcal{F}^Y_{-\infty,0} \) and \( \mathcal{F}^Y_{m,\infty} \) measurable, \( \mathcal{F}^Y_{-\infty,0} \) is the \( \sigma \)-algebra generated by \( \ldots, Y_{-1}, Y_0 \) and \( \mathcal{F}^Y_{m,\infty} \) is the one generated by \( Y_m, Y_{m+1}, \ldots \).

We say that \((Y_i)_{i \in \mathbb{Z}}\) is \( \beta \)-mixing if and only if \( \lim_{m \to \infty} \beta_m = 0 \).

Full details can be found in e.g. [12], [32], [5] and [3].

**Proposition 5.2.** Consider (1) under the assumptions of Section 2. Suppose that

- \((S_i)_{i \in \mathbb{Z}}\) is \( \beta \)-mixing,
- there exists a constant \( C > 0 \) such that

\[
\sum_{m=1}^{\infty} \beta_m \leq C,
\]

(13)
- there exists a constant \( C_\ast > 0 \) such that

\[
\sup_{x \in [0,1]} f_S(x) \leq C_\ast.
\]
Suppose that $f_\sigma^2 \in B_{2,\infty}^s(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^{j_0} = \lceil n^{1/(2s+2\nu+1)} \rceil$. Then there exists a constant $C > 0$ such that

$$
E \left( \int_0^1 \left( \hat{f}(x) - f_\sigma^2(x) \right)^2 \, dx \right) \leq C n^{-2s/(2s+2\nu+1)}.
$$

Since $\beta$-mixing implies strong mixing, Proposition 5.2 shows that the rate of convergence $n^{-2s/(2s+2\nu+1)}$ can be attained by $\hat{f}$ for strong mixing $(S_t)_{t \in \mathbb{Z}}$ without the constraint on $f(S_m, S_0)$ in (12).

### 5.3 Application to the PPQD

**Definition 5.3.** We say that $n$ random variables $S_1, \ldots, S_n$ are pairwise positive quadrant dependent (PPQD) if and only if, for any $(\ell, v) \in \{1, \ldots, n\}^2$ with $\ell \neq v$ and any $(x, y) \in [0, 1]^2$,

$$
\mathbb{P}(S_\ell > x, S_v > y) \geq \mathbb{P}(S_\ell > x)\mathbb{P}(S_v > y).
$$

Further details on PPQD can be found in [20], [23] and [28].

**Proposition 5.3.** Consider (1) under the assumptions of Section 2. Suppose that

- $S_1, \ldots, S_n$ are PPQD,
- there exist two constants, $\delta \in [0, 1)$ and $C > 0$, such that

$$
\sum_{m=1}^n m^3 \text{Cov}(S_m, S_0) \leq C n^\delta, \quad (14)
$$

- (12) is satisfied.

Suppose that $f_\sigma^2 \in B_{2,\infty}^s(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^{j_0} = \lceil n^{1-(1-\delta)/(2s+2\nu+1)} \rceil$. Then there exists a constant $C > 0$ such that

$$
E \left( \int_0^1 \left( \hat{f}(x) - f_\sigma^2(x) \right)^2 \, dx \right) \leq C n^{-(1-\delta)/(2s+2\nu+1)}.
$$

Proposition 5.4 below investigates the MISE properties of $\hat{f}$ in the PPQD case without the constraint on $f(S_m, S_0)$ in (12).

**Proposition 5.4.** Consider (1) under the assumptions of Section 2. Suppose that
• $S_1, \ldots, S_n$ are PPQD,

• there exist two constants, $\delta \in [0,1)$ and $C > 0$, such that

$$\sum_{m=1}^{n} \text{Cov}(S_m, S_0) \leq Cn^\delta,$$

(15)

• there exists a constant $C_* > 0$ such that

$$\sup_{x \in [0,1]} f_S(x) \leq C_*.$$

Suppose that $f_{\sigma^2} \in B_2^{s,\infty}(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^{j_0} = [n^{(1-\delta)/(2s+2\nu+4)}]$. Then there exists a constant $C > 0$ such that

$$\mathbb{E} \left( \int_{0}^{1} \left( \hat{f}(x) - f_{\sigma^2}(x) \right)^2 dx \right) \leq Cn^{-2s(1-\delta)/(2s+2\nu+4)}.$$

