Studentizing weighted sums of linear processes[1]

V. Dalla L. Giraitis H.L. Koul
University of Athens,
Queen Mary, University of London,
Michigan State University

Abstract

This paper presents a general method for studentizing weighted sums of a linear process where weights are arrays of known real numbers and innovations form a martingale difference sequence. Asymptotical normality for such sums was established in Abadir, Distaso, Giraitis and Koul (2013). This paper centers on the estimation of the standard deviation, to make the normal approximation operational. The proposed studentization is easy to apply and robust against unknown type of dependence (short range and long range) in the observations. It does not require the estimation of the parameters controlling the dependence structure. A finite sample Monte-Carlo simulation study shows the applicability of the proposed methodology for moderate sample sizes. Assumptions for studentization are satisfied by the Nadaraya-Watson kernel type weights used for inference in non-parametric regression settings.

Keywords: Linear process, short memory, long memory, weighted sum.

1 Introduction

A large part of the inference in statistics and econometrics is based on studentization of weighted sums

$$S_n = \sum_{j=1}^{n} b_{nj} X_j, \tag{1.1}$$

where $\{X_j\}$ is a covariance stationary process and b_{nj} are arrays of real numbers. The class of weights b_{nj} used in applications is rich and extends from equal weighting in the sample mean to strong down-weighting of a part of the sample in non-parametric regression estimation. The problem of studentizating S_n consists of the two parts:
(i) CLT, showing the asymptotic normality,

$$v_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1), \qquad v_{n,X}^2 := \mathrm{Var}(S_n) = \sum_{j,k=1}^{n} b_{nj} b_{nk} \mathrm{Cov}(X_j, X_k). \tag{1.2}$$

[1]Corresponding author: Liudas Giraitis Queen Mary, University of London; email: L.Giraitis@qmul.ac.uk;
Research of Giraitis and Koul is supported in part by the grants ESRC RES062230790 and NSF DMS-1205271. **April 7, 2013**

(ii) Constructing estimate $\widehat{v}_n^2$ of $\mathrm{Var}(S_n)$, such that

$$\frac{\widehat{v}_n^2}{v_{n,X}^2} \to_p 1, \qquad \widehat{v}_n^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1). \tag{1.3}$$

The main focuss of this paper is to build a computationally tractable estimate of the standard deviation $v_{n,X}$.

For i.i.d. zero mean random variables $X_1, \cdots, X_n$, the self-normalization $(\sum_{j=1}^n X_j)/V_n$, $V_n = \big(\sum_{j=1}^n X_j^2\big)^{1/2}$, allows establishing the standard central limit theorem without requiring any moment conditions and is well-investigated in the literature, see e.g., Shao (1997) and the references therein. Such self-normalization is commonly used in statistics. We show that it is valid also for the sums of martingale differences, see Proposition 2.2 below.

Extension of self-normalization to the weighted sum $S_n = \sum_{j=1}^n b_{nj} X_j$ of dependent random variables (time series), however, is not straightforward. It would be desirable to base studentization on the estimated standard deviation (SD) $\widehat{v}_{n,X}$, the positive square root of the estimated variance

$$\widehat{v}_{n,X}^2 := \sum_{j,k=1}^n b_{nj} b_{nk} \widehat{\gamma}_X(j-k), \qquad \widehat{\gamma}_X(k) = n^{-1} \sum_{j=1}^{n-|k|} (X_j - \bar{X})(X_{j+|k|} - \bar{X}), \tag{1.4}$$

where $\bar{X} = n^{-1} \sum_{j=1}^n X_j$ is the sample mean. Obviously, studentization by $\widehat{v}_{n,X}$ will not apply for all types of weights and requires additional restrictions on b_{nj}, besides those which are sufficient for the CLT (1.2). The simplest example when studentization by $\widehat{v}_{n,X}$ does not hold is that of the equally weighted sum $S_n = \sum_{j=1}^n X_j$. It may satisfy CLT $v_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1)$, but it cannot be standardized by $\widehat{v}_{n,X}$, because under equal weighting, $\widehat{v}_{n,X} \neq v_{n,X}(1 + o_p(1))$. In this case one first replaces $v_{n,X}$ by its asymptotic form $n^{1/2} s_X$, estimates the long run SD s_X by the HAC estimate $\widehat{s}_X$, and then uses $n^{1/2}\widehat{s}_X$ to studentize S_n, see e.g. Robinson (1998, 2005), Philips, Sun and Jin (2007) and Abadir, Distato and Giraitis (2009). One may also use bootstrap methodology, see e.g., Zhao and Li (2012).

In this paper we show that studentization by the standard deviation estimate $\hat{v}_{n,X}$ naturally applies for non-flat weights, as those used in non-parametric kernel regression, strongly downweighting a part of the sample. The asymptotic normality (1.2) for the sum $(S_n - ES_n)/v_{n,X}$ was established in Abadir, Distaso, Giraitis and Koul (2013). We provide simple sufficient conditions on b_{nj} under which $(S_n - ES_n)/\widehat{v}_{n,X}$ satisfies the CLT (1.3) when $\{X_j\}$ is a linear process with martingale difference (m.d.) innovations. Such studentization does not require the knowledge of the dependence structure of $\{X_j\}$, is easy to implement, and is applicable in non-parametric regression setting.

We start with the assumptions on the process $\{X_j\}$. We suppose that,

$$X_j = \mu + \sum_{k=0}^{\infty} a_k \zeta_{j-k} = \mu + \sum_{k=-\infty}^{j} a_{j-k} \zeta_k, \quad j \in \mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}, \quad \sum_{k=0}^{\infty} a_k^2 < \infty, \tag{1.5}$$

where $(\zeta_j, \mathcal{F}_j)$, $j \in \mathbb{Z}$ is a stationary ergodic martingale difference sequence (m.d.s.) with constant variance with respect to $\mathcal{F}_j := \sigma\text{-field}\{\zeta_i, i \le j\}$, i.e., $E(\zeta_j|\mathcal{F}_{j-1}) = 0$, $E\zeta_j^2 = \sigma_\zeta^2 < \infty$, $j \in \mathbb{Z}$. No other conditions on m.d.s. $\{\zeta_j\}$ are needed, which is an additional novelty of the results of this paper. The process $\{X_j\}$ of (1.5) is covariance stationary, with the mean $\mu = EX_j$.

In the sequel, $\gamma_X(k) := \mathrm{Cov}(X_j, X_{j+k})$, $k \in \mathbb{Z}$, $\Pi := [-\pi, \pi]$, and all limits are taken as $n \to \infty$, unless specified otherwise. We write $a_n \sim b_n$ if $a_n/b_n \to 1$, and $a_n \asymp b_n$ if $c_1 \le a_n/b_n \le c_2$, $\exists\, c_1 > 0, c_2 > 0$. Moreover, $\to_p$ and $\to_D$ denote convergence in probability and distribution, respectively, and $\mathcal{N}(0,1)$ stands for the standard normal distribution.

Section 2 contains the main results, section 3 presents studentization lemmas and section 4 includes the proofs. Performance of the theoretical results is illustrated by a small Monte-Carlo study in section 5. Examples 2.2 and 2.3 show applicability of the studentization results in non-parametric regression settings indicating that the suggested studentization method provides a tractable alternative to bootstrap methods for constructing confidence intervals in non-parametric regression, developed by several authors including Hall (1992) and McMurry and Politis (2008).

2 Main results

In this section we describe the asymptotic behavior of the sum S_n studentized by the estimate $\widehat{v}_{n,X}$ of (1.4) of the standard deviation $v_{n,X}$, under various types of dependence assumptions. Section 2.1 discusses the case of weak dependence (short memory), which is common in applications. Sections 2.2 and 2.3 extends these results to long memory processes, and heteroscedastic processes, respectively.

2.1 Weak dependence case

Here we focus on short memory (SM) processes. Accordingly, we assume $\{X_j\}$ in (1.5) is covariance stationary with its auto-covariance function γ_X satisfying

$$(2.1) \qquad \sum_{j\in\mathbb{Z}} |\gamma_X(j)| < \infty, \qquad s_X^2 := \sum_{j\in\mathbb{Z}} \gamma_X(j) > 0.$$

Under condition (2.1), the spectral density function $f(u) = (2\pi)^{-1} \sum_{j\in\mathbb{Z}} e^{\mathbf{i}ju}\gamma_X(j)$, $u \in \Pi$ is continuous and $f(0) = (2\pi)^{-1} s_X^2 > 0$.

The following theorem describes a class of weights b_{nj} for which studentization is possible, when $\{X_j\}$ has short memory. Let $B_n := \sum_{j=1}^n b_{nj}^2$.

Theorem 2.1 *Suppose that the process $\{X_j\}$ satisfies (1.5) and (2.1), and b_{nj} are such that*

$$|b_{n1}| + \sum_{k=2}^{n} |b_{nk} - b_{n,k-1}| = o\Big((\sum_{k=1}^{n} b_{nk}^2)^{1/2} \Big), \tag{2.2}$$

$$\sum_{k=1}^{n} |b_{nk}| = o\Big(n^{1/2} (\sum_{k=1}^{n} b_{nk}^2)^{1/2} \Big). \tag{2.3}$$

Then, with $v_{n,X}$ and $\widehat{v}_{n,X}$ as in (1.2) and (1.4),

$$v_{n,X}^2 \sim s_X^2 B_n, \tag{2.4}$$

$$v_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1), \tag{2.5}$$

$$\widehat{v}_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1), \qquad \widehat{v}_{n,X}/v_{n,X} \to_p 1. \tag{2.6}$$

Proof of Theorem 2.1. Under (1.5), (2.1) and (2.2), S_n satisfies assumption of Proposition 2.2(a) of Abadir *et al.* (2013), which implies the claims (2.4) and (2.5). In Lemma 3.1(i) below it is shown that under the additional condition (2.3), $\widehat{v}_{n,X}^2/v_{n,X}^2 \to_p 1$, which together with (2.5) proves (2.6). □

Condition (2.2) on the first differences of b_{nj}'s is needed to verify the asymptotic normality (2.5) and to derive (2.4). It is satisfied by "smooth", e.g. Nadaraya-Watson kernel weights, see Example 2.2. Additional condition (2.3) is needed to validate studentization (2.6).

The next theorem deals with studentization, when the variance $v_{n,X}^2$ may not satisfy (2.4). Compared to Theorem 2.1, it imposes weaker assumptions on b_{nj}'s and stronger assumptions on the spectral density f, by assuming that it is bounded away from zero. Then (2.2) can be replaced by a weaker condition (2.7) on the levels of b_{nj}'s. Remarkably, it still allows studentization, although it guaranties only $v_{n,X}^2 \asymp B_n$ and does not imply asymptotic of the variance (2.4).

Theorem 2.2 *Suppose $\{X_j\}$ of (1.5) satisfies the SM condition (2.1), $\inf_{u\in\Pi} f(u) > 0$, b_{nj}'s satisfy (2.3), and*

$$\max_{1\le j\le n} |b_{nj}| = o\Big(\big(\sum_{k=1}^{n} b_{nk}^2 \big)^{1/2} \Big). \tag{2.7}$$

Then the normal approximations (2.5) and (2.6) hold. Moreover,

$$2\pi \inf_{u\in\Pi} f(u)\, B_n \le v_{n,X}^2 \le 2\pi \sup_{u\in\Pi} f(u)\, B_n. \tag{2.8}$$

Proof of Theorem 2.2. Under (1.5), (2.1) and (2.7), Proposition 2.2(c) of Abadir *et al.* (2013) implies the asymptotic normality (2.5), which together with the consistency of the SD estimator $\widehat{v}_{n,X}$ shown in Lemma 3.1(i) below proves (2.6). To show (2.8), note that $v_{n,X}^2 = \int_\Pi |\sum_{j=1}^n e^{\mathbf{i}ju} b_{nj}|^2 f(u) du$. Then $v_{n,X}^2 \le \sup_{u\in\Pi} f(u) \int_\Pi |\sum_{j=1}^n e^{\mathbf{i}ju} b_{nj}|^2 du = \sup_{u\in\Pi} f(u) 2\pi \sum_{j=1}^n b_{nj}^2$, which implies the upper bound in (2.8). The lower bound follows using a similar argument. □

Remark 2.1 We shall now discuss some examples of the m.d. noise ζ_j and weights b_{nj} satisfying assumptions of Theorem 2.1. Standard examples of a stationary and ergodic m.d. sequence $\{\zeta_j\}$ appearing in (1.5) are ARCH or GARCH white noise processes ζ_j, see, e.g., Giraitis, Leipus and Surgailis (2007).

Example 2.1 (*Rolling window weights*). Assumptions (2.2) and (2.3) of Theorem 2.1 are satisfied by Nadaraya-Watson weights b_{nj} used in kernel estimation. The common feature of these weights is strong downweighting of data distant from a selected time point t. For example, consider the rolling window weights $b_{nj} = I(|j-t| \le n\omega)$ where $t = [\tau n]$ $(0 < \tau < 1)$, and $\omega \to 0$ is such that the window width $n\omega \to \infty$. Then

$$\frac{|b_{n1}| + \sum_{k=2}^{n} |b_{nk} - b_{n,k-1}|}{(\sum_{k=1}^{n} b_{nk}^2)^{1/2}} \le \frac{3}{(n\omega)^{1/2}} \to 0, \qquad \frac{\sum_{k=1}^{n} |b_{nk}|}{n^{1/2}(\sum_{k=1}^{n} b_{nk}^2)^{1/2}} \sim \frac{n\omega}{n\omega^{1/2}} = \omega^{1/2} \to 0.$$

Hence, these weights satisfy assumptions (2.2) and (2.3).

