

Estimation pitfalls when the noise is not i.i.d.

LIUDAS GIRAITIS

Queen Mary, University of London

MASANOBU TANIGUCHI

Waseda University

MURAD S. TAQQU

Boston University

March 11, 2018

Abstract

This paper extends Whittle estimation to linear processes with a general stationary ergodic martingale difference noise. We show that such estimation is valid for standard parametric time series models with smooth bounded spectral densities, e.g. ARMA models. Furthermore, we clarify the impact of the hidden dependence in the noise on such estimation. We show that although the asymptotic normality of the Whittle estimates may still hold, the presence of dependence in the noise impacts the limit variance. Hence, the standard errors and confidence intervals valid under i.i.d. noise may not be applicable and thus require correction. The goal of this paper is to raise awareness to the impact of a non i.i.d. noise in applied work.

AMS 2000 Subject classification 62E20, 60F05.

Keywords and phrases: Whittle estimation, Asymptotic normality, Quadratic form, Martingale difference noise.

1 Introduction

A variety of stationary models known in statistical and econometric literature can be expressed as a linear/moving average time series. Under minor restrictions a linear stationary time series can be represented, by the Wold decomposition, as a moving average

of infinite order,

$$(1.1) \quad X_t = \sum_{j=0}^{\infty} a_j \eta_{t-j}, \quad t \in \mathbb{Z},$$

where $\{\eta_t\}$ is an uncorrelated noise with zero mean and weights $\{a_j\}$ satisfy $\sum_{j=0}^{\infty} a_j^2 < \infty$. We have $a_j = a_j(\theta)$ and $E\eta_t^2 = \sigma^2$. The goal is to estimate θ and σ^2 .

In applications, the noise $\{\eta_t\}$ thus is often assumed to be i.i.d. or a stationary martingale differences (m.d.) sequence. The estimation of such models is mainly done under the assumption of i.i.d. noise but this may be too restrictive in applications and hard to verify in practice and it excludes ARCH type conditionally heteroskedastic noises $\{\eta_t\}$ which are stationary ergodic m.d. processes. A typical example of a linear process (1.1) with m.d. noise is $X_t = r_t^2 - Er_t^2$ where r_t is a GARCH(p, q) process. It is commonly used for modeling squared returns of assets in financial econometrics, see e.g. the review in Giraitis, Leipus and Surgailis (2007).

In empirical work, parameters of such time series are often estimated using techniques suitable for i.i.d. noises $\{\eta_t\}$ but without proper theoretical validation, see e.g. Wu and Shieh (2007).

In this paper, we examine the validity of a standard Whittle estimation procedure for a linear process (1.1) with a martingale difference noise $\{\eta_t\}$ and we analyze how this m.d. noise impacts the asymptotic behaviour of the estimates.

In his seminal work Hannan (1973) showed that parametric Whittle estimates $(\hat{\sigma}_n^2, \hat{\theta}_n)$ given in (2.7) are consistent estimators of the true value of the parameter (σ_0^2, θ_0) for a large class of ergodic time series $\{X_t\}$. He established the asymptotic normality of $\hat{\theta}_n$ for linear processes with smooth bounded spectral densities $f_{\sigma^2, \theta}$ for a special class of m.d. noises $\{\eta_t\}$. Hannan assumed that $E[\eta_t | \mathcal{F}_{t-1}] = \sigma^2$ is a constant a.s.

Fox and Taqqu (1986), Giraitis and Surgailis (1990), Giraitis, Hidalgo and Robinson (2001) and others extended the parametric Whittle estimation technique to linear long memory time series $\{X_t\}$ with unbounded spectral densities and i.i.d. noise $\{\eta_t\}$; see Chapter 8 in Giraitis, Koul and Surgailis (2012). Hosoya and Taniguchi (1982) showed that Whittle estimation remains valid for linear processes with uncorrelated noise $\{\eta_t\}$ whose fourth-order cumulants are summable and whose conditional moments satisfy some regularity conditions.

Our aim in this paper is to extend Whittle estimation to linear processes with a stationary ergodic m.d. noise $\{\eta_t\}$. We shall show that such estimation is valid for standard parametric time series models with smooth bounded spectral densities, e.g. ARMA models. Furthermore, we shall clarify the impact of the dependence structure of the noise in such estimation. The proof of the asymptotic normality relies on the normal approximation results for quadratic forms in stationary ergodic m.d. noise $\{\eta_t\}$ obtained in Giraitis,

Taniguchi and Taqqu (2016).

We show that differences in inference between modelling with i.i.d. noise and modelling with m.d. noise cannot be ignored. The goal of this paper is thus to raise awareness, when doing applied work, to the fact that although the asymptotic normality of the estimates may still hold, the limit variance might be affected. Hence, the standard errors and confidence intervals valid under i.i.d. noise may not be applicable and thus need to be corrected.

The main results are Theorem 2.1 (consistency), Theorem 2.2 (asymptotic normality) and Theorem 3.1 (asymptotic normality of quadratic forms). More generally, Section 2 examines the impact of m.d. noise on parametric Whittle estimation. We establish asymptotic normality for parametric Whittle estimator $\hat{\theta}_n$ under weak conditions on the noise. This requires deriving the asymptotic normality of quadratic forms for linear processes with m.d. noise and is done in Section 3. Section 4 contains auxiliary results. Section 5 deals with applications.

Throughout the paper, by \rightarrow_p and \rightarrow_D we denote convergence in probability and distribution, respectively, while C denotes generic constants.

2 Parametric Whittle estimation

Denote by $\{\eta_t\}$ a stationary ergodic martingale difference (m.d.) sequence with respect to the natural filtration \mathcal{F}_t defined below, namely $E[\eta_t|\mathcal{F}_{t-1}] = 0$, with moments

$$(2.2) \quad E\eta_t = 0, E\eta_t^2 = \sigma^2 \text{ and } E\eta_t^4 < \infty.$$

By \mathcal{F}_t we denote the σ -field generated by $(\eta_t, \eta_{t-1}, \dots)$ or, more generally, by some underlying noise $(\varepsilon_t, \varepsilon_{t-1}, \dots)$ such that $\eta_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots)$ is a measurable function of ε_t 's. Clearly, $E\eta_t\eta_s = 0$ for $t \neq s$. Indeed, if $s < t$, then $E\eta_t\eta_s = E[E[\eta_t|\mathcal{F}_{t-1}]\eta_s] = 0$.