5.4 Application to the $\rho$-mixing dependence

Definition 5.4. Let $(Y_i)_{i \in \mathbb{Z}}$ be a strictly stationary random sequence. For any $m \in \mathbb{Z}$, we define the $m$-th maximal correlation coefficient of $(Y_i)_{i \in \mathbb{Z}}$ by

$$\rho_m = \sup_{(U,V) \in L^2(\mathcal{F}_{-\infty,0}^Y) \times L^2(\mathcal{F}_{m,\infty}^Y)} \frac{|\text{Cov}(U,V)|}{\sqrt{\mathbb{V}(U)\mathbb{V}(V)}},$$

where $\mathcal{F}_{-\infty,0}^Y$ is the $\sigma$-algebra generated by $\ldots, Y_{-1}, Y_0$, $\mathcal{F}_{m,\infty}^Y$ is the one generated by $Y_m, Y_{m+1}, \ldots$ and, for any $\mathcal{A} \in \{\mathcal{F}_{-\infty,0}^Y, \mathcal{F}_{m,\infty}^Y\}$, $L^2(\mathcal{A}) = \{U \in \mathcal{A}; \mathbb{E}(U^2) < \infty\}$.

We say that $(Y_i)_{i \in \mathbb{Z}}$ is $\rho$-mixing if and only if $\lim_{m \to -\infty} \rho_m = 0$.

For details on $\rho$-mixing, we refer to [18], [12], [29], [19] and [35].

Proposition 5.5. Consider (1) under the assumptions of Section 2. Suppose that

• $(S_i)_{i \in \mathbb{Z}}$ is $\rho$-mixing,

• there exist two constants, $\delta \in [0,1)$ and $C > 0$, such that

$$\sum_{m=1}^{n} \rho_m \leq Cn^\delta,$$

(16)
there exists a constant $C_*>0$ such that
\[ \sup_{x \in [0,1]} f_S(x) \leq C_*. \] (17)

Suppose that $f_{\sigma^2} \in B^{s,\infty}_2(M)$ with $s > 0$ and $M > 0$. Let $\hat{f}$ be (8) with $j_0$ such that $2^{j_0} = [n^{(1-\delta)/(2s+2\nu+1)}]$. Then there exists a constant $C > 0$ such that
\[ \mathbb{E} \left( \int_0^1 \left( \hat{f}(x) - f_{\sigma^2}(x) \right)^2 \, dx \right) \leq C n^{-(1-\delta)/(2s+2\nu+1)}. \]

**General Remark.** Note that, in Propositions 5.1, 5.3 and 5.5, if $\delta = 0$, the rate of convergence $n^{-2s/(2s+2\nu+1)}$ becomes the one attained by $\hat{f}$ when $S_1, \ldots, S_n$ are i.i.d. Therefore, our results extend the good asymptotic performances of $\hat{f}$ in the standard i.i.d. case to the considered dependence structures.

**Conclusions and Perspectives.** We have constructed a new wavelet estimator to estimate a density in a GARCH-type model under various dependence structures. Its asymptotic MISE properties have been investigated and fast rates of convergence have been established.

Due to its construction, $\hat{f}$ is not adaptive with respect to $s$, $\nu$ and $\delta$. Adaptivity can perhaps be achieved by using another wavelet estimator as the hard thresholding one. However, several important technical difficulties arise due to the dependence conditions and it is not immediately clear how to solve them. This aspect needs further investigations that we leave for a future work.

### 6 Proofs

In this section, we consider (1) under the assumptions of Section 2. Moreover, $C$ denotes any constant that does not depend on $j, k$ and $n$. Its value may change from one term to another and may depends on $\phi$.

**Proof of Theorem 4.1.** We expand the function $f_{\sigma^2}$ on $B$ as
\[ f_{\sigma^2}(x) = \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^\infty \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}(x), \quad x \in [0,1], \]

where $\alpha_{j_0,k} = \int_0^1 f_{\sigma^2}(x) \phi_{j_0,k}(x) \, dx$ and $\beta_{j,k} = \int_0^1 f_{\sigma^2}(x) \psi_{j,k}(x) \, dx$. 