Remark 2.2 Assumptions on weights b_{nj} in Theorems 2.1 and 2.2 are independent from assumptions on the process X_j. As indicated in the proof, assumptions (2.2) in Theorem 2.1 and (2.7) in Theorem 2.2 are sufficient for asymptotic normality (2.5). Additional restriction (2.3) on b_{nj}'s is needed to secure consistency of the standard deviation estimator $\widehat{v}_{n,X}$. It is minimal in the sense that it is not satisfied by the equal weights where $\widehat{v}_{n,X}$ is not consistent. The requirements of short memory (2.1) and linearity (1.5) on X_j's are unrestrictive and satisfied, e.g., by stationary ARMA models generated by a m.d. noise ζ_j. Notice that under m.d. noise ζ_j, the auto-covariance function $\gamma_X(k) = \sigma_\zeta^2 \sum_{j=0}^{\infty} a_j a_{k+j}$, $k \ge 0$, depends only on the variance σ_ζ^2 of the noise ζ_j, and is the same as for an i.i.d. noise $\zeta_j \sim IID(0, \sigma_\zeta^2)$.

Remark 2.3 (*Zero mean case*, $EX_j = 0$). In Theorems 2.1 and 2.2, the variance $v_{n,X}^2$ is estimated by $\widehat{v}_{n,X}^2$, and $\widehat{v}_{n,X}^2 = v_{n,X}^2(1 + o_p(1))$. In sample auto-covariances $\widehat{\gamma}_X(k)$, appearing in $\widehat{v}_{n,X}^2$, X_j's are centered by the sample mean $\bar{X}$. If, in addition, the mean $EX_j = 0$, e.g., when X_j's are residuals, $\widehat{v}_{n,X}^2$ in studentization can be replaced by

$$\widetilde{v}_{n,X}^2 := \sum_{j,k=1}^{n} b_{nj} b_{nk} \widetilde{\gamma}_X(j-k), \qquad \widetilde{\gamma}_X(k) = n^{-1} \sum_{j=1}^{n-|k|} X_j X_{j+|k|}. \tag{2.9}$$

Then, by Lemma 3.1(i), $\widetilde{v}_{n,X}$ is a consistent estimate of $v_{n,X}$, and under the assumptions of Theorems 2.1 and 2.2, besides (2.6), we also have

$$\widetilde{v}_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0,1), \qquad \widetilde{v}_{n,X}^2 / v_{n,X}^2 \to_p 1.$$

2.1.1 HAC estimator

In applications, the variance $v_{n,X}^2 = \mathrm{Var}(\sum_{j=1}^n X_j)$ of a sum of equally weighted observations $X_1, \cdots, X_n$ of a short memory linear process $\{X_j\}$ (1.5) is commonly estimated by the HAC estimator

$$\widehat{v}_{HAC,q}^2 := n\widehat{s}_{n,X}^2, \tag{2.10}$$

where

$$\widehat{s}_{n,X}^2 := q^{-1}\widehat{v}_{q,X}^2, \quad \widehat{v}_{q,X}^2 = q\widehat{\gamma}_X(0) + 2\sum_{k=1}^{q}(q-k)\widehat{\gamma}_X(k), \quad q = o(n), \tag{2.11}$$

is the estimate of the long-run variance s_X^2 of (2.1), and $q = q_n \to \infty$ is a bandwidth parameter. The weights $b_{nj} = 1$ do not satisfy assumption (2.3) and hence studentization of the sum $\sum_{j=1}^n X_j$ by the standard deviation estimate $\widehat{v}_{n,X}$ is not applicable.

Consistency of the estimates $\widehat{v}_{HAC,q}^2$ and $\widehat{s}_{n,X}^2$ is well known when X_j's have i.i.d. innovations. Proposition 2.1 extends this result to short memory processes with m.d. innovations.

Proposition 2.1 *Suppose that a short memory linear process $\{X_j\}$, (1.5), with stationary ergodic m.d. innovations $\{\zeta_j\}$ satisfies (2.1). Then*

$$\widehat{v}_{HAC,q}^2 = v_{n,X}^2(1 + o_p(1)), \qquad \widehat{s}_X^2 \to_p s_X^2. \tag{2.12}$$

Proof. Let $S_q := \sum_{j=1}^q b_{nj}X_j$, $b_{nj} := I(1 \le j \le q)$. Under assumption $q \to \infty$, $q = o(n)$, the weights b_{nj} satisfy assumptions (2.2) and (2.3) of Theorem 2.1. Hence, by (2.6), $\widehat{v}_{q,X}^2 = v_{q,X}^2(1 + o_p(1))$, where $v_{q,X}^2 = \mathrm{Var}(S_q^2)$. Since under short memory $v_{n,X}^2 \sim s_X^2 n$ and $v_{q,X}^2 \sim s_X^2 q$, then

$$\begin{aligned}
\widehat{v}_{HAC,q}^2 &= nq^{-1}v_{q,X}^2(1 + o_p(1)) = s_X^2 n(1 + o_p(1)) = v_{n,X}^2(1 + o_p(1)), \\
\widehat{s}_{n,X}^2 &= q^{-1}v_{q,X}^2(1 + o_p(1)) \to_p s_X^2,
\end{aligned}$$

which completes the proof of the proposition. □.

2.2 Robustness to dependence

In this section we show that under slightly stronger assumptions on the weights b_{nj}, the results of asymptotic normality for studentized weighted sums obtained for a short memory processes remain valid also for long memory processes. Thus, studentization is robust to the unknown dependence structure of $\{X_j\}$.

We assume that a linear process $\{X_j\}$ of (1.5) may have either short memory (SM) or long memory (LM). In SM case it satisfies (2.1), while in LM case there exist $0 < d < 1/2$,

$c_\gamma > 0$ and $c_f > 0$, with $\gamma_X(k)$ and the spectral density f satisfy standard long memory asymptotics:

$$\gamma_X(j) \sim c_\gamma j^{-1+2d}, \quad j \to \infty, \tag{2.13}$$

$$f(u) \sim c_f |u|^{-2d}, \quad u \to 0, \tag{2.14}$$

and $f(u)$ is bounded on intervals $[\delta, \pi]$, for any $\delta > 0$.

These LM assumptions are not restrictive and are satisfied by stationary ARFIMA(p,d,q) models. Relations (2.13) and (2.14) hold simultaneously, e.g. if $a_k = c_a k^{-1+d}(1 + O(k^{-1}))$, $k \to \infty$ in (1.5), see Giraitis, Koul and Surgailis (2012, Propositions 3.2.1, 3.2.2).

The following theorem validates normal approximation for studentized weighted sums of short and long memory processes $\{X_j\}$ when studentized by $\widehat{v}_{n,X}$. Such studentization does not require to know the type of dependence. It is easier to implement than the standardizations based on the estimated asymptotic form of $v_{n,X}$ as is done, for example, in Robinson (1997) and Guo and Koul (2007) when dealing with non-parametric regression models. The asymptotic form of $v_{n,X}$ depends on covariance structure of $\{X_j\}$, in particular on the memory parameter d, which makes its estimation relatively a difficult task.

Theorem 2.3 *Suppose that b_{nj}'s in S_n satisfy*

$$\sum_{k=1}^n b_{nk}^2 \to \infty, \qquad \sum_{k=1}^n b_{nk}^2 = o(n), \qquad \sum_{k=1}^n |b_{nk}| = O\Big(\sum_{k=1}^n b_{nk}^2\Big), \tag{2.15}$$

$$|b_{n1}| + \sum_{k=2}^n |b_{nk} - b_{n,k-1}| = O(1). \tag{2.16}$$

Let $\{X_j\}$ be a linear process (1.5). Then (2.5) and the studentized normal approximation (2.6) hold under assumption (2.1) of SM, and under assumptions (2.13) and (2.14) of LM.

Moreover, under SM, $v_{n,X}^2 \asymp B_n$, while under LM, $v_{n,X}^2 \asymp B_n^{1+2d}$.

Proof. *SM case.* Notice that (2.15) and (2.16) imply (2.2) and (2.3). Thus, Theorem 2.1 implies (2.5) and (2.6), while by Lemma 4.3(i), $v_{n,X}^2 \asymp B_n$.

LM case. Under assumptions (2.13) and (2.14) of LM, Lemma 4.3(ii) implies $v_{n,X}^2 \asymp B_n^{1+2d}$, which together with (2.15) and (2.16) implies $|b_{n1}| + \sum_{k=2}^n |b_{nk} - b_{n,k-1}| = o(v_{n,X})$. Hence S_n satisfies assumptions of Theorem 2.2 (iii) of Abadir *et al.* (2013), which yields asymptotic normality (2.5). By Lemma 3.1(ii) below, $v_{n,X} = \widehat{v}_{n,X}(1 + o_p(1))$, which together with (2.5) proves (2.6). □

Remark 2.4 While Theorem 2.1 establishes the normal approximation for studentized weighted sums of SM processes $\{X_j\}$, Theorem 2.3 extends it to SM and LM processes covering much wider range of dependence scenarios. Verification of its assumptions (2.15) and (2.16) on weights b_{nj} in specific cases is straightforward. They also imply assumptions

(2.2), (2.3) and (2.7) on b_{nj} of Theorems 2.1 and 2.2 and do not depend on $\{X_j\}$. Structural assumptions on $\{X_j\}$ in Theorem 2.3 allow for extremely rich dependence structure, are unrestrictive and practically do not require verification. They are satisfied by standard time series models, e.g. ARMA and ARFIMA.

Example 2.2 (*Non-parametric regression weights*). Nadaraya-Watson nonparametric regression estimation uses kernel type weights $b_{nj} = K(|j - nx|/nw)$, $j = 1, \cdots, n$, $x \in (0, 1)$. Here, $w \equiv w_n \to 0$, $nw \to \infty$, is a sequence of bandwidths and $K \geq 0$ is a density function on $(-1, 1)$, vanishing off $(-1, 1)$, and having bounded first derivative. It is straightforward to verify that these weights satisfy assumptions (2.15) and (2.16) of Theorem 2.3. Indeed,

$$\sum_{k=1}^{n} b_{nk} \sim nw, \quad \sum_{k=1}^{n} b_{nk}^2 \sim cnw, \quad |b_{n1}| + \sum_{k=2}^{n} |b_{nk} - b_{n,k-1}| = O(1).$$

To establish the first claim, note that $\sum_{k=1}^{n} b_{nk} \sim \sum_{k=1-[nx]}^{n-[nx]} K(|k|/n\omega) \sim (n\omega) \int K(u)du = n\omega$. Similarly, the second claim follows with $c = \int K^2(u)du$. Next, by the mean value theorem and since $K(|u|) = 0$ for $|u| \geq 1$, it holds $|b_{nk} - b_{n,k-1}| \leq C(n\omega)^{-1} I(|k - nx|/n\omega \leq 2)$, which implies $|b_{n1}| + \sum_{k=2}^{n} |b_{nk} - b_{n,k-1}| \leq C\big(1 + \sum_{k=2}^{n} (n\omega)^{-1} I(|k - nx|/n\omega \leq 2)\big) \leq C$, which proves the third claim. These weights clearly also satisfy assumptions (2.2) and (2.3) of Theorem 2.1, and (2.7) of Theorem 2.2. This example shows that the studentization results of this paper apply in non-parametric regression settings.

Condition $w \to 0$ is essential. For example, if $w = 1$, then the weights $b'_{nj} = K(|j|/n)$, $j = 1, \cdots, n$ have property $\sum_{k=1}^{n} b'_{nk} \sim cn$ and $\sum_{k=1}^{n} b'^2_{nk} \sim cn$, and, hence, they do not satisfy assumptions (2.15) and (2.3). For such weights studentization results of this paper do not apply.

Remark 2.5 If X_j's have zero mean, $EX_j = 0$, then Theorem 2.3 holds with $\widehat{v}_{n,X}$ replaced by $\widetilde{v}_{n,X}$ of (2.9):

$$\widetilde{v}_{n,X}^{-1}(S_n - ES_n) \to_D \mathcal{N}(0, 1), \qquad \widetilde{v}_{n,X} = v_{n,X}(1 + o_p(1)).$$

The second claim is proved in Lemma 3.1 while the first follows from (2.5) and the second claim.

2.2.1 Self-normalizing sums of martingale differences

The self-normalization property of the sum $\sum_{j=1}^{n} \varepsilon_j$ is of interest in residual analysis. It is often assumed that residuals ε_j form a martingale difference sequence with unknown possibly non-constant variances. Asymptotic theory for sums of m.d.s. is well developed, see, e.g., Hall and Heyde (1980). The following proposition establishes the self-normalization property for sums of martingale differences with non-constant variances. Differently from weighted sums discussed above, the weights d_{nj} below are unknown.