In this section we study the parametric Whittle estimation for a *linear* process

$$(2.3) \quad X_t := \sum_{k=0}^{\infty} a_k(\theta)\eta_{t-k}, \quad a_0 = a_0(\theta) = 1, \quad t \in \mathbb{Z},$$

with

$$\sum_{k=0}^{\infty} a_k^2(\theta) < \infty, \quad \theta \in \Theta.$$

The real-valued coefficients $a_k(\theta)$, $k = 0, 1, \dots$ are parameterized by the parameter $\theta \in \Theta$ taking values in a compact set $\Theta \subset \mathbb{R}^q$. Throughout this paper, σ_0^2, θ_0 denote the true parameter values of σ^2, θ , respectively. In this paper, we prove asymptotic normality of

the Whittle estimator of θ_0 and consistency of the estimator of σ^2 .

The spectral density of the process $\{X_t\}$ has a parametric form

$$(2.4) \quad f(u) \equiv f_{\sigma^2, \theta}(u) = \frac{\sigma^2}{2\pi} s_\theta(u),$$

where

$$s_\theta(u) := \left| \sum_{k=0}^{\infty} a_k(\theta) e^{iku} \right|^2, \quad u \in \Pi, \quad \theta \in \Theta,$$

where $\Pi = [-\pi, \pi)$. By Wold decomposition, the class of stationary processes having linear representation (2.3) with an uncorrelated noise η_t is very large. In this paper the class of possible noises is reduced by supposing that they form stationary ergodic m.d. sequences.

Assume that observations X_1, X_2, \dots, X_n are from the linear process (2.3). Denote the periodogram based on X_1, X_2, \dots, X_n by

$$(2.5) \quad I_n(u) := \frac{1}{2\pi n} \left| \sum_{j=1}^n X_j e^{iju} \right|^2.$$

Parametric Whittle inference procedures involve the integrated weighted periodogram

$$(2.6) \quad Q_n(\theta) := \int_{\Pi} \frac{I_n(u)}{s_\theta(u)} du$$

with weight function $s_\theta(u)$. In view of (2.5), $Q_n(\theta)$ equals the quadratic form

$$Q_n(\theta) = \frac{1}{n} \sum_{j,k=1}^n b_{j-k}(\theta) X_j X_k,$$

where

$$b_j(\theta) := \frac{1}{2\pi} \int_{\Pi} \frac{e^{iju}}{s_\theta(u)} du, \quad j \in \mathbb{Z}.$$

Whittle estimates of σ_0, θ_0 based on X_1, X_2, \dots, X_n are defined as

$$(2.7) \quad \hat{\sigma}_n^2 = Q_n(\hat{\theta}_n), \quad \hat{\theta}_n = \operatorname{argmin}_{\theta \in \Theta} Q_n(\theta).$$

These estimators were introduced by Whittle (1953) and are obtained by minimizing the approximate Gaussian log-likelihood. The approximate Gaussian log-likelihood is known as "the Whittle Gaussian log-likelihood".

We shall first address the consistency. Consider the following assumption.

Assumption (a0) *The parameter space Θ is compact, parameter (σ^2, θ) determines the*

spectral density

$$f_{\sigma^2, \theta}(u) = \sigma^2 s_{\theta}^2(u) / 2\pi$$

uniquely. The function $s_{\theta}(u)$ is continuous in $(u, \theta) \in \Pi \times \Theta$ and for some $c_1 > 0, c_2 > 0$,

$$0 < c_1 \leq s_{\theta}(u) \leq c_2 < \infty, \quad (u, \theta) \in \Pi \times \Theta.$$

Thus the spectral density is bounded away from the origin and infinity.

THEOREM 2.1 (Consistency of Whittle estimators). *Suppose an observable linear process $\{X_t\}$ of (2.3) has the spectral density*

$$f(u) = \frac{\sigma_0^2}{2\pi} s_{\theta_0}(u),$$

and suppose the functions $s_{\theta}, \theta \in \Theta$, satisfy Assumption (a0). Then, as $n \rightarrow \infty$,

$$(2.8) \quad \hat{\theta}_n \rightarrow \theta_0, \quad \hat{\sigma}_n^2 \rightarrow \sigma_0^2, \quad a. s.$$

Proof. By assumption, $\{\eta_t\}$ is a stationary ergodic sequence. Thus, Theorem 3.5.8 in Stout (1974) implies that the sequence $\{X_t\}$ in (2.3) is also stationary ergodic. Hence (2.8) follows from Theorem 1 in Hannan (1973). For more details of the proof, see Theorem 8.2.1 in Giraitis *et al.* (2012). \square

In general, the asymptotic normality of the Whittle estimates requires stronger modelling assumptions on $\{X_t\}$ in (2.3). We introduce the following conditions on the functions $s_{\theta}(u)$ and the weights $a_k(\theta_0)$ of $\{X_t\}$.

Denote by ∇_{θ} the partial derivative operator with respect to a vector θ and by $'$ the transposition operator. Set

$$(2.9) \quad \begin{aligned} W_{\theta_0} &= \int_{\Pi} \nabla_{\theta} \log s_{\theta_0}(u) \nabla'_{\theta} \log s_{\theta_0}(u) du, \\ V_{\theta_0, \eta} &= 4E \left[\left(\sum_{k=1}^{\infty} \beta_{k, \theta_0} \eta_{-k} \right) \left(\sum_{k=1}^{\infty} \beta_{k, \theta_0} \eta_{-k} \right)' \eta_0^2 \right], \\ \beta_{k, \theta_0} &:= (2\pi)^{-1} \int_{\Pi} e^{iku} \nabla_{\theta} \log s_{\theta_0}(u) du, \quad k \in \mathbb{Z}. \end{aligned}$$

Noting that $\beta_{k, \theta_0} = \beta_{-k, \theta_0}$ and $\beta_{0, \theta_0} = 0$, Parseval's identity implies

$$(2.10) \quad W_{\theta_0} = 2\pi \sum_{k \in \mathbb{Z}} \beta_{k, \theta_0} \beta'_{k, \theta_0} = 4\pi \sum_{k=1}^{\infty} \beta_{k, \theta_0} \beta'_{k, \theta_0}.$$

Assumption (a1) (i) *The true value of parameter (σ_0^2, θ_0) lies in the interior $(0, \infty) \times \Theta$.*

The weights $a_k(\theta_0)$ in (2.3) have property $\sum_{k=0}^{\infty} ka_k^2(\theta_0) < \infty$.