We have, for any $x \in [0,1]$, 
\[
\hat{f}(x) - f_{\sigma^2}(x) = \sum_{k=0}^{2^{j_0} - 1} (\alpha_{j_0,k} - \tilde{\alpha}_{j_0,k}) \phi_{j_0,k}(x) - \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j - 1} \alpha_{j_0,k} \psi_{j,k}(x).
\]

Since $B$ is an orthonormal basis of $L^2([0,1])$, we have 
\[
E \left( \int_0^1 \left( \hat{f}(x) - f_{\sigma^2}(x) \right)^2 dx \right) = 
\sum_{k=0}^{2^{j_0} - 1} E \left( (\alpha_{j_0,k} - \tilde{\alpha}_{j_0,k})^2 \right) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j - 1} \beta_{j,k}^2.
\]

Let us now bound these two terms in turn.

Using (6), we have 
\[
\alpha_{j_0,k} = \int_0^1 f_{\sigma^2}(x) \phi_{j_0,k}(x) dx = \int_0^1 f_S(x) T_\nu(\phi_{j_0,k})(x) dx
\]
\[
= E(T_\nu(\phi_{j_0,k})(S_0)) = E(\tilde{\alpha}_{j_0,k}).
\]

So 
\[
E \left( (\tilde{\alpha}_{j_0,k} - \alpha_{j_0,k})^2 \right) = \mathbb{V}(\tilde{\alpha}_{j_0,k}) = \frac{1}{n^2} \mathbb{V} \left( \sum_{i=1}^{n} T_\nu(\phi_{j_0,k})(S_i) \right).
\]

We have 
\[
\mathbb{V} \left( \sum_{i=1}^{n} T_\nu(\phi_{j_0,k})(S_i) \right)
\]
\[
= \sum_{v=1}^{n} \sum_{\ell=1}^{n} \mathbb{C}_{av}(T_\nu(\phi_{j_0,k})(S_v), T_\nu(\phi_{j_0,k})(S_\ell))
\]
\[
= n \mathbb{V}(T_\nu(\phi_{j_0,k})(S_0)) + 2 \sum_{v=2}^{n} \sum_{\ell=1}^{v-1} \mathbb{C}_{av}(T_\nu(\phi_{j_0,k})(S_v), T_\nu(\phi_{j_0,k})(S_\ell))
\]
\[
\leq n \mathbb{V}(T_\nu(\phi_{j_0,k})(S_0)) + 2 \sum_{v=2}^{n} \sum_{\ell=1}^{v-1} \mathbb{C}_{av}(T_\nu(\phi_{j_0,k})(S_v), T_\nu(\phi_{j_0,k})(S_\ell)).
\]
The stationarity of \((S_i)_{i\in\mathbb{Z}}\) implies that
\[
\left| \sum_{v=2}^{n} \sum_{\ell=1}^{v-1} \text{Cov}(T_{\nu}(\phi_{j_0,k})(S_v), T_{\nu}(\phi_{j_0,k})(S_\ell)) \right| \\
= \left| \sum_{m=1}^{n} (n-m) \text{Cov}(T_{\nu}(\phi_{j_0,k})(S_m), T_{\nu}(\phi_{j_0,k})(S_0)) \right| \\
\leq n \sum_{m=1}^{n} |\text{Cov}(T_{\nu}(\phi_{j_0,k})(S_m), T_{\nu}(\phi_{j_0,k})(S_0))|.
\]

(21)

It follows from (19), (20) and (21) that
\[
\begin{align*}
\mathbb{E}\left( (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2 \right) \\
&\leq C \frac{1}{n} \left( \mathbb{V}(T_{\nu}(\phi_{j_0,k})(S_0)) + \sum_{m=1}^{n} |\text{Cov}(T_{\nu}(\phi_{j_0,k})(S_m), T_{\nu}(\phi_{j_0,k})(S_0))| \right).
\end{align*}
\]