Proposition 2.2 *Let $\varepsilon_j = d_{nj}\zeta_j$, $j \geq 1$, where $\{\zeta_j\}$ is a stationary ergodic m.d.s., $E\zeta_1^2 = \sigma_\zeta^2 < \infty$, and d_{nj}'s are real numbers such that*

$$|d_{n1}| + \sum_{j=2}^n |d_{nj} - d_{n,j-1}| = o\big((\sum_{j=1}^n d_{nj}^2)^{1/2}\big). \tag{2.17}$$

Then

$$\frac{\sum_{j=1}^n \varepsilon_j}{(\sum_{k=1}^n \varepsilon_k^2)^{1/2}} \to_D \mathcal{N}(0,1). \tag{2.18}$$

Proof. Let $D_n := \sum_{j=1}^n d_{nj}^2$, $p_{nj} := D_n^{-1/2} d_{nj}$, $S_{n,\varepsilon} := \sum_{j=1}^n p_{nj}\zeta_j$ and $V_n^2 = \sum_{j=1}^n p_{nj}^2 \zeta_j^2$. Note that $EV_n^2 = \sigma_\zeta^2$. Then the l.h.s. of (2.18) can be written as $S_{n,\varepsilon}/V_n$. Under (2.17), the weights p_{nk} satisfy assumption (2.2) of Theorem 2.1, and, therefore, (2.5) implies $S_{n,\varepsilon} \to_D \mathcal{N}(0, \sigma_\zeta^2)$. To prove (2.18), it remains to show that $V_n^2 \to_p \sigma_\zeta^2$. Note that $\sum_{k=1}^n p_{nk}^2 = 1$, while $|p_{nk}| \leq |p_{n1}| + \sum_{j=2}^k |p_{nj} - p_{n,j-1}|$ together with (2.17) implies $\max_{k=1,\cdots,n} |p_{nk}| = o(1)$. So, $p_{n1}^2 + \sum_{k=2}^n |p_{nk}^2 - p_{n,k-1}^2| \leq o(1)\big(|p_{n1}| + \sum_{k=2}^n |p_{nk} - p_{n,k-1}|\big) = o(1)$, by (2.17). Hence, the weights $c_{nj} := p_{nj}^2$ satisfy assumptions of Lemma 4.7(i) below with $\alpha_n = 1$, which implies $V_n^2 \to_p \sigma_\zeta^2$ and completes the proof. $\square$

2.3 Heteroscedastic case

A number of application require extending studentized normal approximation for the weighted sum of heteroscedastic observations

$$S_{n,Y} = \sum_{j=1}^n b_{nj} Y_j, \qquad Y_j := h_{nj} X_j, \qquad j = 1, \cdots, n, \tag{2.19}$$

where $\{X_j\}$ is a covariance stationary process (1.5), $EX_1 = 0$, $EX_1^2 = 1$, and h_{nj} are unknown real numbers. Such variables Y_j have zero mean and non-constant variance $\mathrm{Var}(Y_j) = h_{nj}^2$. Technically, the sum $S_{n,Y} = \sum_{j=1}^n (b_{nj} h_{nj}) X_j$ is a weighted sum of a homoscedastic process X_j, and asymptotic normality $S_{n,Y}/(\mathrm{Var}(S_{n,Y}))^{1/2} \to_D \mathcal{N}(0,1)$ holds as long as the weights $b_{nj} h_{nj}$ satisfy conditions established in the previous sections. Studentization, however, is complicated by the fact that now the weights $(b_{nj} h_{nj})$ are unknown.

We shall now address the question of estimation of the standard deviation $v_{n,Y} := (\mathrm{Var}(S_{n,Y}))^{1/2}$. Since Y_j's have zero mean we will use the estimate $\widetilde{v}_{n,Y}^2 := \sum_{j,k=1}^n b_{nj} b_{nk} \widetilde{\gamma}_Y(j-k)$, where the sample autocovariance function $\widetilde{\gamma}_Y(k)$ is defined as in (2.9), see Remark 2.3.

Under additional smoothness assumptions on h_{nj}, described in Assumption 2.1, the estimate $\widetilde{v}_{n,Y}^2$ asymptotically factors as $\sigma_h^2 v_{n,X}^2$, with σ_h^2 as in (2.20). If, in addition, $v_{n,Y}^2 \sim \theta^2 v_{n,X}^2$, for some $\theta > 0$, then studentization can be carried out. In non-parametric regression settings, the latter as the rule holds and θ can be consistently estimated, see Example 2.3.

We start with assumptions on h_{nj}. Define $\sigma_h^2 := n^{-1} \sum_{j=1}^n h_{nj}^2$.

Assumption 2.1 *Y_j, $j = 1, \cdots, n$ is as in (2.19), and*

$$|h_{n1}| + \sum_{j=2}^{n} |h_{n,j} - h_{n,j-1}| = O(1), \quad \sigma_h^2 \geq c > 0, \quad n \to \infty, \qquad \exists\, c > 0. \tag{2.20}$$

To establish $v_{n,Y}^2 \sim \theta^2 v_{n,X}^2$, let $B_{nh} := \sum_{j=1}^n b_{nj}^2 h_{nj}^2$, $B_n = \sum_{j=1}^n b_{nj}^2$, and additionally assume that

$$\sum_{j=1}^{n} b_{nj}^2 (h_{nj} - \theta)^2 = o\big(B_n\big), \quad \exists\, \theta \neq 0. \tag{2.21}$$

Theorem 2.3 deals with weighted sums $S_{n,X}$ of SM/LM processes X_j with constant variance. The next theorem extends it to sums $S_{n,Y}$ of heteroscadastic observations $Y_j = h_{nj} X_j$ as in (2.19), under the same conditions on b_{nj} and X_j.

Theorem 2.4 *Let Y_j be as in (2.19) satisfying Assumption 2.1 and let $S_{n,Y} = \sum_{j=1}^n b_{nj} Y_j$. Suppose, also, that b_{nj}'s and X_j's satisfy assumptions of Theorem 2.3.*
(a) *If $B_{nh} \geq c B_n$, for some $c > 0$, then $B_{nh} \asymp B_n$, and*

$$\frac{S_{n,Y}}{v_{n,Y}} \ \to_D \ \mathcal{N}(0,1), \quad \widetilde{v}_{n,Y} = \sigma_h v_{n,X} (1 + o_p(1)). \tag{2.22}$$

(b) *If (2.21) holds, then $B_{nh} = \theta^2 B_n (1 + o(1))$, and*

$$v_{n,Y} \sim \theta v_{n,X}, \qquad \widetilde{v}_{n,Y} = \sigma_h v_{n,X} (1 + o_p(1)), \quad \widetilde{\gamma}_Y(0) = \sigma_h^2 (1 + o_p(1)), \tag{2.23}$$

$$\frac{1}{\theta} \Big(\frac{\widetilde{\gamma}_Y(0)}{\widetilde{v}_{n,Y}^2} \Big)^{1/2} S_{n,Y} \to_D \mathcal{N}(0,1). \tag{2.24}$$

Proof. (a) Write $S_{n,Y} = \sum_{j=1}^n \bar{b}_{nj} X_j$, $\bar{b}_{nj} := b_{nj} h_{nj}$. By Lemma 4.6(ii), $\bar{b}_{nj}$ satisfy assumptions of Theorem 2.3. Thus, Theorem 2.3 implies $B_{nh} \asymp B_n$ and the first claim of (2.22). The second claim is established in Lemma 3.2(i) below.

(b) To verify $B_{nh} = \theta^2 B_n + o(B_n)$, use $b_{nj}^2 - \theta^2 = (b_{nj} - \theta)^2 + 2\theta (b_{nj} - \theta)$, the Cauchy–Schwarz inequality and (2.21), to obtain

$$\begin{aligned} |B_{nh} - \theta^2 B_n| &= \ |\sum_{j=1}^n b_{nj}^2 (h_{nj}^2 - \theta^2)| \leq \sum_{j=1}^n b_{nj}^2 (h_{nj} - \theta)^2 + 2|\theta| \sum_{j=1}^n b_{nj}^2 |h_{nj} - \theta| \\ &\leq \ o(B_n) + 2|\theta| \big(\sum_{j=1}^n b_{nj}^2 (h_{nj} - \theta)^2 \big)^{1/2} \big(\sum_{j=1}^n b_{nj}^2 \big)^{1/2} = o(B_n). \end{aligned}$$

Hence, condition of (a) holds implying (2.22). It suffices to verify (2.23), which together with (2.22) proves (2.24). The second and third claims in (2.23) are shown in Lemma 3.2(i) below. We prove the first claim.

Let $r_n = S_{n,Y} - \theta S_{n,X} = \sum_{j=1}^n b_{nj}(h_{nj} - \theta)X_j$. Then, by Lemma 4.2,

$$|v_{n,Y}^2 - \theta^2 v_{n,X}^2| \le Er_n^2 + 2E|r_n||\theta S_{n,X}| \le Er_n^2 + 2|\theta|(Er_n^2)^{1/2} v_{n,X}. \tag{2.25}$$

In SM case, parts (i) and (iii) of Lemma 4.3 imply $v_{n,X}^2 \asymp B_n$, while by (2.21), $Er_n^2 \le C\sum_{j=1}^n b_{nj}^2(h_{nj} - \theta)^2 = o(B_n)$, which proves (2.23). In LM case, by Lemma 4.3(ii), $v_{n,X}^2 \asymp B_n^{1+2d}$ and $Er_n^2 \le C\{\sum_{j=1}^n b_{nj}^2(h_{nj} - \theta)^2\}^{1+2d} = o(B_n^{1+2d})$, which together with (2.25) proves (2.23). □

Observe that conditions on b_{nj} and h_{nj} enabling studentization (2.24) are independent of the covariance structure of $\{X_j\}$, and the same under short and long memory.

Corollary 2.1 *Theorem 2.4 (a)-(b) holds if, instead of assumptions of Theorem 2.3, b_{nj} and X_j satisfy assumptions of Theorem 2.1.*

Proof. Similarly as in the proof of Theorem 2.4, by Lemma 4.6(i), $\bar{b}_{nj}$'s satisfy assumptions of Theorem 2.1, which implies the first claim of (2.22), while the second claim follows by Lemma 3.2(ii).

Part (b) follows by the same argument as in Theorem 2.4. □

Example 2.3 Non-parametric regression estimation requires establishing asymptotic normality for studentized weighted sums $S_{n,Y} = \sum_{j=1}^n b_{nj}Y_j$ of heteroscedastic observations $Y_j = h_{nj}X_j$, where $h_{nj} := \sigma(j/n)$ and $\sigma(x) \ge 0$, $x \in [0,1]$, is a continuous function with the bounded first derivative, while the weights $b_{nj} = K(|j - nx|/n\omega)$ are as in Example 2.2. Using the mean value theorem, it is straightforward to verify (2.20) for h_{nj}'s. Moreover, they satisfy (2.21) with $\theta = \sigma(x)$:

$$\sum_{j=1}^n b_{nj}^2(h_{nj} - \sigma(x))^2 = o(1)\sum_{j=1}^n b_{nj}^2.$$

Indeed, by continuity of σ, $\forall\varepsilon > 0$ one can find $\delta > 0$ such for $|h_{nj} - \sigma(x)| = |\sigma(j/n) - \sigma(x)| \le \varepsilon$ as $|j/n - x| \le \delta$, while $b_{nj} = K(|j - nx|/n\omega) = 0$ for $|j/n - x| \ge \varepsilon$ as $\omega \to 0$ because K has bounded support. In Example 2.2 we showed that b_{nj}'s satisfy (2.15) and (2.16). Hence, if $\{X_j\}$ has property (2.1) of SM, or properties (2.13) and (2.14) of LM, then by Theorem 2.4,

$$\frac{1}{\sigma(x)}\Big(\frac{\widetilde{\gamma}_Y(0)}{\widetilde{v}_{n,Y}^2}\Big)^{1/2} S_{n,Y} \to_D \mathcal{N}(0,1).$$

This anables studentization, since in non-parametric regression settings $\sigma(x)$ can be often consistently estimated, see, e.g., Guo and Koul (2007).

3 Consistency Lemmas

This section consists of the three lemmas that prove the consistency of $\widehat{v}_{n,X}$ for $v_{n,X}$ for homoscedastic, heteroscedastic and signal plus noise type observations. These lemmas are crucial for the studentization as is seen in Sections 2.1 and 2.2. The process $\{X_j\}$ is assumed to have SM in part (i) and SM or LM in part (ii).

Lemma 3.1 *Let $v^2_{n,X}$, $\widehat{v}^2_{n,X}$, and $\widetilde{v}^2_{n,X}$ be as in (1.2), (1.4), and (2.9), respectively.*
(i) *Suppose that assumptions of Theorem 2.1 or 2.2 hold. Then*

$$
\begin{aligned}
\text{(3.1)} \qquad \widehat{v}^2_{n,X}/v^2_{n,X} &\to_p 1, \\
\widetilde{v}^2_{n,X}/v^2_{n,X} &\to_p 1, \quad \text{if } EX_j = 0.
\end{aligned}
$$

(ii) *Suppose that assumptions of Theorem 2.3 are satisfied. Then also (3.1) holds.*

Proof. Without loss of generality, assume that $EX_j = 0$. Since $|\widetilde{v}^2_{n,X} - v^2_{n,X}| \le |E\widetilde{v}^2_{n,X} - v^2_{n,X}| + |\widetilde{v}^2_{n,X} - E\widetilde{v}^2_{n,X}|$ and $|\widehat{v}^2_{n,X} - v^2_{n,X}| \le |E\widetilde{v}^2_{n,X} - v^2_{n,X}| + |\widetilde{v}^2_{n,X} - E\widetilde{v}^2_{n,X}| + |\widehat{v}^2_{n,X} - \widetilde{v}^2_{n,X}|$, to prove (3.1), it suffices to show that

$$
\text{(3.2)} \quad |E\widetilde{v}^2_{n,X} - v^2_{n,X}| = o(v^2_{n,X}), \quad E|\widetilde{v}^2_{n,X} - E\widetilde{v}^2_{n,X}| = o(v^2_{n,X}), \quad E|\widehat{v}^2_{n,X} - \widetilde{v}^2_{n,X}| = o(v^2_{n,X}).
$$

(i) By Lemma 4.3(i), $B_n = O(v^2_{n,X})$. Moreover, by (4.6), $|E\widetilde{v}^2_{n,X} - v^2_{n,X}| = o(B_n)$ and $E|\widehat{v}^2_{n,X} - \widetilde{v}^2_{n,X}| = o(B_n)$, whereas by Lemma 4.4(i), $E|\widetilde{v}^2_{n,X} - E\widetilde{v}^2_{n,X}| = o(B_n)$, which proves (3.2).

(ii) Let $\{X_j\}$ has SM. Since (2.15) and (2.16) imply (2.2) and (2.3), then (3.1) follows from part (i).

Let $\{X_j\}$ has LM. By Lemma 4.3(ii), $B_n^{1+2d} = O(v^2_{n,X})$. In addition, by (4.7) $|E\widetilde{v}^2_{n,X} - v^2_{n,X}| = o(B_n^{1+2d})$ and $E|\widehat{v}^2_{n,X} - \widetilde{v}^2_{n,X}| = o(B_n^{1+2d})$, whereas by Lemma 4.4(ii) $E|\widetilde{v}^2_{n,X} - E\widetilde{v}^2_{n,X}| = o(B_n^{1+2d})$, proving (3.2). $\square$

The next lemma analyzes the asymptotic behavior of the variance estimate $\widetilde{v}^2_{n,Y}$ in the case of heteroscedastic observations.