(ii) The partial derivatives $\nabla_{\theta}s_{\theta}(u), \nabla_u s_{\theta}(u), \nabla_u \nabla_{\theta}s_{\theta}(u), \nabla_{\theta} \nabla'_{\theta}s_{\theta}(u)$, exist and are bounded and continuous functions of $u \in \Pi$ and $\theta \in \Theta$.

(iii) The matrix W_{θ_0} is positive definite.

The following theorem establishes the asymptotic normality of the Whittle estimate $\hat{\theta}_n$. Remarkably, besides (2.2), no additional conditions on the stationary ergodic martingale difference noise $\{\eta_t\}$ are needed.

Hidden dependence of the noise variables η_t , however, will have an impact on the asymptotic variance matrix of the Whittle estimate $\hat{\theta}_n$ in (2.11). By *hidden* dependence, we mean for example, a situation where the η_t may be uncorrelated but their square η_t^2 are correlated. The matrix has the form

$$\Omega_{\theta_0, \eta} = 4\pi^2 \sigma_0^{-4} W_{\theta_0}^{-1} V_{\theta_0, \eta} W_{\theta_0}^{-1}.$$

We show that for i.i.d. noise $\{\eta_j\}$, this matrix reverts to the standard asymptotic variance matrix of the Whittle estimate $\Omega_{\theta_0} = 4\pi W_{\theta_0}^{-1}$ given in Theorem 2 in Hannan (1973), which does not depend on $\{\eta_t\}$.

THEOREM 2.2 (Asymptotic normality of Whittle estimators). *Let $\{X_t\}$ be a linear process (2.3) having parametric spectral density $f = \sigma_0^2 s_{\theta_0}/2\pi$, with s_{θ} satisfying Assumptions (a0) and (a1). Then*

$$(2.11) \quad n^{1/2}(\hat{\theta}_n - \theta_0) \rightarrow_D \mathcal{N}(0, \Omega_{\theta_0, \eta}), \quad \Omega_{\theta_0, \eta} = 4\pi^2 \sigma_0^{-4} W_{\theta_0}^{-1} V_{\theta_0, \eta} W_{\theta_0}^{-1}.$$

Moreover,

(i) If $\{\eta_t\}$ are i.i.d. random variables, then

$$(2.12) \quad \Omega_{\theta_0, \eta} = 4\pi W_{\theta_0}^{-1}, \quad V_{\theta_0, \eta} = \pi^{-1} \sigma_0^4 W_{\theta_0},$$

and

$$n^{1/2}(\hat{\theta}_n - \theta_0) \rightarrow_D \mathcal{N}(0, \Omega_{\theta_0}).$$

(ii) If m.d. noise $\{\eta_t\}$ is such that $E[\eta_0^2 \eta_k \eta_s] = 0$ for any $s < k < 0$, then

$$(2.13) \quad \begin{aligned} \Omega_{\theta_0, \eta} &= 4\pi W_{\theta_0}^{-1} + \Omega_{\theta_0, \eta}^*, & \Omega_{\theta_0, \eta}^* &:= 4\pi^2 \sigma_0^{-4} W_{\theta_0}^{-1} \Delta_{\eta} W_{\theta_0}^{-1}, \\ V_{\theta_0, \eta} &= \pi^{-1} \sigma_0^4 W_{\theta_0} + \Delta_{\eta}, & \Delta_{\eta} &:= 4 \sum_{k=1}^{\infty} \beta_{k, \theta_0} \beta'_{k, \theta_0} \text{cov}(\eta_0^2, \eta_{-k}^2). \end{aligned}$$

REMARK 2.1 Obviously (2.13) reduces to (2.12) if the η_t are i.i.d. Whittle estimation is robust with respect to an i.i.d. noise η_t in the sense that its variance (2.12) does not depend on η_t . In particular, Whittle estimation does not have pitfalls in the sense that

(2.11) holds with (2.12) when the m.d. noise η_t is Gaussian since such η_t 's are i.i.d. Theorem 2.2 shows that this property is not valid anymore for m.d. noises η_t which are not i.i.d. Although asymptotic normality may still hold, the standard errors of the Whittle estimates are affected by the presence of the hidden dependence in the noise.

EXAMPLE 2.1 Consider, for example, the m.d. noise

$$(2.14) \quad \eta_t = \varepsilon_t \varepsilon_{t-1}$$

where $\{\varepsilon_t\}$ are i.i.d. Gaussian random variables with zero mean and variance $E\varepsilon_0^2 = \sigma_\varepsilon^2$. Then $\{\eta_t\}$ is a stationary ergodic m.d. sequence with respect to sigma field \mathcal{F}_t generated by variables $\varepsilon_t, \varepsilon_{t-1}, \dots$ with

$$\eta_t = 0, \quad \sigma_0^2 = E\eta_0^2 = (E\varepsilon_0^2)^2 \quad \text{and} \quad E[\eta_0^2 \eta_k \eta_s] = 0$$

for $s < k < 0$. Moreover,

$$\text{cov}(\eta_0^2, \eta_{-1}^2) = E[\eta_0^2 \eta_{-1}^2] - E[\eta_0^2]E[\eta_{-1}^2] = 2\sigma_0^4$$

and

$$\text{cov}(\eta_0^2, \eta_{-k}^2) = 0, \quad \text{for } k \geq 2.$$

Hence, (2.13) holds with $\Delta_\eta := 8\sigma_0^4 \beta_{1,\theta_0} \beta'_{1,\theta_0}$,

$$\Omega_{\theta_0, \eta} = 4\pi W_{\theta_0}^{-1} + \Omega_{\theta_0, \eta}^*, \quad \Omega_{\theta_0, \eta}^* := 32\pi^2 W_{\theta_0}^{-1} \beta_{1,\theta_0} \beta'_{1,\theta_0} W_{\theta_0}^{-1}.$$

Here, the dependence in m.d. noise contributed an additional term $\Omega_{\theta_0, \eta}^*$ to the variance $\Omega_{\theta_0, \eta}$ compared with an i.i.d. noise.