Using [8, Proposition 1] i.e. thanks to (10), \(\mathbb{V}(T_{\nu}(\phi_{j_0,k})(S_0)) \leq C 2^{2j_0} \), (9) and \(\gamma \geq 2\nu\), we have
\[
\sum_{k=0}^{2^{j_0}-1} \mathbb{E}\left( (\hat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2 \right) \leq C \frac{1}{n} \left( 2^{j_0} 2^{2j_0} + 2^{j_0}(\gamma+1)n^\delta \right) \\
\leq C 2^{(\gamma+1)j_0} n^{\delta-1} \\
\leq C n^{-2s(1-\delta)/(2s+\gamma+1)}.
\]

(22)

Using \(f_\sigma^2 \in B_{2,\infty}^s(M)\), we obtain
\[
\sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C \sum_{j=j_0}^{\infty} 2^{-2js} \leq C 2^{-2j_0s} \leq C n^{-2s(1-\delta)/(2s+\gamma+1)}.
\]

(23)

It follows from (18), (22) and (23) that
\[
\mathbb{E} \left( \int_0^1 \left( \hat{f}(x) - f_\sigma^2(x) \right)^2 \, dx \right) \leq C n^{-2s(1-\delta)/(2s+\gamma+1)}.
\]

The proof of Theorem 4.1 is complete. \(\square\)
Proof of Proposition 5.1. Using a standard covariance equality and (12), for any \( g \in L^2([0,1]) \) and any \( m \in \{1, \ldots, n\} \), we have
\[
|\text{Cov}(g(S_m), g(S_0))| = \left| \int_0^1 \int_0^1 (f_{S_mS_0}(x,y) - f_S(x)f_S(y))g(x)g(y)\,dx\,dy \right|
\leq \int_0^1 \int_0^1 |f_{S_mS_0}(x,y) - f_S(x)f_S(y)||g(x)||g(y)|\,dx\,dy
\leq 2C_* \int_0^1 \int_0^1 |g(x)||g(y)|\,dx\,dy = 2C_* \left( \int_0^1 |g(x)|\,dx \right)^2. \tag{24}
\]
Moreover, using \((\phi_{j,k})^{(u)}(x) = 2^{(2\nu+1)j/2}\phi^{(u)}(2^j x - k)\) and doing the change of variables \( y = 2^j x - k \), we have
\[
\int_0^1 |T_\nu(\phi_{j,k})(x)|\,dx \leq \nu! \sum_{u=0}^{\nu} \int_0^1 |x^u(\phi_{j,k})^{(u)}(x)|\,dx
\leq \nu! \sum_{u=0}^{\nu} \int_0^1 |(\phi_{j,k})^{(u)}(x)|\,dx = C2^{-j/2} \nu! \sum_{u=0}^{\nu} 2^{uj} \int |\phi^{(u)}(y)|\,dy
\leq C2^{\nu j} 2^{-j/2}. \tag{25}
\]
It follows from (24) and (25) that
\[
|\text{Cov}(T_\nu(\phi_{j,k})(S_m), T_\nu(\phi_{j,k})(S_0))| \leq C2^{\nu j} 2^{-j}. \tag{26}
\]
By [8, equation before eq. (20)], we have
\[
|\text{Cov}(T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m))| \leq C2^{(2\nu+q)j} \alpha^q_m. 
\]
Therefore
\[
|\text{Cov}(T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m))| \leq C \min\left(2^{\nu j} 2^{-j}, 2^{(2\nu+q)j} \alpha^q_m\right). 
\]
Hence, by (11),
\[
\sum_{m=1}^n |\text{Cov}(T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m))| 
\leq C \sum_{m=1}^n \min\left(2^{\nu j} 2^{-j}, 2^{(2\nu+q)j} \alpha^q_m\right)
\leq C \left( \sum_{m=1}^{2^j-1} 2^{\nu j} 2^{-j} + \sum_{m=2^j}^n 2^{(2\nu+q)j} \alpha^q_m \right)
\leq C2^{\nu j} \left(1 + \sum_{m=1}^n m^q \alpha^q_m \right) \leq C2^{\nu j} \left(1 + n^\delta \right) \leq C2^{\nu j} n^\delta.
\]
Therefore

\[
\sum_{k=0}^{2^j-1} \sum_{m=1}^{n} |C_{ov} (T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m))| \leq C 2^{(2\nu+1)j} n^\delta.
\]