Lemma 3.2 *Let $Y_j = h_{nj}X_j$, $j = 1, \cdots, n$ satisfy Assumption 2.1.*

(i) *Suppose $\{X_j\}$ is a SM process satisfying (2.1), or LM process satisfying (2.13) and (2.14), and b_{nj}'s satisfy (2.15) and (2.16). Then*

$$
\text{(3.3)} \qquad \widetilde{v}^2_{n,Y} = \sigma^2_h v^2_{n,X}(1 + o_p(1)), \qquad \widetilde{\gamma}_Y(0) = \sigma^2_h(1 + o_p(1)).
$$

(ii) *Let b_{nj}'s and $\{X_j\}$ be as in Theorem 2.1. Then (3.3) holds.*

Proof. (i) The first claim follows from $\widetilde{v}^2_{n,Y}=\sigma^2_h v^2_{n,X}+(E\widetilde{v}^2_{n,Y}-\sigma^2_h v^2_{n,X})+(\widetilde{v}^2_{n,Y}-E\widetilde{v}^2_{n,Y})$, using

$$|E\widetilde{v}^2_{n,Y}-\sigma^2_h v^2_{n,X}|=o(v^2_{n,X}),\quad E|\widetilde{v}^2_{n,Y}-E\widetilde{v}^2_{n,Y}|=o(v^2_{n,X}). \tag{3.4}$$

For SM $\{X_j\}$, by Lemma 4.3(i), $B_n=O(v^2_{n,X})$, (4.6) implies $|E\widetilde{v}^2_{n,Y}-\sigma^2_h v^2_{n,X}|=o(B_n)$, and by Lemma 4.4(i) $E|\widetilde{v}^2_{n,Y}-E\widetilde{v}^2_{n,Y}|=o(B_n)$, which proves (3.4).

For LM $\{X_j\}$, by Lemma 4.3(ii), $B_n^{1+2d}=O(v^2_{n,X})$, (4.7) implies $|E\widetilde{v}^2_{n,Y}-\sigma^2_h v^2_{n,X}|=o(B_n^{1+2d})$, and by Lemma 4.4(ii), $E|\widetilde{v}^2_{n,Y}-E\widetilde{v}^2_{n,Y}|=o(B_n^{1+2d})$, proving (3.4).

Finally, by Lemma 4.1, $\widetilde{\gamma}_Y(0)=\sigma^2_h\gamma_X(0)+o_p(1)=\sigma^2_h(1+o_p(1))$, since $EX_1^2=1$ and $\sigma^2_h\asymp 1$, which proves the second claim in (3.3).

(ii) The claims in (3.3) follows using the same argument as in SM case in (i). $\square$

The next lemma focuses on the estimation of the SD in the signal plus noise setting $\hat{Y}_j:=Y_j+\xi_j$, $j=1,\cdots,n$, when the signal $Y_j=h_{nj}X_j$. It describes a class of noises ξ_j for which $v^2_{n,X}$ can still be consistently estimated, while $\{X_j\}$ may have either SM or LM. Such setting is common in non-parametric regression estimation, when Y_j are not directly observed and computation of standard errors is based on residuals $\hat{Y}_j=Y_j+\xi_j$, containing noise ξ_j. Recall the notation $\widetilde{\gamma}_\xi(0)=n^{-1}\sum_{j=1}^n\xi_j^2$.

Lemma 3.3 *Let $\hat{Y}_j=Y_j+\xi_j$, where $Y_j=h_{nj}X_j$, $j=1,\cdots,n$ satisfies Assumption 2.1, $\{X_j\}$ is a SM process satisfying (2.1), or LM process satisfying (2.13) and (2.14), and the noise ξ_j, $j=1,\cdots,n$ is such that*

$$\begin{aligned} E\widetilde{\gamma}_\xi(0) &= o(B_n^{-1}), && \text{if } \{X_j\} \text{ has SM,} \\ &= o(B_n^{-1+2d}), && \text{if } \{X_j\} \text{ has LM.} \end{aligned} \tag{3.5}$$

Furthermore, suppose that b_{nj}'s satisfy (2.15) and (2.16). Then

$$\widetilde{v}_{n,\hat{Y}}=\sigma_h v_{n,X}(1+o_p(1)),\qquad \widetilde{\gamma}_{\hat{Y}}(0)=\sigma^2_h(1+o_p(1)). \tag{3.6}$$

Proof. It suffices to show that

$$E|\widetilde{v}^2_{n,\hat{Y}}-\widetilde{v}^2_{n,Y}| = o(v^2_{n,X}), \tag{3.7}$$

$$E|\widetilde{\gamma}_{\hat{Y}}(0)-\widetilde{\gamma}_Y(0)| = o(1), \tag{3.8}$$

which, in view of (3.3), implies (3.6).

Proof of (3.7). By Lemma 4.2,

$$E|\widetilde{v}^2_{n,\hat{Y}}-\widetilde{v}^2_{n,Y}|\le E\widetilde{v}^2_{n,\xi}+2(E\widetilde{v}^2_{n,Y}E\widetilde{v}^2_{n,\xi})^{1/2}. \tag{3.9}$$

In the proof of Lemma 3.2 it was shown, that in SM case $E\widetilde{v}^2_{n,Y}=O(B_n)$ and $v^2_{n,X}=O(B_n)$, while in the LM case, $E\widetilde{v}^2_{n,Y}=O(B_n^{1+2d})$ and $v^2_{n,X}=O(B_n^{1+2d})$. Since $|\widetilde{\gamma}_\xi(k)|\le\widetilde{\gamma}_\xi(0)$, then,

$E\widetilde{v}^2_{n,\xi} = E[\widetilde{\gamma}_\xi(0)](\sum_{j=1}^n |b_{nj}|)^2 \le CE[\widetilde{\gamma}_\xi(0)]B_n^2$. Together with (3.5) and (3.9) this implies that in SM case $E|\widetilde{v}^2_{n,\hat{Y}} - \widetilde{v}^2_{n,Y}| = o(B_n)$, while in LM case it is of order $o(B_n^{1+2d})$, which proves (3.7).

Proof of (3.8). The bound $E|\widetilde{\gamma}_{\hat{Y}}(0) - \widetilde{\gamma}_Y(0)| \le E\widetilde{\gamma}_\xi(0) + 2(E\widetilde{\gamma}_Y(0)E\widetilde{\gamma}_\xi(0))^{1/2}$, together with (3.5) and $E\widetilde{\gamma}_Y(0) = O(1)$ implies (3.8). □

4 Auxiliary results

In this section we present the main proofs and some auxiliary results.

Preliminaries. Let

$$\phi(v) := \sum_{j=1}^n e^{\mathbf{i}jv} b_{nj}, \quad I_X(v) := (2\pi n)^{-1}\Big|\sum_{j=1}^n e^{\mathbf{i}vj} X_j\Big|^2, \quad v \in \Pi.$$

The sample auto-covariances $\widetilde{\gamma}_X(k) = \widetilde{\gamma}_X(-k)$, $\widehat{\gamma}_X(k) = \widehat{\gamma}_X(-k)$, $\widetilde{v}^2_{n,X}$ of (2.9) and $\widehat{v}^2_{n,X}$ of (1.4) can be written as

$$\begin{aligned} \widetilde{\gamma}_X(k) &= \int_\Pi e^{\mathbf{i}kv} I_X(v)dv, \qquad \widehat{\gamma}_X(k) = \int_\Pi e^{\mathbf{i}kv} I_{X-\bar{X}}(v)dv, \\ \widetilde{v}^2_{n,X} &= \sum_{j,k=1}^n b_{nj}b_{nk}\widetilde{\gamma}_X(j-k) = \int_\Pi |\phi(v)|^2 I_X(v)dv, \\ \widehat{v}^2_{n,X} &= \int_\Pi |\phi(v)|^2 I_{X-\bar{X}}(v)dv. \end{aligned} \tag{4.1}$$

Note that, in view of (4.1),

$$\begin{aligned} \widetilde{v}^2_{n,X} &= \frac{1}{2\pi n}\int_\Pi |\phi(u)|^2 \sum_{t,s=1}^n e^{\mathbf{i}(t-s)u} X_t X_s du = \frac{1}{n}\sum_{t,s=1}^n d_n(t-s)X_tX_s, \\ d_n(s) &:= (2\pi)^{-1}\int_\Pi e^{\mathbf{i}su}|\phi(u)|^2 du = \sum_{j=1}^{n-|s|} b_{nj}b_{n,j+|s|}, \quad |s| \le n-1. \end{aligned} \tag{4.2}$$

Furthermore, we have the following facts:

$$\begin{aligned} \sum_{s=1}^n d_n^2(s) &= \sum_{s=1}^n \Big(\sum_{j=1}^{n-|s|} b_{nj}b_{n,j+|s|}\Big)^2 \le \Big(\sum_{j=1}^n |b_{nj}|\Big)^2 \sum_{s=1}^n b^2_{ns}, \\ \sum_{s=0}^n |d_n(s)| &\le (\sum_{s=1}^n |b_{ns}|)^2, \qquad \max_{1\le s\le n} |d_n(s)| \le \sum_{j=1}^n b^2_{nj}. \end{aligned} \tag{4.3}$$

Lemma 4.1 *Let $Y_j = h_{nj}X_j$, $j = 1, \cdots, n$ satisfy Assumption 2.1.*

(i) *Then*

$$\widetilde{\gamma}_Y(k) = \sigma_h^2 \gamma_X(k) + o_p(1), \tag{4.4}$$

$$\widetilde{\gamma}_Y(k)/\widetilde{\gamma}_Y(0) = \gamma_X(k)/\gamma_X(0) + o_p(1), \quad k = 0, 1, 2, \cdots . \tag{4.5}$$

(ii) *Moreover, if $\{X_j\}$ is a SM process satisfying (2.1) and (2.3) holds, then*

$$\big|E\widetilde{v}_{n,Y}^2 - \sigma_h^2\, v_{n,X}^2\big| = o(B_n), \quad E\big|\widehat{v}_{n,Y}^2 - \widetilde{v}_{n,Y}^2\big| = o(B_n). \tag{4.6}$$

In addition, if $\{X_j\}$ is a LM process satisfying (2.13) and (2.15) holds, then

$$\big|E\widetilde{v}_{n,Y}^2 - \sigma_h^2\, v_{n,X}^2\big| = o\big(B_n^{1+2d}\big), \quad E\big|\widehat{v}_{n,Y}^2 - \widetilde{v}_{n,Y}^2\big| = o(B_n^{1+2d}). \tag{4.7}$$

Proof. (i) We show that

$$E\big|\widetilde{\gamma}_Y(k) - E\widetilde{\gamma}_Y(k)\big| = o(1), \tag{4.8}$$

$$|E\widetilde{\gamma}_Y(k) - \sigma_h^2\gamma_X(k)| \le C|\gamma_X(k)|(k/n) = o(1). \tag{4.9}$$

Since $\widetilde{\gamma}_Y(k) - \sigma_h^2\gamma_X(k) = \widetilde{\gamma}_Y(k) - E\widetilde{\gamma}_Y(k) + E\widetilde{\gamma}_Y(k) - \sigma_h^2\gamma_X(k)$, this implies (4.4).

Proof of (4.8). Write $\widetilde{\gamma}_Y(k) = \sum_{j=1}^{n-k} c_{nj}Z_j$, where $c_{nj} := n^{-1}h_{nj}h_{n,j+k}$ and $Z_j := X_jX_{j+k}$. Since the noise $\{\zeta_k\}$ in definition of X_j, (1.5), is stationary ergodic, by Theorem 3.5.8 in Stout (1974), the process $\{X_j\}$ and $\{Z_j\}$ are stationary and ergodic. Assumption (2.20) implies $\max_{1\le j\le n}|h_{nj}| = O(1)$, $\sum_{j=1}^{n-k}|c_{nj}| \le n^{-1}\sum_{j=1}^{n} h_{nj}^2 = O(1)$ and $|c_{n1}| + \sum_{j=2}^{n-k}|c_{nj} - c_{n,j-1}| \le Cn^{-1}\{|h_{n1}| + \sum_{j=2}^{n}|h_{nj} - h_{n,j-1}|\} = o(1)$. Thus, by Lemma 4.7 below, $E\big|\widetilde{\gamma}_Y(k) - E[\widetilde{\gamma}_Y(k)]\big| = o(1)$.

Proof of (4.9). From $E[Y_jY_{j+k}] = h_{nj}h_{n,j+k}\gamma_X(k)$,

$$\begin{aligned} E\widetilde{\gamma}_Y(k) &= n^{-1}\sum_{j=1}^{n-k} E[Y_jY_{j+k}] = \gamma_X(k)n^{-1}\sum_{j=1}^{n-k} h_{nj}h_{n,j+k} \\ &=: \gamma_X(k)\widetilde{\gamma}_h(k), \qquad k = 0, 1, \cdots . \end{aligned}$$

Moreover,

$$\begin{aligned} |\widetilde{\gamma}_h(k) - \widetilde{\gamma}_h(0)| &= \big|n^{-1}\sum_{j=1}^{n-k} h_{nj}h_{n,j+k} - n^{-1}\sum_{j=1}^{n} h_{nj}^2\big| \\ &\le n^{-1}\sum_{j=1}^{n-k}|h_{n,j+k} - h_{nj}||h_{nj}| + n^{-1}\sum_{j=n-k+1}^{n} h_{nj}^2 \le C(k/n), \end{aligned}$$

because by (2.20), $\sum_{j=1}^{n-k}|h_{n,j+k} - h_{nj}| \le \sum_{j=1}^{n-k}|\{h_{n,j+k} - h_{n,j+k-1}\} + \cdots + \{h_{n,j+1} - h_{nj}\}| \le k\sum_{j=1}^{n-1}|h_{n,j+1} - h_{nj}| \le Ck$ and $\max_j|h_{nj}| = O(1)$. The above bound and the fact $\sigma_h^2 = \widetilde{\gamma}_h(0)$, in turn yield

$$\big|E\widetilde{\gamma}_Y(k) - \sigma_h^2\gamma_X(k)| \le |\gamma_X(k)|\big|\widetilde{\gamma}_h(k) - \widetilde{\gamma}_h(0)\big| \le C|\gamma_X(k)|(k/n),$$

which completes the proof of (4.9) and (4.4).