EXAMPLE 2.2 The class of stationary ergodic m.d. processes $\{\eta_t\}$ is very rich. It covers conditional heteroscedastic ARCH models, stochastic volatility models and others. Such processes can usually be written in the form

$$\eta_t = \varepsilon_t \sigma_t, \quad \sigma_t = f(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with $E\varepsilon_t = 0$, $E\varepsilon_t^2 < \infty$ and f is a measurable function of $(\varepsilon_t, \varepsilon_{t-1}, \dots)$. Clearly, such $\{\eta_t\}$ process is a stationary m.d. sequence. Since $\{\varepsilon_t\}$ is an ergodic process, then by Theorem 3.5.8 in Stout (1974), $\{\eta_t\}$ is a stationary ergodic m.d. sequence. It satisfies (2.2) as long as $E\varepsilon_t^4 < \infty$, $E\sigma_t^4 < \infty$.

REMARK 2.2 Verification of the asymptotic normality for $\hat{\theta}_n$ is reduced in this paper to the asymptotic normality of a quadratic form $\sum_{j,k=1, j \neq k}^n \beta_{j-k} \eta_j \eta_k$ of an m.d. noise η_t with

a zero diagonal. Normal approximation for $\hat{\sigma}_n^2$ however would require establishing the asymptotic normality of a quadratic form with a non-zero diagonal. Since $\{\eta_t^2\}$ is not a m.d. noise, proof of the asymptotic normality for $\hat{\sigma}_n^2$ and $(\hat{\theta}_n, \hat{\sigma}_n^2)$ would require additional assumptions on $\{\eta_t\}$.

Under additional assumptions, Hosoya and Taniguchi (1982) established asymptotic normality of Gaussian maximum likelihood estimate for linear processes with uncorrelated noise $\{\eta_t\}$ and Taniguchi (1982) suggested consistent estimates for its asymptotic variance.

Proof of Theorem 2.2. By Theorem 2.1, $\hat{\theta}_n \rightarrow \theta_0$ a.s. Hence, $\nabla_{\theta} Q_n(\hat{\theta}_n) = 0$ with probability tending to 1. Thus, by the continuity and differentiability of function $Q_n(\hat{\theta})$ which is guaranteed by Assumption (a0), by the mean-value theorem there exists $\theta_n^* \in \Theta$ such that $\|\theta_n^* - \theta_0\| \leq \|\hat{\theta}_n - \theta_0\|$ and

$$(2.15) \quad 0 = \nabla_{\theta} Q_n(\hat{\theta}_n) = \nabla_{\theta} Q_n(\theta_0) + \nabla_{\theta} \nabla'_{\theta} Q_n(\theta_n^*)(\hat{\theta}_n - \theta_0).$$

Since the components of $\nabla_{\theta} s_{\theta}^{-1}(u)$ are continuous functions in θ and u , by Lemma 8.2.2. in Giraitis *et al.* (2012),

$$\nabla_{\theta} \nabla'_{\theta} Q_n(\theta) \rightarrow \frac{\sigma_0^2}{2\pi} \int_{\Pi} s_{\theta_0}(u) \nabla_{\theta} \nabla'_{\theta} s_{\theta_0}^{-1}(u) du \quad a.s.$$

uniformly in $\theta \in \Theta$. Together with the consistency of $\hat{\theta}_n$ this implies

$$(2.16) \quad \nabla_{\theta} \nabla'_{\theta} Q_n(\theta_n^*) \rightarrow_p \frac{\sigma_0^2}{2\pi} \int_{\Pi} s_{\theta_0}(u) \nabla_{\theta} \nabla'_{\theta} s_{\theta_0}^{-1}(u) du = \frac{\sigma_0^2}{2\pi} W(\theta_0).$$

By Kolmogorov formula, the parameterization assumption $a_0 = 1$ is equivalent to

$$\int_{\Pi} \log s_{\theta}(u) du = 0, \quad \theta \in \Theta.$$

The use of the latter yields the last equality in (2.16), see Hannan (1973) or, e.g., page 216 in Giraitis *et al.* (2012).

Because of (2.15) and (2.16), (2.11) will follow from the convergence

$$(2.17) \quad -n^{1/2} \nabla_{\theta} Q_n(\theta_0) \rightarrow_D \mathcal{N}(0, V_{\theta_0, \eta}).$$

Let $c \in \mathbb{R}^q$. Denote

$$S_{n,c} := -nc' \nabla_{\theta} Q_n(\theta_0).$$

By Cramér-Wold device, to prove (2.17) it suffices to verify that for any c ,

$$(2.18) \quad n^{-1/2} S_{n,c} \rightarrow \mathcal{N}(0, v_c^2), \quad v_c^2 = c' V_{\theta_0, \eta} c.$$

Express $S_{n,c}$ as a quadratic form

$$(2.19) \quad S_{n,c} = \sum_{t,s=1}^n g_{t-s} X_t X_s, \quad g_t := (2\pi)^{-1} \int_{\Pi} e^{itu} \hat{g}(u) du, \quad t \in \mathbb{Z},$$

where

$$\hat{g}(u) := -c' \nabla_{\theta} s_{\theta_0}^{-1}(u).$$

We prove (2.18) by showing that $S_{n,c}$ satisfies assumptions of Theorem 3.1 below. By Assumptions (a0) and (a1)(ii), the partial derivative $\nabla_u \nabla_{\theta} s_{\theta_0}^{-1}$ is a bounded continuous function which implies that the series $\sum_{j \in \mathbb{Z}} |g_j| < \infty$ converges. Clearly, $g_j = g_{-j}$. By Assumption (a1)(i), $\sum_{k=1}^{\infty} k a_k^2(\theta_0) < \infty$. By (3.26) and (3.34) we obtain

$$(2.20) \quad \begin{aligned} \beta_{k,\theta_0} &= (2\pi)^{-1} \int_{\Pi} e^{iku} \hat{g}(u) s_{\theta_0}(u) du = -(2\pi)^{-1} \int_{\Pi} e^{iku} c' \nabla_{\theta} s_{\theta_0}^{-1}(u) s_{\theta_0}(u) du \\ &= (2\pi)^{-1} \int_{\Pi} e^{iku} c' \nabla_{\theta} \log s_{\theta_0}(u) du, \quad k \in \mathbb{Z}. \end{aligned}$$

Moreover, property $\int_{\Pi} \log s_{\theta}(u) du = 0$ implies that

$$(2.21) \quad \int_{\Pi} \nabla_{\theta} \log s_{\theta}(u) du = 0, \quad \beta_{0,\theta_0} = (2\pi)^{-1} \int_{\Pi} c' \nabla_{\theta} \log s_{\theta_0}(u) du = 0.$$

Hence, Theorem 3.1 implies (2.18) with $V_{\theta_0,\eta}$ as in (2.9).