Proposition 5.1 follows from Theorem 4.1 with \( \gamma = 2\nu. \)

**Proof of Proposition 5.2.** Since \((S_i)_{i \in \mathbb{Z}}\) is \( \beta \)-mixing, for any bounded function \( g \), [32, equation line 12 p. 479 and Lemma 4.2. with \( p = 1 \)] imply that

\[
\sum_{m=0}^{n} |Cov(g(S_m), g(S_0))| \leq 2 \int_{0}^{1} b(x)g^2(x)f_S(x)dx,
\]

where \( b \) is a function such that, by (13), \( \int_{0}^{1} b(x)f_S(x)dx \leq \sum_{m=0}^{\infty} \beta_m \leq C. \)

We have \( (\phi_{j,k})^{(u)}(x) = 2^{(2u+1)j/2}\phi^{(u)}(2^j x - k). \) Since \( \phi \) is compactly supported, we have \( \sup_{x \in [0,1]} \sum_{k=0}^{2^j-1} (\phi^{(u)}(2^j x - k))^2 \leq C. \) Therefore, for any integer \( j \geq \tau, \)

\[
\sum_{k=0}^{2^j-1} (T_\nu(\phi_{j,k})(x))^2 \leq \nu (\nu!)^2 \sum_{k=0}^{2^j-1} \sum_{u=0}^{\nu} x^{2u} (\phi_{j,k})^{(u)}(x))^2
\]

\[
\leq \nu (\nu!)^2 \sum_{k=0}^{2^j-1} \sum_{u=0}^{\nu} ((\phi_{j,k})^{(u)}(x))^2
\]

\[
\leq \nu (\nu!)^2 2^{(2\nu+1)j/2} \sum_{u=0}^{\nu} \sum_{k=0}^{2^j-1} (\phi^{(u)}(2^j x - k))^2
\]

\[
\leq C 2^{(2\nu+1)j}.
\]

Putting (27) and (28) together, we obtain

\[
\sum_{k=0}^{2^j-1} \sum_{m=1}^{n} |C_{ov} (T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m))| \leq 2 \int_{0}^{1} b(x)f_S(x)dx \sum_{k=0}^{2^j-1} (T_\nu(\phi_{j,k}))^2(x)dx \leq C 2^{(2\nu+1)j} \int_{0}^{1} b(x)f_S(x)dx
\]

\[
\leq C 2^{(2\nu+1)j}.
\]

Proposition 5.2 follows from Theorem 4.1 with \( \gamma = 2\nu. \) \( \square \)
Proof of Proposition 5.3. Since $S_1, \ldots, S_n$ are PPQD, the Newman inequality (see [23]) yields: for any $m \in \{1, \ldots, n\}$ and any $g \in C^1([0,1])$,

$$|C_{ov}(g(S_m), g(S_0))| \leq \left( \sup_{x \in [0,1]} |g'(x)| \right)^2 C_{ov}(S_m, S_0). \quad (29)$$

Moreover, since $(\phi_{j,k})^{(u)}(x) = 2^{(2u+1)j/2} \phi^{(u)}(2^j x - k)$, we have

$$\sup_{x \in [0,1]} |T^i_{\nu}(\phi_{j,k})(x)| \leq (\nu + 1)! \sum_{n=0}^\nu |x^n(\phi_{j,k})^{(u)}(x)| \leq C \sum_{n=0}^{\nu+1} 2^{(2u+1)j/2} \sup_{x \in [0,1]} |\phi^{(u)}(x)| \leq C 2^{(2\nu+1)j/2} = C 2^{(2\nu+3)j/2}. \quad (30)$$

Therefore, by (26) (which holds thanks to (12)), (29) and (30), we have

$$|C_{ov}(T_{\nu}(\phi_{j,k})(S_m), T_{\nu}(\phi_{j,k})(S_0))| \leq C \min(2^{2\nu j} 2^{-j}, 2^{(2\nu+3)j} C_{ov}(S_m, S_0)).$$