The claim (4.5) follows from (4.4), noting that $\widetilde{\gamma}_Y(0) = \sigma_h^2\gamma_X(0) + o_p(1) = \sigma_h^2\gamma_X(0)(1 + o_p(1))$, because by (2.20), $\sigma_h^2 \geq c > 0$ as $n \to \infty$.

(ii) *Proof of (4.6).* By (4.9),

$$
\begin{aligned}
(4.10) \qquad \left|E\widetilde{v}_{n,Y}^2 - \sigma_h^2 v_{n,X}^2\right| &= \Big|\sum_{j,\,k=1}^{n} b_{nj}b_{nk}\left(E\widetilde{\gamma}_Y(j-k) - \sigma_h^2\gamma_X(j-k)\right)\Big| \\
&\leq C\sum_{j,\,k=1}^{n} |b_{nj}b_{nk}|\,|\gamma_X(j-k)|(|j-k|/n) =: q_n.
\end{aligned}
$$

Then $q_n \leq C\big(\sum_{j=1}^n b_{nj}^2\big)\sum_{k=-n}^n |\gamma_X(k)|(|k|/n) = o(B_n)$, because (2.1) implies $\sum_{|j|\leq n}|\gamma_X(j)|(|j|/n) = o(1)$, which proves the first claim of (4.6).

To show the second claim, use $\widehat{v}_{n,X}^2 = \widetilde{v}_{n,X-\bar{X}}^2$ and Lemma 4.2 below, to get

$$
(4.11) \qquad E|\widehat{v}_{n,X}^2 - \widetilde{v}_{n,X}^2| = E|\widetilde{v}_{n,X-\bar{X}}^2 - \widetilde{v}_{n,X}^2| \leq E\widetilde{v}_{n,\bar{X}}^2 + 2(E\widetilde{v}_{n,X}^2 E\widetilde{v}_{n,\bar{X}}^2)^{1/2}.
$$

By Lemma 4.3(i) below, $\sigma_h^2 = O(1)$ and the first claim of (4.6) imply $E\widetilde{v}_{n,X}^2 = O(B_n)$. Therefore, it suffices to show $E\widetilde{v}_{n,\bar{X}}^2 = o(B_n)$. Now, $E\widetilde{v}_{n,\bar{X}}^2 \leq E\bar{X}^2(\sum_{j=1}^n b_{nj})^2$, where $E\bar{X}^2 = n^{-2}\sum_{j,k=1}^n \gamma_X(j-k) \leq n^{-1}\sum_{j\in\mathbb{Z}}|\gamma_X(j)| = O(n^{-1})$ by (2.1). Thus, $E\widetilde{v}_{n,\bar{X}}^2 \leq Cn^{-1}(\sum_{j=1}^n |b_{nj}|)^2 = o(B_n)$ by (2.3), which completes the proof of the second claim in (4.6).

Proof of (4.7). Since $\gamma_X(k)$ satisfies (2.13), then $|\gamma_X(k)||k| \leq Ck^{2d} \leq Cn^{2d}$, for $k = 1, \cdots, n$. Use this together with $\sum_{j=1}^n |b_{nj}| = O(B_n) = o(n)$ by (2.15), to obtain

$$
q_n \leq Cn^{-1+2d}\big(\sum_{j=1}^n |b_{nj}|\big)^2 \leq Cn^{-1+2d}B_n^2 = CB_n^{1+2d}(B_n/n)^{1-2d} = o(B_n^{1+2d}).
$$

This proves the first claim of (4.7).

To show the second claim, note that by Lemma 4.3(ii), $\sigma_h^2 = O(1)$ and the first claim of (4.7) imply $E\widetilde{v}_{n,X}^2 = O(B_n^{1+2d})$. Thus, by (4.11), it suffices to show $E\widetilde{v}_{n,\bar{X}}^2 = o(B_n^{1+2d})$. Under (2.13),

$$
E\bar{X}^2 \leq Cn^{-1}\sum_{j=0}^n |\gamma_X(j)| \leq Cn^{-1}\sum_{j=1}^n j^{-1+2d} \leq Cn^{-1+2d}.
$$

Therefore, by (2.15),

$$
E\widetilde{v}_{n,\bar{X}}^2 \leq Cn^{-1+2d}(\sum_{j=1}^n |b_{nj}|)^2 \leq Cn^{-1+2d}B_n^2 = CB_n^{1+2d}(B_n/n)^{1-2d} = o(B_n^{1+2d}),
$$

which completes the proof of the second claim in (4.7). □

Lemma 4.2 *For any r.v.'s $U_1, \cdots, U_n$ and $V_1, \cdots, V_n$, define $\widetilde{v}_{n,V}^2$ and $\widetilde{v}_{n,U}^2$ analogous to (2.9). Then*

$$
(4.12) \qquad E|\widetilde{v}_{n,U+V}^2 - \widetilde{v}_{n,V}^2| \leq E\widetilde{v}_{n,V}^2 + 2(E\widetilde{v}_{n,U}^2 E\widetilde{v}_{n,V}^2)^{1/2}.
$$

Proof. By (4.1), $\widetilde{v}^2_{n,U+V} - \widetilde{v}^2_{n,U} = \int_\Pi |\phi(v)|^2 (I_{U+V}(v) - I_U(v))dv$. Apply the elementary inequality $\big||a|^2 - |b|^2\big| \le |a-b||a+b| \le |a-b|^2 + 2|b||a-b|$, for any complex numbers a, b, to obtain $|I_{U+V}(v) - I_V(v)| \le I_V(v) + 2(I_U(v)\, I_V(v))^{1/2}$. Therefore, by the Cauchy–Schwarz inequality,

$$
\begin{aligned}
|\widetilde{v}^2_{n,U+V} - \widetilde{v}^2_{n,U}| &\le \int_\Pi |\phi(v)|^2 (I_V(v) + 2(I_U(v)\, I_V(v))^{1/2}) dv \le \widetilde{v}^2_{n,V} + 2\widetilde{v}_{n,U}\widetilde{v}_{n,V}, \\
E|\widetilde{v}^2_{n,U+V} - \widetilde{v}^2_{n,U}| &\le E\widetilde{v}^2_{n,V} + 2(E\widetilde{v}^2_{n,U} E\widetilde{v}^2_{n,V})^{1/2},
\end{aligned}
$$

which completes the proof. □

In the next lemma we obtain the order of the variance $v^2_{n,X}$, given by (1.2). The process $\{X_j\}$ has SM in part (i) and LM in part (ii).

Lemma 4.3 (i) *Let $\{X_j\}$ and b_{nj} satisfy assumptions of Theorem 2.1 or 2.2. Then*

$$v^2_{n,X} \asymp B_n. \tag{4.13}$$

(ii) *Let $\{X_j\}$ of (1.5) satisfy (2.14), b_{nj}'s satisfy (2.16) and $\sum_{j=1}^n |b_{nj}| = O(B_n)$. Then,*

$$v^2_{n,X} = O(B_n^{1+2d}); \qquad v^2_{n,X} \asymp B_n^{1+2d} \quad \textit{if } B_n \to \infty. \tag{4.14}$$

(iii) *If $\{X_j\}$ of (1.5) has the bounded spectral density f, then $v^2_{n,X} = O(B_n)$.*

Proof. (i) Under assumptions of Theorem 2.1, by Proposition 2.2 (a) of Abadir *et al.* (2013), $v^2_{n,X} \sim s^2_X B_n$ which implies (4.13), while under assumptions of Theorem 2.2, it is shown in the proof of Theorem 2.2.

(ii) *Upper bound*. To prove the first claim in (4.14), recall that by (2.14), $f(u) \sim c_f |u|^{-2d}$, $u \to 0$. Let $\phi(u) = \sum_{j=1}^n e^{\mathbf{i}ju} b_{nj}$, $u \in \Pi$, $k_n := 1/(\epsilon B_n)$, where $0 < \epsilon < 1$. Use $\big|\sum_{l=1}^j e^{\mathbf{i}lu}\big| = |\sin(ju/2)/\sin(u/2)| \le \pi u^{-1}$, $|u| \le \pi$, $j \ge 1$, summation by parts and (2.16), to obtain

$$
\begin{aligned}
|\phi(u)| &= \Big|\sum_{j=1}^{n-1} \big(\sum_{l=1}^{j} e^{\mathbf{i}lu}\big)(b_{nj} - b_{n,j+1}) + b_{nn} \sum_{l=1}^{n} e^{\mathbf{i}lu}\Big| \\
&\le \pi |u|^{-1} \big(\sum_{j=1}^{n-1} |b_{nj} - b_{n,j+1}| + |b_{nn}|\big) \le C|u|^{-1}, \quad u \in \Pi,
\end{aligned} \tag{4.15}
$$

where $C < \infty$ does not depend on n and b_{nj}.

From (4.15), $|\phi(u)|^2 \le (\sum_{j=1} |b_{nj}|)^2 \le CB_n^2$ by assumptions of (ii). Hence,

$$
\begin{aligned}
v^2_{n,X} &= \int_\Pi f(u)|\phi(u)|^2 du \le C \int_\Pi |\phi(u)|^2 |u|^{-2d} du \\
&\le CB_n^2 \int_{|u| \le k_n} |u|^{-2d} du + C \int_{k_n \le |u| \le \pi} |u|^{-2-2d} du \le C\{B_n^2 k_n^{1-2d} + k_n^{-1-2d}\} \\
&\le C\{B_n^2 (\epsilon B_n)^{-1+2d} + (\epsilon B_n)^{1+2d}\} \le C(\epsilon) B^{1+2d},
\end{aligned}
$$

which proves the first claim in (4.14).

Lower bound. Let $f_* := f(k_n)/2$. To establish the second claim in (4.14), we shall show that for some $\epsilon > 0$ and $c > 0$,

$$v_{n,X}^2 \geq c f_* B_n, \quad n \to \infty. \tag{4.16}$$

Since $f_* \sim (c_f/2)(\epsilon B_n)^{2d}$, (4.16) implies $v_{n,X}^2 \geq C(\epsilon) B_n^{1+2d}$, $n \to \infty$ with $C(\epsilon) > 0$, which, as easily seen, verifies the seond claim of (ii).

To prove (4.16) note that, for $n \to \infty$, $k_n \to 0$, $\inf_{|u| \leq k_n} f(u) \geq f_*$ and $\sup_{k_n < |u| \leq \pi} f(u) \leq 4 f_*$. Setting $j_n := \int_{k_n \leq |u| \leq \pi} |\phi(u)|^2 du$, since $\gamma_X(k) = \int_\Pi e^{\mathbf{i} k u} f(u)\, du$, we obtain

$$\begin{aligned} v_{n,X}^2 &= \sum_{j,k=1}^n b_{nj} b_{nk} \gamma_X(j-k) = \int_\Pi f(u) |\phi(u)|^2 du \\ &\geq f_* \int_{|u| \leq k_n} |\phi(u)|^2 du - 4 f_* j_n = f_* \Big(\int_{|u| \leq \pi} |\phi(u)|^2 du - 5 j_n \Big) \\ &= f_* \, (2\pi B_n - 5 j_n) \geq f_* \, 2\pi (B_n - j_n). \end{aligned} \tag{4.17}$$

By (4.15), $j_n \leq C \int_{k_n \leq |u| \leq \pi} u^{-2} du \leq C k_n^{-1} = C \epsilon B_n \leq B_n/2$, if $\epsilon \leq 1/(2C)$, which together with (4.17) implies (4.16).

(iii) Note that $v_{n,X}^2 = \int_\Pi \big| \sum_{j=1}^n e^{\mathbf{i} j u} b_{nj} \big|^2 f(u) du \leq C \int_\Pi \big| \sum_{j=1}^n e^{\mathbf{i} j u} b_{nj} \big|^2 du = C B_n$. $\Box$

In the next lemma, in part (i) $\{X_j\}$ has SM, and in part (ii) SM or LM. In (ii), for SM process $\{X_j\}$ we set $d = 0$ in (4.19).