Proof of (2.12). Recall that $\beta_{k,\theta_0} = \beta_{-k,\theta_0}$ and by (2.21),

$$\beta_{0,\theta_0} = (2\pi)^{-1} \int_{\Pi} \nabla_{\theta} \log s_{\theta_0}(u) du = 0.$$

Therefore, for i.i.d. random variables $\{\eta_t\}$, $V_{\theta_0,\eta}$ in (2.11) takes the form

$$(2.22) \quad V_{\theta_0,\eta} = 4\sigma_0^4 \sum_{k=1}^{\infty} \beta_{k,\theta_0} \beta'_{k,\theta_0} = 2\sigma_0^4 \sum_{k=-\infty}^{\infty} \beta_{k,\theta_0} \beta'_{k,\theta_0}.$$

Hence, using definition (2.9) of β_{k,θ_0} , by Parseval's identity, we obtain (2.12):

$$(2.23) \quad V_{\theta_0,\eta} = \pi^{-1} \sigma_0^4 \int_{\Pi} \nabla_{\theta} \log s_{\theta_0}(u) \nabla'_{\theta} \log s_{\theta_0}(u) du = \pi^{-1} \sigma_0^4 W_{\theta_0}.$$

Finally, under assumption $E[\eta_0^2 \eta_k \eta_s] = 0$ for any $s < k < 0$, (2.13) follows straightforwardly from the definition of $V_{\theta_0,\eta}$ in (2.9) noting that $E[\eta_{-k}^2 \eta_0^2] = (E[\eta_{-k}^2])^2 + \text{cov}(\eta_0^2, \eta_{-k}^2)$ and using (2.22) and (2.23). This completes the proof of the theorem. \square

3 Quadratic forms of m.d. noise

We study here the quadratic form

$$(3.24) \quad S_n = \sum_{t,s=1}^n g_{t-s} X_t X_s$$

with symmetric real weights $g_k = g_{-k}$, $k \in \mathbb{Z}$. We assume that

$$(3.25) \quad X_t = \sum_{j=0}^{\infty} a_j \eta_{t-j}, \quad \sum_{j=0}^{\infty} |a_j| < \infty,$$

is a linear process and $\{\eta_t\}$ is a stationary ergodic m.d. sequence satisfying (2.2). Such a quadratic form appeared in (2.19).

Quadratic forms appear in numerous statistical applications. Asymptotic normality for quadratic forms of linear processes was widely investigated in the statistical and probabilistic literature, see e.g. Hannan (1973), Fox and Taqqu (1987), Giraitis and Surgailis (1990), Robinson (1995), Bhansali, Giraitis and Kokoszka (2007) and others. Sufficient general conditions for asymptotic normality of quadratic forms in i.i.d. random variables were established in Rotar (1973), De Jong (1987) and Guttorp and Lockhart (1988). For quadratic forms in m.d. random variables such conditions were derived in Giraitis, Tanuguchi and Taqqu (2016).

For simplicity set $a_j = 0$ for $j \leq 0$. Denote

$$(3.26) \quad \gamma_k = \sum_{j \in \mathbb{Z}} a_j a_{j+k}, \quad \beta_k = \sum_{j \in \mathbb{Z}} g_j \gamma_{j+k}, \quad k \in \mathbb{Z}.$$

The following theorem establishes the asymptotic normality of a quadratic form S_n . Its proof is based on the results of the paper by Giraitis, Tanuguchi and Taqqu (2016), whose application requires additional technical effort.

THEOREM 3.1 *Suppose that the g_k 's and a_k 's are such that*

$$(3.27) \quad a) \sum_{k \in \mathbb{Z}} |g_k| < \infty, \quad b) \sum_{k=1}^{\infty} k a_k^2 < \infty, \quad c) \beta_0 = 0.$$

Then the quadratic form S_n in (3.24) satisfies

$$(3.28) \quad n^{-1/2}(S_n - ES_n) \rightarrow_D \mathcal{N}(0, v^2), \quad v^2 := 4E \left[\left(\sum_{k=1}^{\infty} \beta_k \eta_{-k} \right)^2 \eta_0^2 \right],$$

$$(3.29) \quad n^{-1/2}ES_n \rightarrow 0.$$

Moreover, v^2 has the following properties.

a) If in addition, $\{\eta_t\}$ is a sequence of i.i.d. random variables, then

$$(3.30) \quad v^2 = 4(E\eta_0^2)^2 \sum_{k=1}^{\infty} \beta_k^2.$$

b) If in addition, $E[\eta_0^2 \eta_s \eta_k] = 0$ for any $s < k < 0$, then

$$(3.31) \quad v^2 = 4 \sum_{k=1}^{\infty} \beta_k^2 E[\eta_0^2 \eta_{-k}^2] = 4(E\eta_0^2)^2 \sum_{k=1}^{\infty} \beta_k^2 + 4 \sum_{k=1}^{\infty} \beta_k^2 \text{cov}(\eta_0^2, \eta_{-k}^2).$$

c) If in addition, $\{\eta_t\}$ satisfies assumption that there exists a positive constant $c > 0$ such that

$$(3.32) \quad E[\eta_k^2 | \mathcal{F}_{k-1}] \geq c > 0, \quad k \in \mathbb{Z}, \quad \text{a.s.},$$

then there exists $c_0 > 0$ such that

$$(3.33) \quad v^2 \geq c_0 \sum_{k=1}^{\infty} \beta_k^2 > 0.$$

REMARK 3.1 Suppose that $g_k = (2\pi)^{-1} \int_{\Pi} e^{ikx} \hat{g}(x) dx$, $k \in \mathbb{Z}$ where $\hat{g}(x)$, $x \in \Pi$ is an even continuous bounded function. Notice that γ_k in (3.26) can be expressed as $\gamma_k = (2\pi)^{-1} \int_{\Pi} e^{ikx} s(x) dx$ with $s(x) = |\sum_{j=0}^{\infty} e^{ijx} a_j|^2$, $x \in \Pi$. Then, by Parseval's identity,

$$(3.34) \quad \beta_k = \sum_{j \in \mathbb{Z}} g_j \gamma_{j+k} = (2\pi)^{-1} \int_{\Pi} e^{ikx} \hat{g}(x) s(x) dx, \quad k \in \mathbb{Z}$$

and the condition $\beta_0 = 0$ is equivalent to

$$(3.35) \quad \int_{\Pi} \hat{g}(x) s(x) dx = 0.$$

Furthermore, for an i.i.d. noise η_j , by Parseval's identity, v^2 in (3.30) can be written as

$$(3.36) \quad v^2 = 2(E\eta_0^2)^2 \sum_{k \in \mathbb{Z}} \beta_k^2 = (E\eta_0^2)^2 \pi^{-1} \int_{\Pi} \hat{g}^2(x) s^2(x) dx.$$