Hence, by (14),

$$\sum_{m=1}^n |C_{ov}(T_{\nu}(\phi_{j,k})(S_0), T_{\nu}(\phi_{j,k})(S_m))| \leq C 2^{2\nu j} \left( 1 + \sum_{m=1}^n m^3 C_{ov}(S_m, S_0) \right) \leq C 2^{2\nu j} \left( 1 + n^6 \right) \leq C 2^{2\nu j} n^6.$$ 

Therefore

$$\sum_{k=0}^{2^j-1} \sum_{m=1}^n |C_{ov}(T_{\nu}(\phi_{j,k})(S_0), T_{\nu}(\phi_{j,k})(S_m))| \leq C 2^{(2\nu+1)j} n^6.$$ 

Proposition 5.3 follows from Theorem 4.1 with $\gamma = 2\nu$. □

Proof of Proposition 5.4. Proceeding similarly to the proof of Proposition 5.3, we have

$$|C_{ov}(T_{\nu}(\phi_{j,k})(S_m), T_{\nu}(\phi_{j,k})(S_0))| \leq C 2^{(2\nu+3)j} C_{ov}(S_m, S_0).$$
Therefore, by (16),
\[
\sum_{k=0}^{2^j-1} \sum_{m=1}^{n} |\text{Cov} \left( T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m) \right) | \leq C 2^{(2\nu+4)j} n^\delta.
\]

Proposition 5.4 follows from Theorem 4.1 with \( \gamma = 2\nu + 3 \).

\( \square \)

**Proof of Proposition 5.5.** A standard covariance inequality for \( \rho \)-mixing gives: for any \( m \in \{1, \ldots, n\} \) and any \( g \in L_2([0,1], f_S(x)dx) \),
\[
|\text{Cov} (g(S_m), g(S_0))| \leq \mathbb{E}((g(S_0))^2) \rho_m. \tag{31}
\]

See, for instance, [35, Lemma 1.2.7.].

Since \( S_0(\Omega) = [0,1] \), for any integer \( j \geq \tau \) and any \( k \in \{0, \ldots, 2^j-1\} \), we have
\[
\mathbb{E}((T_\nu(\phi_{j,k})(S_0))^2) \leq \nu(\nu!)^2 \sum_{u=0}^{\nu} \mathbb{E} \left( S_0^{2u} ((\phi_{j,k})^{(u)}(S_0))^2 \right)
\leq \nu(\nu!)^2 \sum_{u=0}^{\nu} \mathbb{E} \left( (\phi_{j,k})^{(u)}(S_0))^2 \right). \tag{32}
\]

Using (17), \( (\phi_{j,k})^{(u)}(x) = 2^{(2u+1)j/2} \phi^{(u)}(2^j x - k) \) and doing the change of variables \( y = 2^j x - k \), we obtain
\[
\mathbb{E} \left( (\phi_{j,k})^{(u)}(S_0))^2 \right) = \int_0^1 ((\phi_{j,k})^{(u)}(x))^2 f_S(x)dx
\leq C_s \int_0^1 ((\phi_{j,k})^{(u)}(x))^2 dx
\leq C_s 2^{2uj} \int (\phi^{(u)}(y))^2 dy. \tag{33}
\]

Putting (32) and (33) together, we obtain
\[
\mathbb{E}((T_\nu(\phi_{j,k})(S_0))^2) \leq C_s \nu(\nu!)^2 \sum_{u=0}^{\nu} 2^{2uj} \int_{1-N}^{N} (\phi^{(u)}(y))^2 dy \leq C 2^{2uj}. \tag{34}
\]

Hence, by (31), (34) and (16),
\[
\sum_{m=1}^{n} |\text{Cov} (T_\nu(\phi_{j,k})(S_0), T_\nu(\phi_{j,k})(S_m)) | \leq C 2^{2uj} \sum_{m=1}^{n} \rho_m \leq C 2^{2uj} n^\delta.
\]
Therefore
\[ \sum_{k=0}^{2^l-1} \sum_{m=1}^{n} |Cov(T_{\nu}(\phi_{j,k})(S_0), T_{\nu}(\phi_{j,k})(S_m))| \leq C 2^{(2\nu+1)j} n^\delta. \]

Proposition 5.5 follows from Theorem 4.1 with \( \gamma = 2\nu \). □

**Acknowledgments**

We thank the referee for insightful comments that helped us improve the paper significantly.

**References**