Lemma 4.4 *Let $Y_j = h_{nj} X_j$, $j = 1, \cdots, n$ satisfy Assumption 2.1, and $\widetilde{v}_{n,Y}^2$ be as in (2.9).*
(i) *If b_{nj} and $\{X_j\}$ satisfy assumptions of Theorem 2.1 or 2.2, then*

$$E \big| \widetilde{v}_{n,Y}^2 - E \widetilde{v}_{n,Y}^2 \big| = o(B_n). \tag{4.18}$$

(ii) *If b_{nj} and $\{X_j\}$ satisfy assumptions of Theorem 2.3, then*

$$E \big| \widetilde{v}_{n,Y}^2 - E \widetilde{v}_{n,Y}^2 \big| = o(B_n^{1+2d}). \tag{4.19}$$

Proof. (i) Since $\{X_j\}$ has SM, by (4.6) and (4.13),

$$E \widetilde{v}_{n,Y}^2 \leq \sigma_h^2 v_{n,X}^2 + o(B_n) = O(B_n). \tag{4.20}$$

We split the proof into two cases, c1) and c2).
c1) Suppose $\max_j E \zeta_j^4 < \infty$. Letting $a_j = 0$ for $j < 0$, by (4.2), one can write using notation $\beta_{n,jk} := n^{-1} \sum_{t,s=1}^n d_n(t-s) h_{nt} a_{t-j} h_{ns} a_{s-k}$,

$$\widetilde{v}_{n,Y}^2 = n^{-1} \sum_{t,s=1}^n d_n(t-s) Y_t Y_s = \sum_{j,k \in \mathbb{Z}} \beta_{n,jk} \zeta_j \zeta_k. \tag{4.21}$$

By Lemma 4.5, if the sums

$$s_{n,1} \;:=\; \sum_{j\in\mathbb{Z}} |\beta_{n,jj}|, \qquad s_{n,2} := \sum_{j\in\mathbb{Z}} |\beta_{n,jj} - \beta_{n,j-1,j-1}|$$

satisfy $s_{n,1} = O(\alpha_n)$ and $s_{n,2} = o(\alpha_n)$, for some α_n, then,

$$E|\widetilde{v}^2_{n,Y} - E\widetilde{v}^2_{n,Y}| \;=\; O(t_n) + o(\alpha_n), \quad \text{where}$$

$$t_n^2 \;:=\; n^{-2}(h_n^*)^4 \sum_{j_1,k_1,j_2,k_2=1}^{n} |d_n(j_1-k_1)d_n(j_2-k_2)||\gamma_X(j_1-j_2)\gamma_X(k_1-k_2)|, \tag{4.22}$$

where $h_n^* := \max_{1\le j\le n}|h_{nj}| = O(1)$, by (2.20). Thus, to prove (4.18), it suffices to verify

$$t_n = o(B_n), \quad s_{n,1} = O(B_n), \quad s_{n,2} = o(B_n). \tag{4.23}$$

To bound t_n, use (4.22), $|d_n(j_1-k_1)d_n(j_2-k_2)| \le d_n^2(j_1-k_1) + d_n^2(j_2-k_2)$ and (2.1), to obtain

$$\begin{aligned} t_n^2 \;&\le\; Cn^{-1}\sum_{s_1=-n}^{n} d_n^2(s_1)\sum_{s_2=-n}^{n}|\gamma_X(s_2)|\sum_{s_3=-n}^{n}|\gamma_X(s_3)| \\ &\le\; Cn^{-1}\sum_{s_1=-n}^{n} d_n^2(s_1) \le Cn^{-1}(\sum_{j=1}^{n}|b_{nj}|)^2 B_n = o(B_n^2), \end{aligned} \tag{4.24}$$

by (4.3) and (2.3), which verifies $t_n = o(B_n)$.

To bound $s_{n,1}$, notice that by definition the matrix $(d_n(t-s))_{t,s,=1,\cdots,n}$ is non-negative definite, and therefore $\beta_{n,jj} \ge 0$. Hence, $s_{n,1} = \sum_{j\in\mathbb{Z}}\beta_{n,jj} = \sigma_\zeta^{-2}n^{-1}\sum_{t,s,=1}^{n} d_n(t-s)h_{nt}h_{ns}\gamma_X(t-s) = \sigma_\zeta^{-2}E\widetilde{v}^2_{n,Y} = O(B_n)$, by (4.20).

To bound $s_{n,2}$, set $h_{n0} = h_{n,n+1} = 0$. Then

$$\begin{aligned} \beta_{n,jj} - \beta_{n,j-1,j-1} \;&=\; n^{-1}\sum_{t,s=1}^{n} d_n(t-s)h_{nt}h_{ns}(a_{t-j}a_{s-j} - a_{t-j+1}a_{s-j+1}) \\ &=\; n^{-1}\sum_{t,s=1}^{n+1} d_n(t-s)(h_{nt}h_{ns} - h_{n,t-1}h_{n,s-1})a_{t-j}a_{s-j}. \end{aligned}$$

Hence, by $\sum_{j=0}^{\infty} a_j^2 < \infty$, (4.3) and (2.20),

$$\begin{aligned} s_{n,2} \;&=\; \sum_{j\in\mathbb{Z}}|\beta_{n,jj} - \beta_{n,j-1,j-1}| \le n^{-1}\sum_{j=0}^{\infty} a_j^2 \sum_{t,s=1}^{n+1}|d_n(t-s)(h_{nt}h_{ns} - h_{n,t-1}h_{n,s-1})| \\ &\le\; Cn^{-1}h_n^*\sum_{|k|\le n}|d_n(k)|\sum_{t=1}^{n+1}|h_{nt} - h_{n,t-1}| \le Cn^{-1}(\sum_{j=1}^{n}|b_{nj}|)^2 = o(B_n), \end{aligned} \tag{4.25}$$

which completes the proof of (4.23) in case c1).

c2) Let $\max_j E\zeta_j^4 = \infty$. We establish (4.18) by using truncation argument. For $K > 0$, let $\zeta_j^- := \zeta_j I(|\zeta_j| \le K)$, $\zeta_j^+ := \zeta_j I(|\zeta_j| > K)$. Then $\zeta_j = \zeta_j^- + \zeta_j^+ = \{\zeta_j^- - E[\zeta_j^-|\mathcal{F}_{j-1}]\} + \{\zeta_j^+ - E[\zeta_j^+|\mathcal{F}_{j-1}]\} =: \zeta_{j,1} + \zeta_{j,2}$. Note that $E\zeta_{j,1}^2 \le 2E(\zeta_1^-)^2 \le 2\sigma_\zeta^2$ and $E\zeta_{j,2}^2 \le 2E(\zeta_1^+)^2 = \delta_K \to 0$ as $K \to \infty$. Then

$$X_t = \sum_{k=0}^{\infty} a_k \zeta_{t-k} = \sum_{k=0}^{\infty} a_k \zeta_{t-k,1} + \sum_{k=0}^{\infty} a_k \zeta_{t-k,2} =: X_{1,t} + X_{2,t}.$$

Setting $Y_t = h_{nt} X_t = h_{nt} X_{1,t} + h_{nt} X_{2,t} =: U_t + Z_t$, we have $\widetilde{v}_{n,Y}^2 = \widetilde{v}_{n,U+Z}^2$. Use the equality $\widetilde{v}_{n,Y}^2 - E\widetilde{v}_{n,Y}^2 = \widetilde{v}_{n,Y}^2 - \widetilde{v}_{n,U}^2 + \widetilde{v}_{n,U}^2 - E\widetilde{v}_{n,U}^2 + E[\widetilde{v}_{n,U}^2 - \widetilde{v}_{n,Y}^2]$, to bound

$$E|\widetilde{v}_{n,Y}^2 - E\widetilde{v}_{n,Y}^2| \le E|\widetilde{v}_{n,U}^2 - E\widetilde{v}_{n,U}^2| + 2E|\widetilde{v}_{n,Y}^2 - \widetilde{v}_{n,U}^2|.$$

To prove (4.18), it suffices to show that the two terms on the r.h.s. are of order $o(B_n)$, as $n \to \infty$, $K \to \infty$. Since $\{\zeta_{j,1}\}$ is a stationary ergodic m.d.s. with all moments finite, then by c1) above, $E|\widetilde{v}_{n,U}^2 - E\widetilde{v}_{n,U}^2| = o(E\widetilde{v}_{n,U}^2) = o(B_n)$, $\forall K > 0$. To bound the second term, use Lemma 4.2, to obtain

$$E|\widetilde{v}_{n,Y}^2 - \widetilde{v}_{n,U}^2| = E|\widetilde{v}_{n,U+Z}^2 - \widetilde{v}_{n,U}^2| \le E\widetilde{v}_{n,Z}^2 + 2(E\widetilde{v}_{n,U}^2 E\widetilde{v}_{n,Z}^2)^{1/2}.$$

By the definition of U, (4.21) and (4.20), $E\widetilde{v}_{n,U}^2 \le CE\widetilde{v}_{n,Y}^2 = O(B_n)$, while $E\widetilde{v}_{n,Z}^2 = (E\zeta_{1,2}^2/\sigma_\zeta^2)E\widetilde{v}_{n,Y}^2 \le C\delta_K B_n$, where $\delta_K \to 0$, as $K \to \infty$. Hence $E|\widetilde{v}_{n,Y}^2 - \widetilde{v}_{n,U}^2| = o(B_n)$, as $n \to \infty$, $K \to \infty$, which completes the proof of (i) in case c2).

(ii) Let $\{X_j\}$ have SM. Since (2.15) and (2.16) imply (2.2) and (2.3), (4.18) follows from part (i) above.

It remains to prove (2.3) in the case when $\{X_j\}$ has LM. Observe that by (4.7) and (4.14),

$$E\widetilde{v}_{n,Y}^2 \le \sigma_h^2 v_{n,X}^2 + o(B_n^{1+2d}) = O(B_n^{1+2d}). \tag{4.26}$$

We verify (4.19) in case c1), $\max_j E\zeta_j^4 < \infty$. (In case c2), $\max_j E\zeta_j^4 = \infty$, (4.19) follows by the same argument as in (i), and using (4.26) instead of (4.20).)

Because of (4.22), it suffices to verify

$$t_n = o(B_n^{1+2d}), \quad s_{n,1} = O(B_n^{1+2d}), \quad s_{n,2} = o(B_n^{1+2d}). \tag{4.27}$$

Recall that by assumptions (2.13) and (2.15), $\gamma_X(k) = O(k^{-1+2d})$, $d \in (0, 1/2)$ and $\sum_{j=1}^n |b_{nj}| = O(B_n) = o(n)$.

We split the proof for t_n into three cases, $0 < d < 1/4$, $1/4 < d < 1/2$ and $d = 1/4$.

Consider the case $0 < d < 1/4$. Applying $\sum_{s=0}^n |\gamma_X(s)| \le C\sum_{s=1}^n s^{-1+2d} \le Cn^{2d}$ and (4.3) in the first bound of (4.24), gives

$$\begin{aligned} t_n^2 &\le Cn^{-1+4d} \sum_{s=-n}^{n} d_n^2(s) \le Cn^{-1+4d} (\sum_{j=1}^{n} |b_{nj}|)^2 B_n \\ &\le Cn^{-1+4d} B_n^3 = CB_n^{2+4d} (B_n/n)^{1-4d} = o(B_n^{2+4d}), \end{aligned}$$

which verifies (4.27).

Next, consider the case $1/4 < d < 1/2$. Use $|\gamma_X(j_1-j_2)\gamma_X(k_1-k_2)| \le \gamma_X^2(j_1-j_2) + \gamma_X^2(k_1-k_2)$ in (4.22), and $\sum_{s=1}^n \gamma_X^2(s) \le C\sum_{s=1}^n s^{-2+4d} \le Cn^{-1+4d}$, to obtain

$$\begin{aligned} t_n^2 &\le Cn^{-1}\big(\sum_{j=-n}^{n} |d_n(j)|\big)^2 \sum_{t=-n}^{n} \gamma_X^2(t) \le Cn^{-2+4d}(\sum_{j=1}^{n}|b_{nj}|)^4 \\ &\le Cn^{-2+4d}B_n^4 \le CB_n^{2+4d}(B_n/n)^{2-4d} = o(B_n^{2+4d}). \end{aligned}$$

Finally, consider the case $d = 1/4$. By (4.22),

$$\begin{aligned} t_n^2 &\le Cn^{-2} \sum_{j_1,k_1,j_2,k_2=1}^{n} \big|d_n(j_1-k_1)d_n(j_2-k_2)\big|\big|\gamma_X(j_1-j_2)\gamma_X(k_1-k_2)\big| \\ &\le Cn^{-1} \sum_{s_1,s_2,s_3=-n}^{n} \big|d_n(s_1)d_n(s_2)\big|\big|\gamma_X(s_3)\gamma_X(-s_1-s_2-s_3)\big|. \end{aligned} \tag{4.28}$$

Recall the following inequality, see e.g., Lemma 4.2 in Giraitis and Taqqu (1998). Let $p_1 \ge 1, \cdots, p_k \ge 1$ be real numbers such that $p_1^{-1} + \cdots + p_k^{-1} = k-1$, $k \ge 2$. Then, for any $b_m(j)$'s,

$$\sum_{s_1,\cdots,s_{k-1}\in\mathbb{Z}} \big|b_1(s_1)\cdots b_{k-1}(s_{k-1})b_k(s_1+\cdots+s_{k-1})\big| \le \big(\sum_{s\in\mathbb{Z}}|b_1(s)|^{p_1}\big)^{1/p_1}\cdots\big(\sum_{s\in\mathbb{Z}}|b_k(s)|^{p_k}\big)^{1/p_k}$$

assuming that the r.h.s. is finite. Applying this inequality in (4.28) with $k=4$, $p_1 = \cdots = p_4 = 4/3$, gives

$$t_n^2 \le Cn^{-1}\big(\sum_{|s|\le n}|d_n(s)|^{4/3}\big)^{3/2}\big(\sum_{|s|\le 3n}|\gamma_X(s)|^{4/3}\big)^{3/2}.$$

Since $\sum_{s=1}^n |\gamma_X(s)|^{4/3} \le C\sum_{s=1}^n s^{-2/3} \le Cn^{1/3}$, and by (4.3) and (2.15), $\sum_{|s|\le n}|d_n(s)|^{4/3} \le B_n^{1/3}\sum_{|s|\le n}|d_n(s)| \le B_n^{1/3+2}$, we obtain

$$t_n^2 \le Cn^{-1}\{n^{1/3}B_n^{1/3+2}\}^{3/2} = CB_n^3(B_n/n)^{1/2} = o(B_n^3) = o(B_n^{2(1+2d)}),$$

because, $1+2d = 3/2$, which completes the proof of $t_n = o(B_n^{1+2d})$.

The bound $s_{n,1} = o(B_n^{1+2d})$ follows by the same argument as in (i), using (4.26).