EXAMPLE 3.1 Consider the m.d. noise $\eta_t = \varepsilon_t \varepsilon_{t-1}$ where $\{\varepsilon_t\}$ are independent Gaussian random variables with zero mean and variance $E\varepsilon_0^2 = \sigma_\varepsilon^2$, see Example 2.1. Then

$\text{cov}(\eta_0^2, \eta_{-1}^2) = 2\sigma_0^4$ and $\text{cov}(\eta_0^2, \eta_{-k}^2) = 0$ for $k \geq 2$. Thus, (3.31) becomes

$$v^2 = 4\sigma_0^4 \sum_{k=1}^{\infty} \beta_k^2 + 8\sigma_0^4 \beta_1^2 > 4\sigma_0^4 \sum_{k=1}^{\infty} \beta_k^2.$$

Hence, compared to an i.i.d. noise, dependence in the noise η_j increased the variance v^2 .

Proof of Theorem 3.1. We shall start with the proof of (3.28). Set $S_{\eta,n} = \sum_{j,k=1}^n \beta_{j-k} \eta_j \eta_k$. We shall show that

$$(3.37) \quad S_n - ES_n = S_{\eta,n} + o_p(n^{1/2}),$$

$$(3.38) \quad n^{-1/2} S_{\eta,n} \rightarrow_D \mathcal{N}(0, v^2),$$

which implies (3.28).

First we verify (3.37). Write S_n as

$$(3.39) \quad S_n = \sum_{j,k=-\infty}^n c_{n,jk} \eta_j \eta_k \quad \text{with} \quad c_{n,jk} = \sum_{t,s=1}^n g_{t-s} a_{t-j} a_{s-k}.$$

Then we split S_n into two sums

$$S_n = S_n^\Delta + S_n^o, \quad S_n^\Delta := \sum_{j=-\infty}^n c_{n,jj} \eta_j^2, \quad S_n^o := \sum_{j,k=-\infty: j \neq k}^n c_{n,jk} \eta_j \eta_k,$$

where S_n^Δ is a sum of diagonal term and S_n^o a sum of quadratic term with zero diagonal.

To verify (3.37), it suffices to show that

$$(3.40) \quad E|S_n^\Delta| = o(n^{1/2}),$$

$$(3.41) \quad E(S_n^o - S_{\eta,n})^2 = o(n).$$

We have $E|S_n^\Delta| \leq E\eta_1^2 \sum_{j=-\infty}^n |c_{n,jj}| = o(n^{1/2})$ by (4.48) which proves (3.40) and (3.29).

To show (3.41), write

$$S_n^o - S_{\eta,n} = \sum_{j,k=-\infty: j \neq k}^n a_{n,jk} \eta_j \eta_k, \quad a_{n,jk} := c_{n,jk} - \beta_{j-k} I(1 \leq j, k \leq n).$$

Then by Lemma 3.1(ii) below,

$$\begin{aligned} E(S_n^o - S_{\eta,n})^2 &\leq C \sum_{j,k=-\infty:j \neq k}^n a_{n,jk}^2 = C \sum_{j,k=-\infty:j \neq k}^n (c_{n,jk} - \beta_{j-k}I(1 \leq j, k \leq n))^2 \\ &\leq C \left\{ \sum_{j,k=1:j \neq k}^n (c_{n,jk} - \beta_{j-k})^2 + \sum_{j=-\infty}^0 \sum_{k=-\infty}^n c_{n,jk}^2 \right\} = o(n) \end{aligned}$$

by (4.50) and (4.49) of Lemma 4.1 below.

It remains to show (3.38). Observe that $ES_{\eta,n} = 0$ and

$$\sum_{k \in \mathbb{Z}} |\beta_k| \leq \sum_{j \in \mathbb{Z}} |g_j| \sum_{k \in \mathbb{Z}} |\gamma_k| \leq \sum_{j \in \mathbb{Z}} |g_j| \left(\sum_{k=0}^{\infty} |a_k| \right)^2 < \infty.$$

Therefore, under the additional assumption (3.32) on the m.d. noise η_t , Corrolary 1.1 (i) in Giraitis, Taniguchi and Taqqu (2016) implies that

$$(3.42) \quad \text{var}(S_{\eta,n})^{-1/2} S_{\eta,n} \rightarrow_D \mathcal{N}(0, 1).$$

Assumption (3.32) is required only to show that $\text{var}(S_{\eta,n})$ has property

$$(3.43) \quad \text{var}(S_{\eta,n}) \geq cB_n^2, \quad n \rightarrow \infty, \quad B_n^2 := \sum_{j,k=1}^n \beta_{j-k}^2$$

for some $c > 0$ where B_n is the Euclidean norm of the (symmetric) matrix $(\beta_{j-k}, j, k = 1, \dots, n)$. Since $B_n^2 \sim n \sum_{j \in \mathbb{Z}} \beta_j^2$, condition (3.43) is equivalent to

$$(3.44) \quad \text{var}(S_{\eta,n}) \geq cn, \quad n \rightarrow \infty \quad (\exists c > 0).$$

Therefore, assumption (3.32) can be replaced by (3.44).

To verify (3.44), we will show that under the assumptions of our theorem,

$$(3.45) \quad n^{-1} \text{var}(S_{\eta,n}) \rightarrow v^2, \quad v^2 = 4E \left[\left(\sum_{k=1}^{\infty} \beta_k \eta_{-k} \right)^2 \eta_0^2 \right].$$

If $v^2 > 0$, then (3.44) is valid and (3.42) holds which in turn implies (3.38).

Finally, if $v^2 = 0$, then the normal approximation $n^{-1/2} S_{\eta,n} \rightarrow 0 = \mathcal{N}(0, 0)$ in (3.38) holds with a degenerate limit.

Proof of (3.45). Since η_t is a stationary m.d. sequence, $E\eta_t^4 < \infty$, and $\beta_0 = 0$, then

$$\begin{aligned} ES_{\eta,n}^2 &= E\left(2 \sum_{j=1}^n \eta_j \sum_{k=1}^{j-1} \beta_{j-k} \eta_k\right)^2 = 4 \sum_{j=1}^n E\left[\eta_j^2 \left(\sum_{k=1}^{j-1} \beta_{j-k} \eta_k\right)^2\right] \\ &= 4 \sum_{j=1}^n E\left[\eta_j^2 \left(\sum_{l=1}^{j-1} \beta_l \eta_{j-l}\right)^2\right] = 4 \sum_{j=1}^n E\left[\eta_0^2 \left(\sum_{l=1}^{j-1} \beta_l \eta_{-l}\right)^2\right] = 4 \sum_{j=1}^n \left(\frac{v^2}{4} + r_j\right), \end{aligned}$$

where

$$r_j := E\left[\eta_0^2 \left[\left(\sum_{l=1}^{j-1} \beta_l \eta_{-l}\right)^2 - \left(\sum_{l=1}^{\infty} \beta_l \eta_{-l}\right)^2\right]\right].$$

Hence, $n^{-1}ES_{\eta,n}^2 = v^2 + 4R_n$, where $R_n = n^{-1} \sum_{j=1}^n r_j$. To verify (3.45), it suffices to show that

$$R_n = o(1).$$

Indeed, using equality $a^2 - b^2 = (a-b)^2 + 2(a-b)b$ with $a = \sum_{l=1}^{j-1} \beta_l \eta_{-l}$ and $b = \sum_{l=1}^{\infty} \beta_l \eta_{-l}$, we obtain

$$\begin{aligned} |r_j| &\leq E\left[\eta_0^2 \left\{ \left(\sum_{l=j}^{\infty} \beta_l \eta_{-l}\right)^2 + 2 \left|\sum_{l=j}^{\infty} \beta_l \eta_{-l}\right| \left|\sum_{l=1}^{\infty} \beta_l \eta_{-l}\right| \right\}\right] \\ &\leq (E\eta_0^4)^{1/2} \left\{ \left(E\left(\sum_{l=j}^{\infty} \beta_l \eta_{-l}\right)^4\right)^{1/2} + 2 \left(E\left(\sum_{l=j}^{\infty} \beta_l \eta_{-l}\right)^4\right)^{1/4} \left(E\left(\sum_{l=1}^{\infty} \beta_l \eta_{-l}\right)^4\right)^{1/4} \right\}. \end{aligned}$$

Denote $P_j = \sum_{l=j}^{\infty} \beta_l^2$, $j \geq 1$. By (3.46),

$$E\left(\sum_{l=j}^{\infty} \beta_l \eta_{-l}\right)^4 \leq CP_j^2, \quad E\left(\sum_{l=1}^{\infty} \beta_l \eta_{-l}\right)^4 \leq CP_1^2.$$

Hence, $|r_j| \leq C\{P_j + P_j^{1/2}\}$ for $j \geq 1$. Since $P_j \leq C \sum_{l=j}^{\infty} |\beta_l| \rightarrow 0$ as $j \rightarrow \infty$, then

$$|R_n| \leq n^{-1} \sum_{j=1}^n |r_j| \leq Cn^{-1} \sum_{j=1}^n \{P_j + P_j^{1/2}\} = o(1).$$

This proves (3.45) and completes the proof of (3.28).

The claim (3.30) is obvious. Equality (3.31) follows straightforwardly noting that $E[\eta_0^2 \eta_{-k}^2] = (E\eta_0^2)^2 + \text{cov}(\eta_0^2, \eta_{-k}^2)$.

To verify (3.33), notice that under (3.32),

$$E[\eta_0^2 (\sum_{l=1}^{\infty} \beta_l \eta_{-l})^2] = E[E[\eta_0^2 | \mathcal{F}_{-1}] (\sum_{l=1}^{\infty} \beta_l \eta_{-l})^2] \geq cE[(\sum_{l=1}^{\infty} \beta_l \eta_{-l})^2] = cE\eta_1^2 \sum_{l=1}^{\infty} \beta_l^2$$

which implies (3.33). This completes the proof of the theorem. \square

The following lemma was used in the proof of Theorem 3.1 to show (3.41).

LEMMA 3.1 (i) *If the m.d. sequence η_t satisfies $\max_t E|\eta_t|^p < \infty$, for some $p \geq 2$, then*

$$(3.46) \quad E|\sum_{j \in \mathbb{Z}} d_j \eta_j|^p \leq C(\sum_{j \in \mathbb{Z}} d_j^2)^{p/2},$$

for any d_j 's such that $\sum_{j \in \mathbb{Z}} d_j^2 < \infty$, where $C < \infty$.

(ii) *If in addition, $p \geq 4$, then*

$$(3.47) \quad E\left(\sum_{j,k=-\infty: j \neq k}^{\infty} a_{jk} \eta_j \eta_k\right)^2 \leq C \sum_{j,k=-\infty}^{\infty} a_{jk}^2$$

for any a_{jk} 's such that $\sum_{j,k \in \mathbb{Z}: j \neq k} a_{jk}^2 < \infty$, where $C < \infty$.