Finally, by (4.25) and (2.15), $s_{n,2} \le Cn^{-1}(\sum_{j=1}^n |b_{nj}|)^2 \le Cn^{-1}B_n^2 = o(B_n) = o(B_n^{1+2d})$, which completes the proof of (4.27) and of the part (ii) of the lemma. □

The next auxiliary lemma derives the upper bound for $E|Q_n - EQ_n|$ for a general quadratic form $Q_n = \sum_{t,s=1}^n c_{n,ts}X_tX_s$, where $\{X_j\}$ is a linear process (1.5) and $c_{n,jk}$ are real numbers. Setting $a_k = 0$, $k<0$, write $X_j = \sum_{k\in\mathbb{Z}} a_{j-k}\zeta_k$, and

$$Q_n = \sum_{j,k\in\mathbb{Z}} \beta_{n,jk}\zeta_j\zeta_k, \qquad \beta_{n,jk} := \sum_{t,s=1}^{n} c_{n,ts}a_{t-j}a_{s-k}.$$

It is an extension of Lemma 4.5.2 of Giraitis *et al.* (2012), where ζ_j's are assumed to be i.i.d., to m.d.s. ζ_j's.

Lemma 4.5 *Let $\{\zeta_j\}$ be a stationary and ergodic m.d.s. with $E\zeta_1^4 < \infty$. Then*

$$
\begin{aligned}
(4.29) \qquad E|Q_n - EQ_n| &= O(t_n) + o(\alpha_n), \\
t_n &:= \sum_{j_1,k_1,j_2,k_2=1}^{n} c_{n,j_1k_1} c_{n,j_2k_2} \gamma_X(j_1 - j_2)\gamma_X(k_1 - k_2),
\end{aligned}
$$

for any sequence $0 < \alpha_n$ such that

$$(4.30) \qquad \sum_{j\in\mathbb{Z}} |\beta_{n,jj}| = O(\alpha_n), \quad \sum_{j\in\mathbb{Z}} |\beta_{n,jj} - \beta_{n,j-1,j-1}| = o(\alpha_n).$$

Proof. Denote $Z_j' := \sum_{k=-\infty}^{j-1} \beta_{n,jk}\zeta_k$ and $Z_j'' := \sum_{k=-\infty}^{j-1} \beta_{n,kj}\zeta_k$, $j \in \mathbb{Z}$. Then

$$(4.31) \qquad Q_n - EQ_n = \sum_{j\in\mathbb{Z}} (Z_j' + Z_j'')\zeta_j + \sum_{j\in\mathbb{Z}} \beta_{n,jj}(\zeta_j^2 - E\zeta_j^2) =: q_{n1} + q_{n2}.$$

To estimate q_{n1}, note that $(Z_j' + Z_j'')\zeta_j$ are uncorrelated zero mean r.v.'s., since $\{\zeta_k\}$ is m.d.s., and $E\zeta_j^4 = E\zeta_1^4 < \infty$. Therefore,

$$E[(Z_j' + Z_j'')^2\zeta_j^2] \le 2\{E(Z_j')^4 + E(Z_j'')^4\}^{1/2}(E\zeta_j^4)^{1/2}.$$

By (4.36), $E(Z_j')^4 \le C(\sum_{k=-\infty}^{j-1} \beta_{n,jk}^2)^2$ and $E(Z_j'')^4 \le C(\sum_{k=-\infty}^{j-1} \beta_{n,kj}^2)^2$. Hence,

$$
\begin{aligned}
(4.32) \qquad Eq_{n1}^2 = \sum_{j\in\mathbb{Z}} E[(Z_j' + Z_j'')^2\zeta_j^2] &\le C \sum_{j\in\mathbb{Z}} \{(\sum_{k=-\infty}^{j-1} \beta_{n,jk}^2)^2 + (\sum_{k=-\infty}^{j-1} \beta_{n,kj}^2)^2\}^{1/2} \\
&\le C \sum_{j,k\in\mathbb{Z}} \beta_{n,jk}^2 = Ct_n^2,
\end{aligned}
$$

where the last equality follows by definition of $\beta_{n,jk}$ and t_n.

Since q_{n2} is a weighted sum of the process $\zeta_j^2 - E\zeta_j^2$, which is stationary and ergodic, by (4.30) and Lemma 4.7(i), $E\big|q_{n2} - E[q_{n2}]\big| = o(\alpha_n)$. Since $E[q_{n2}] = 0$, this together with (4.31) and (4.32) implies (4.29). □

Lemma 4.6 *Let $\bar{b}_{nj} = b_{nj}h_{nj}$, $j = 1, \cdots, n$, where h_{nj}, $j = 1, \cdots, n$ satisfy (2.20), and $B_{nh} \ge cB_n$, $n \to \infty$ for some $c > 0$.*

(i) *If b_{nj} satisfy (2.2) and (2.3), then $\bar{b}_{nj}$ satisfy (2.2) and (2.3).*

(ii) *If b_{nj} satisfy (2.15) and (2.16), then $\bar{b}_{nj}$ satisfy (2.15) and (2.16).*

Proof. By (2.20), $\max_{1\le j\le n}|h_{nj}| = O(1)$. Therefore, $B_{nh} = O(B_n)$ and $B_{nh} \asymp B_n$.

(i) To verify (2.3) for $\bar{b}_{nk}$, notice that

$$\sum_{k=1}^{n} |\bar{b}_{nk}| \le C\sum_{k=1}^{n} |b_{nk}| = o\Big(n^{1/2}B_n^{1/2}\Big) = o\Big(n^{1/2}B_{nh}^{1/2}\Big).$$

To verify (2.2) for $\bar{b}_{nk}$, note that (2.2) implies $\max_{1\le j\le n}|b_{nj}| = o(B_n^{1/2})$. Use (2.2), (2.20) and $|\bar{b}_{nk} - \bar{b}_{n,k-1}| \le |b_{nk} - b_{n,k-1}|\,|h_{nk}| + |b_{n,k-1}|\,|h_{nk} - h_{n,k-1}|$, to obtain

$$\begin{aligned} (4.33)\quad |\bar{b}_{n1}| + \sum_{k=2}^{n} |\bar{b}_{nk} - \bar{b}_{n,k-1}| \;&\le\; C\{|b_{n1}| + \sum_{k=2}^{n} |b_{nk} - b_{n,k-1}|\} \\ &+\; \max_j |b_{nj}| \sum_{k=2}^{n} |h_{nk} - h_{n,k-1}| = o(B_n^{1/2}) = o\big(B_{nh}^{1/2}\big). \end{aligned}$$

(ii) By (2.15), $\sum_{k=1}^{n} |b_{nk}| = O(B_n)$, $B_n \to \infty$, $B_n = o(n)$. Therefore, $B_{nh} \to \infty$, $B_{nh} \le CB_n = o(n)$ and $\sum_{k=1}^{n} |\bar{b}_{nk}| \le C\sum_{k=1}^{n} |b_{nk}| = O(B_n) = O(B_{nh})$, which proves (2.15) for $\bar{b}_{nk}$.

To show (2.16) for $\bar{b}_{nk}$, note that (4.33) together with (2.16) and (2.20) implies $|\bar{b}_{n1}| + \sum_{k=2}^{n} |\bar{b}_{nk} - \bar{b}_{n,k-1}| = O(1)$. □

Part (a) of the next lemma provides a useful bound for weighted sums of a stationary ergodic process $\{V_j\}$. Part (b) provides a bound for the moments of the weighted sum of m.d.s.

Lemma 4.7 (i) *Let* $T_n = \sum_{j\in\mathbb{Z}} c_{nj}V_j$, *where* $\{V_j\}$ *is a stationary ergodic sequence,* $E|Y_1| < \infty$, *and* c_{nj} *are real numbers such that for some* $0 < \alpha_n < \infty$, $n \ge 1$,

$$(4.34)\qquad \sum_{j\in\mathbb{Z}} |c_{nj}| = O(\alpha_n), \quad \sum_{j\in\mathbb{Z}} |c_{nj} - c_{n,j-1}| = o(\alpha_n).$$

Then

$$(4.35)\qquad E|T_n - ET_n| = o(\alpha_n).$$

In particular, if $\alpha_n = 1$, *then* $T_n = ET_n + o_p(1)$.

(ii) *If m.d.s.* $\{\xi_j\}$ *satisfies* $\max_j E\xi_j^p < \infty$, *for some* $p \ge 2$, *then*

$$(4.36)\qquad E\big(\sum_{j\in\mathbb{Z}} d_j\xi_j\big)^p \le C\big(\sum_{j\in\mathbb{Z}} d_j^2\big)^{p/2},$$

for any d_j*'s such that* $\sum_{j\in\mathbb{Z}} d_j^2 < \infty$, *where* $C < \infty$ *does not depend on* d_j*'s.*

Proof. (i) Without loss of generality, it suffices to prove (4.35) in case $EV_1 = 0$. Then $T_n - ET_n = T_n = \sum_{j\in\mathbb{Z}} c_{nj}V_j$. Letting $k \geq 1$, write

$$T_n = \sum_{j\in\mathbb{Z}} c_{nj}\{V_j - k^{-1}\sum_{l=1}^{k} V_{j+l}\} + \sum_{j\in\mathbb{Z}} c_{nj}\{k^{-1}\sum_{l=1}^{k} V_{j+l}\} =: q_{n1} + q_{n2}.$$

Then $E|T_n| \leq E|q_{n1}| + E|q_{n2}|$, where

$$\begin{aligned} E|q_{n1}| &\leq k^{-1}E\Big|\sum_{j\in\mathbb{Z}} c_{nj}kV_j - \sum_{l=1}^{k}\sum_{j\in\mathbb{Z}} c_{n,j-l}V_j\Big| \\ &\leq k^{-1}E|V_1|\sum_{j\in\mathbb{Z}} |kc_{nj} - c_{n,j-1} - \cdots - c_{n,j-k}| \\ &\leq kE|V_1|\sum_{j\in\mathbb{Z}} |c_{nj} - c_{n,j-1}| = o(\alpha_n), \qquad \forall k \text{ fixed}, \end{aligned}$$

by (4.34), and because $|kc_{nj} - c_{n,j-1} - \cdots - c_{n,j-k}| \leq |c_{nj} - c_{n,j-k}| + |c_{n,j} - c_{n,j-k+1}| + \cdots + |c_{n,j} - c_{n,j-1}| \leq k\{|c_{nj} - c_{n,j-1}| + |c_{n,j-1} - c_{n,j-2}| + \cdots + |c_{n,j-k+1} - c_{n,j-k}|\}$.

It remains to bound q_{n2}. Since $\{V_j\}$ is stationary and ergodic with $EV_1 = 0$, by the Ergodic Theorem, see e.g. Stout (1974, Corollary 3.5.2), $E\big|k^{-1}\sum_{l=1}^{k} V_{j+l}\big| = E\big|k^{-1}\sum_{l=1}^{k} V_l\big| = \delta_k \to 0$, $k \to \infty$. Thus, by (4.34), $E|q_{n2}| \leq \delta_k \sum_{j\in\mathbb{Z}} |c_{nj}| \leq C\delta_k\alpha_n$, which implies $E|T_n| = o(\alpha_n)$ and completes the proof of (4.35) and (i).

(ii) The bound (4.36) is shown in Lemma 3.2 in Abadir *et al.* (2013) $\square$

5 Monte-Carlo simulation

In the following Monte-Carlo simulation we study the finite-sample performance of the above proposed studentization, for the sample size $n = 512$ using 1,000 replications.

In Tables 1 and 2 we evaluate the fit of the normal approximation to studentized sum $t_{n,X} = S_{n,X}/\widetilde{v}_{n,X}$, $S_{n,X} = \sum_{j=1}^{n} b_{nj}X_j$. We use two sets of standartized weights $b_{nj} = k_{nj}/\sum_{l=1}^{n} k_{nl}$, $j = 1, \cdots, n$, where $k_{nj} = K(j/nw)$ is generated by Epanechnikov kernel $K(x) = (3/4\sqrt{5})(1 - x^2/\sqrt{5})I(|x| \leq \sqrt{5})$ and by flat kernel $K(x) = I(|x| \leq 1)$. The bandwidth parameter w is set to take values $H = nw = n, n^{0.9}, \cdots, n^{0.2}$. The weights are normalized to sum up to 1 for comparison purposes. For $nw \to \infty$, $w = o(1)$, the weights k_{nj} satisfy assumptions (2.2), (2.3), (2.15) and (2.16), but do not satisfy some of them when $w = 1$.

Tables 1 and 2 report the average value of the estimator $\widetilde{v}_{n,X}$ of the SD $v_{n,X}$ of $S_{n,X}$, and the probability $P(|t_{n,X}| > 1.96)$. In Table 1, $X_j = rX_{j-1} + \zeta_j$ follows AR(1) model with $r = 0, 0.5, 0.9$ and ARCH(1) m.d. noise $\zeta_j = \varepsilon_j\sigma_j$, $\sigma_j^2 = 1 + a\zeta_{j-1}^2$, with parameter $a = 0.7$, where ε_j is a standard normal i.i.d. noise. In Table 2, X_j is simulated using ARFIMA(1,d,0) model with AR parameter $r = 0, 0.5$, long memory parameter $d = 0.3$ and i.i.d noise. It is evident that in Tables 1 and 2, in all cases, the average of $\widetilde{v}_{n,X}$ is very close to $v_{n,X}$. Moreover, the probability $P(|t_{n,X}| > 1.96)$ is close to 0.05 for $H = nw = n^{0.6}, \cdots, n^{0.2}$ when using the Epanechnikov kernel and for $H = nw = n^{0.7}, \cdots, n^{0.2}$ in the case of flat kernel, indicating good fit of normal approximation. The findings lend evidence to the applicability of the asymptotic results for finite samples of moderate sizes.

The averages $\widetilde{v}_{n,X}$ and the SD $v_{n,X}$ of $S_{n,X}$ are higher for the flat weights and tend to increase when the correlation in the data becomes stronger, for both types of weights considered. For the models in Table 1, we had performed simulations with iid noise, and the results were very similar to those with the ARCH(1) noise presented here.

Table 3 reports simulations in the case of heteroscedastic data. It presents the results on the fit of normal approximation to $t_{n,Y} = S_{n,Y}/\widetilde{v}_{n,Y}$, $S_{n,Y} = \sum_{j=1}^{n} b_{nj} Y_j$, where $Y_j = \sigma_j X_j$ and $\sigma_j^2 = 0.5(\sin(4\pi(j/n)) + 2)$, for some of the X_j models. Theoretical results imply that $v_{n,Y}^2 \sim \theta^2 v_{n,X}^2$ with $\theta = 1$. The average $\widetilde{v}_{n,Y}$ is close to $v_{n,X}$ and $v_{n,Y}$, as predicted by theory, but slightly closer to $v_{n,X}$, and therefore the probability $P(|t_{n,Y}| > 1.96)$ is higher than in the homoscedastic cases. This is somehow expected, taking into account the heteroscedasticity. Here, we only considered the Epanechnikov kernel.

Table 4 reports simulation results for the self-normalized sum $t_n = \sum_{j=1}^{n} \varepsilon_j/\tilde{v}_n$, $\tilde{v}_n^2 = \sum_{j=1}^{n} \varepsilon_j^2$, $\varepsilon_j = d_{nj}\zeta_j$, where the weights d_{nj} are the Epanechnikov and flat weights as above, and ζ_j is the ARCH(1) noise with parameter $a = 0.7$. It is evident that for all nw, the average $\widetilde{v}_n$ is very close to v_n and the probability $P(|t_n| > 1.96)$ is close to 0.05 for the bandwidths $nw = n, n^{0.9}, \cdots, n^{0.3}$, as suggested by Proposition 2.2.

Overall, the Monte-Carlo simulations justify the use of $\widetilde{v}_{n,X}$ for estimating the SD $v_{n,X}$ of $S_{n,X}$. They show that for moderate bandwidth parameters, the normalized versions of $S_{n,X}$ give approximately the right probability when using critical values from the standard normal distribution.

Model: $X_j \sim \zeta_j$ (m.d.s.)						
	Epanechnikov			Flat		
$H = nw$	average $\widetilde{v}_{n,X}$	$v_{n,X}$	$P(\lvert t_{n,X}\rvert > 1.96)$	average $\widetilde{v}_{n,X}$	$v_{n,X}$	$P(\lvert t_{n,X}\rvert > 1.96)$
n	0.040	0.044	0.000	0.040	0.043	0.000
$n^{0.9}$	0.041	0.045	0.000	0.056	0.060	0.006
$n^{0.8}$	0.056	0.061	0.000	0.079	0.086	0.029
$n^{0.7}$	0.080	0.085	0.028	0.110	0.114	0.045
$n^{0.6}$	0.111	0.115	0.043	0.151	0.156	0.052
$n^{0.5}$	0.154	0.160	0.051	0.210	0.221	0.055
$n^{0.4}$	0.211	0.222	0.060	0.286	0.292	0.052
$n^{0.3}$	0.303	0.311	0.051	0.407	0.415	0.060
$n^{0.2}$	0.440	0.467	0.056	0.577	0.618	0.057
Model: $X_j \sim$AR(1), $r = 0.5$						
n	0.068	0.075	0.000	0.068	0.074	0.000
$n^{0.9}$	0.071	0.078	0.000	0.097	0.103	0.007
$n^{0.8}$	0.097	0.105	0.000	0.136	0.146	0.027
$n^{0.7}$	0.137	0.146	0.026	0.189	0.193	0.047
$n^{0.6}$	0.189	0.194	0.041	0.257	0.261	0.052
$n^{0.5}$	0.262	0.269	0.051	0.351	0.366	0.058
$n^{0.4}$	0.353	0.370	0.055	0.465	0.475	0.061
$n^{0.3}$	0.490	0.505	0.057	0.620	0.646	0.053
$n^{0.2}$	0.656	0.701	0.055	0.779	0.834	0.060
Model: $X_j \sim$AR(1), $r = 0.9$						
n	0.174	0.193	0.000	0.173	0.192	0.000
$n^{0.9}$	0.180	0.201	0.000	0.243	0.261	0.005
$n^{0.8}$	0.244	0.266	0.000	0.336	0.360	0.026
$n^{0.7}$	0.340	0.360	0.032	0.453	0.463	0.045
$n^{0.6}$	0.456	0.468	0.052	0.585	0.599	0.062
$n^{0.5}$	0.595	0.615	0.061	0.724	0.755	0.065
$n^{0.4}$	0.726	0.758	0.059	0.831	0.865	0.060
$n^{0.3}$	0.847	0.888	0.063	0.915	0.971	0.067
$n^{0.2}$	0.928	0.989	0.061	0.968	1.030	0.066

Table 1: Monte-Carlo average of $\widetilde{v}_{n,X}$, standard deviation $v_{n,X}$ of $S_{n,X}$ and $P(|t_{n,X}| > 1.96)$ with $t_n = S_{n,X}/\widetilde{v}_{n,X}$ for AR(1) models $r = 0, 0.5, 0.9$ with ARCH(1) $a = 0.7$ noise.

Model: $X_j \sim$ARFIMA(0,d,0), $d = 0.3$						
	Epanechnikov			Flat		
$H = nw$	average $\widetilde{v}_{n,X}$	$v_{n,X}$	$P(\lvert t_{n,X}\rvert > 1.96)$	average $\widetilde{v}_{n,X}$	$v_{n,X}$	$P(\lvert t_{n,X}\rvert > 1.96)$
n	0.215	0.285	0.000	0.214	0.285	0.000
$n^{0.9}$	0.221	0.288	0.000	0.269	0.318	0.004
$n^{0.8}$	0.270	0.319	0.000	0.327	0.367	0.030
$n^{0.7}$	0.329	0.368	0.030	0.387	0.411	0.047
$n^{0.6}$	0.388	0.414	0.048	0.449	0.466	0.044
$n^{0.5}$	0.453	0.474	0.049	0.519	0.541	0.057
$n^{0.4}$	0.520	0.545	0.060	0.592	0.613	0.051
$n^{0.3}$	0.607	0.631	0.046	0.687	0.711	0.048
$n^{0.2}$	0.710	0.738	0.046	0.798	0.822	0.047
Model: $X_j \sim$ARFIMA(1,d,0), $r = 0.5$, $d = 0.3$						
n	0.293	0.390	0.000	0.292	0.390	0.000
$n^{0.9}$	0.302	0.395	0.000	0.367	0.434	0.004
$n^{0.8}$	0.368	0.436	0.000	0.445	0.499	0.029
$n^{0.7}$	0.447	0.501	0.031	0.526	0.559	0.046
$n^{0.6}$	0.527	0.563	0.046	0.609	0.633	0.048
$n^{0.5}$	0.615	0.644	0.050	0.701	0.732	0.061
$n^{0.4}$	0.703	0.737	0.059	0.792	0.825	0.055
$n^{0.3}$	0.809	0.846	0.053	0.893	0.931	0.052
$n^{0.2}$	0.913	0.953	0.062	0.977	1.011	0.054

Table 2: Monte-Carlo average of $\widetilde{v}_{n,X}$, standard deviation $v_{n,X}$ of $S_{n,X}$ and $P(|t_{n,X}| > 1.96)$ with $t_n = S_{n,X}/\widetilde{v}_{n,X}$ for ARFIMA(1,d,0) models $d = 0.3$ and $r = 0, 0.5$ with i.i.d. noise.

	Model: X_j ~AR(1), $r = 0.5$				Model: X_j ~AR(1), $r = 0.9$			
$H = nw$	average $\widetilde{v}_{n,Y}$	$v_{n,X}$	$v_{n,Y}$	$P(\lvert t_{n,Y}\rvert > 1.96)$	average $\widetilde{v}_{n,Y}$	$v_{n,X}$	$v_{n,Y}$	$P(\lvert t_{n,Y}\rvert > 1.96)$
n	0.069	0.075	0.076	0.000	0.169	0.193	0.195	0.000
$n^{0.9}$	0.072	0.078	0.081	0.000	0.176	0.201	0.207	0.000
$n^{0.8}$	0.098	0.105	0.112	0.000	0.240	0.266	0.282	0.001
$n^{0.7}$	0.139	0.148	0.169	0.050	0.335	0.368	0.420	0.067
$n^{0.6}$	0.191	0.196	0.222	0.072	0.450	0.476	0.543	0.098
$n^{0.5}$	0.265	0.270	0.293	0.068	0.588	0.626	0.683	0.088
$n^{0.4}$	0.357	0.371	0.390	0.065	0.717	0.771	0.815	0.093
$n^{0.3}$	0.493	0.512	0.527	0.059	0.836	0.895	0.923	0.071
$n^{0.2}$	0.658	0.681	0.692	0.056	0.915	0.971	0.988	0.070
	Model: X_j ~ARFIMA(0,d,0), $d = 0.3$				Model: X_j ~ARFIMA(1,d,0), $r = 0.5$, $d = 0.3$			
n	0.200	0.285	0.283	0.000	0.262	0.390	0.387	0.001
$n^{0.9}$	0.206	0.288	0.290	0.000	0.270	0.395	0.398	0.003
$n^{0.8}$	0.253	0.319	0.330	0.002	0.332	0.436	0.451	0.008
$n^{0.7}$	0.310	0.368	0.417	0.086	0.407	0.501	0.568	0.112
$n^{0.6}$	0.368	0.414	0.474	0.106	0.484	0.563	0.645	0.122
$n^{0.5}$	0.432	0.474	0.519	0.092	0.568	0.644	0.705	0.104
$n^{0.4}$	0.498	0.545	0.575	0.087	0.653	0.737	0.779	0.103
$n^{0.3}$	0.582	0.631	0.650	0.068	0.754	0.846	0.871	0.087
$n^{0.2}$	0.682	0.738	0.750	0.058	0.852	0.953	0.969	0.076

Table 3: Monte-Carlo average of $\widetilde{v}_{n,Y}$, standard deviation $v_{n,X}$ of $S_{n,X}$, standard deviation $v_{n,Y}$ of $S_{n,Y}$ and $P(|t_{n,Y}| > 1.96)$ with $t_{n,Y} = S_{n,Y}/\widetilde{v}_{n,Y}$ for $Y_j = \sigma_j X_j$ with X_j following the AR(1) $r = 0.5, 0.9$ model, ARFIMA(1,d,0) model $d = 0.3$ and $r = 0, 0.5$ with i.i.d. noise and Epanechnikov weights.

	Epanechnikov			Flat		
$H = nw$	average $\widetilde{v}_n$	v_n	$P(\lvert t_n\rvert > 1.96)$	average $\widetilde{v}_n$	v_n	$P(\lvert t_n\rvert > 1.96)$
n	0.044	0.043	0.047	0.044	0.043	0.047
$n^{0.9}$	0.046	0.045	0.047	0.060	0.060	0.042
$n^{0.8}$	0.060	0.061	0.044	0.082	0.086	0.062
$n^{0.7}$	0.082	0.085	0.052	0.110	0.114	0.054
$n^{0.6}$	0.109	0.114	0.041	0.147	0.156	0.043
$n^{0.5}$	0.148	0.160	0.052	0.196	0.221	0.047
$n^{0.4}$	0.196	0.222	0.050	0.259	0.292	0.040
$n^{0.3}$	0.272	0.311	0.041	0.355	0.415	0.031
$n^{0.2}$	0.376	0.467	0.025	0.478	0.618	0.000

Table 4: Monte-Carlo average of $\widetilde{v}_n$, standard deviation v_n of S_n and $P(|t_n| > 1.96)$ with $t_n = S_n/\widetilde{v}_n$ for ARCH(1) $a = 0.7$ model.

References

Abadir, K.M., Distaso, W., and Giraitis, L. (2009). Two estimators of the long-run variance: beyond short memory. *J. of Econometrics*, **150**, 56–70.

Abadir, K.M., Distaso, W., Giraitis, L. and Koul, H.L. (2013). Asymptotic normality for weighted sums of linear processes, *Econometric Theory*, forthcoming.

Giraitis, L. and Taqqu, M.S. (1998). Central limit theorem for quadratic forms with time-domain conditions. *Ann. Probability*, **26**, 377-398.

Giraitis, L., Koul, H.L. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes.* Imperial College Press, London.

Giraitis, L., Leipus, R., and Surgailis, D. (2007). Recent advances in ARCH modelling. In: G. Teyssiere and A.P. Kirman, Editors, *Long Memory in Economics*, Springer, 3-38.

Guo, H. and Koul, H.L. (2007). Nonparametric regression with heteroscedastic long memory errors. *J. Statist. Plann. Inference*, **137**, 379–404.

Hall, P. (1992). On bootstrap confidence intervals in non-parametric regression. *Ann. Statist.*, **20**, 695–711.

Hall, P. and Heyde, C.C. (1980). *Martingale limit theory and applications.* Academic Press, New York.

McMurry, T.L. and Politis, D.N. (2008). Bootstrap confidence intervals in non-parametric regression with built-in bias correction. *Statist. Probab. Lett.*, **78**, 2463-2469.

Phillips, P.C.B., Sun, Y. and Jin, S. (2007). Long run variance estimation and robust regression testing using sharp origin kernels with no truncation. *J. Statist. Plann. Inference*, **137**, 985–1023.

Robinson, P.M. (1997). Large-sample inference for non-parametric regression with dependent errors. *Ann. Statist.*, **25**, 2054–2083.

Robinson, P.M. (1998). Inference-without-smoothing in the presence of non-parametric autocorrelation. *Econometrica*, **66**, 1163–1182.

Robinson, P.M. (2005). Robust covariance matrix estimation: HAC estimates with long memory/antipersistence correction. *Econometric Theory*, **21**, 171–180.

Stout, W. (1974). *Almost Sure Convergence.* Academic Press, New York.

Shao, Q.M. (1997). Self-normalized standard deviations. *Ann. Probability*, **25**, 285–328.

Zhao, Z. and Li, X. (2012) Inference for modulated stationary processes. *Bernoulli*, to appear.