Proof of Lemma 3.1. (i) The bound (3.46) is known, see e.g. Lemma 2.5.2 in Giraitis *et al.* (2012). (ii) Since η_t is a m.d. sequence, then by (3.46),

$$\begin{aligned} E\left(\sum_{j,k=-\infty: j \neq k}^{\infty} a_{jk} \eta_j \eta_k\right)^2 &\leq E\left(\sum_{j=-\infty}^{\infty} \eta_j \sum_{k=-\infty}^{j-1} (a_{jk} + a_{kj}) \eta_k\right)^2 \\ &\leq \sum_{j=-\infty}^{\infty} E\left[\eta_j^2 \left(\sum_{k=-\infty}^{j-1} (a_{jk} + a_{kj}) \eta_k\right)^2\right] \leq \sum_{j=-\infty}^{\infty} (E\eta_j^4)^{1/2} \left(E\left(\sum_{k=-\infty}^{j-1} (a_{jk} + a_{kj}) \eta_k\right)^4\right)^{1/2} \\ &\leq C \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{j-1} (a_{jk} + a_{kj})^2 \leq C \sum_{j,k=-\infty}^{\infty} a_{jk}^2. \end{aligned}$$

This completes the proof of the lemma. \square

4 Properties of the weights

In this section we derive auxiliary results of properties of the weights $c_{n,jk}$ and β_k given in (3.39) and (3.26). We used them in the proof of Theorem 3.1.

LEMMA 4.1 *Let the g_k 's and a_k 's satisfy (3.27) a)-c). Then,*

$$(4.48) \quad \sum_{j=-\infty}^n |c_{n,jj}| = o(n^{1/2}),$$

$$(4.49) \quad \sum_{j=-\infty}^0 \sum_{k=-\infty}^n c_{n,jk}^2 = o(n),$$

$$(4.50) \quad \sum_{j,k=1}^n (c_{n,jk} - \beta_{j-k})^2 = o(n).$$

Proof. To prove (4.48), write

$$\sum_{j=-\infty}^n |c_{n,jj}| = \sum_{j=1}^n |c_{n,jj}| + \sum_{j=-\infty}^0 |c_{n,jj}| =: s_{n,1} + s_{n,2}.$$

It remains to show that as $n \rightarrow \infty$,

$$(4.51) \quad s_{n,k} = o(n^{1/2}), \quad k = 1, 2.$$

First we consider $s_{n,1}$. To evaluate $c_{n,jj}$ for $1 \leq j \leq n$, recall that $a_j = 0$, $j \leq 0$. After the change of summation variables $u = t - j$, $v = s - j$, we obtain

$$c_{n,jj} = \sum_{t,s=1}^n g_{t-s} a_{t-j} a_{s-j} = \sum_{u,v=0}^{n-j} g_{u-v} a_u a_v.$$

Observe that

$$\sum_{u,v=0}^{\infty} g_{u-v} a_u a_v = \sum_{s \in \mathbb{Z}} g_s \gamma_s = \beta_0 = 0,$$

where the last equality holds by assumption (3.27) c). Hence,

$$\begin{aligned} |c_{n,jj}| &= |c_{n,jj} - \beta_0| \\ &= \left| \left(\sum_{u,v=0}^{n-j} - \sum_{u,v=0}^{\infty} \right) g_{u-v} a_u a_v \right| \\ &\leq \left(\sum_{u=n-j+1}^{\infty} \sum_{v=0}^{\infty} + \sum_{u=0}^{\infty} \sum_{v=n-j+1}^{\infty} \right) |g_{u-v} a_u a_v|. \end{aligned}$$

This, together with the change of the summation order

$$\sum_{j=1}^n \sum_{u=n-j+1}^{\infty} = \sum_{u=1}^{\infty} \sum_{j=\max(n-u+1,1)}^n$$

yields

$$s_{n,1} = \sum_{j=1}^n |c_{n,jj}| \leq 2 \sum_{j=1}^n \sum_{u=n-j+1}^{\infty} \sum_{v=0}^{\infty} |g_{u-v} a_u a_v| \leq 2 \sum_{u=1}^{\infty} \sum_{v=0}^{\infty} |g_{u-v} (u \wedge n) a_u a_v|.$$

where $u \wedge n = \min(u, n)$. Recall the inequality

$$(4.52) \quad \sum_{u,v \in \mathbb{Z}} |f_{u-v} h_u \nu_v| \leq \left(\sum_{u \in \mathbb{Z}} |f_u| \right) \left(\sum_{s \in \mathbb{Z}} h_s^2 \sum_{v \in \mathbb{Z}} \nu_v^2 \right)^{1/2}$$

which holds for any sequences (f_t) , (h_t) and (ν_t) of real numbers such that the right hand side r.h.s. of (4.52) is finite. Applying (4.52) with $f_{u-v} = g_{u-v}$, $h_u = (u \wedge n) |a_u|$ and $\nu_v = |a_v|$, we obtain

$$s_{n,1} \leq \left(\sum_{u \in \mathbb{Z}} |g_u| \right) \left(\sum_{s=0}^{\infty} (s \wedge n)^2 a_s^2 \right)^{1/2} \left(\sum_{v=0}^{\infty} a_v^2 \right)^{1/2} \leq C \left(\sum_{s=1}^{\infty} (s \wedge n)^2 a_s^2 \right)^{1/2}$$

since by (3.27) a)-b), $\sum_{u \in \mathbb{Z}} |g_u| < \infty$ and $\sum_{v=0}^{\infty} a_v^2 < \infty$. Set $L = \log n$. We shall bound $(s \wedge n)^2 \leq ns$ for $s \geq L$; $(s \wedge n)^2 \leq Ls$ for $1 \leq s < L$. Then,

$$s_{n,1} \leq C(L \sum_{s=1}^{L-1} s a_s^2)^{1/2} + C(n \sum_{s=L}^{\infty} s a_s^2)^{1/2} = o(n^{1/2})$$

since $\sum_{s=L}^{\infty} s a_s^2 = o(1)$ as $L \rightarrow \infty$ by (3.27) b).

Next we consider $s_{n,2}$. To evaluate $c_{n,jj}$ for $j \leq 0$, we apply inequality (4.52) with $f_u = g_u$, $h_t = a_{t-j} I(1 \leq t \leq n)$ and $\nu_s = a_{s-j} I(1 \leq s \leq n)$ which yields

$$|c_{n,jj}| = \left| \sum_{t,s=1}^n g_{t-s} a_{t-j} a_{s-j} \right| \leq \left(\sum_{u \in \mathbb{Z}} |g_u| \right) \left(\sum_{t=1}^n a_{t-j}^2 \right)^{1/2} \left(\sum_{s=1}^n a_{s-j}^2 \right)^{1/2} \leq C \left(\sum_{t=1}^n a_{t-j}^2 \right).$$

Observe that $a_{t-j}^2 \leq a_{t-j}^2(t-j)/(1-j)$ for $j \leq 0$. Hence,

$$\begin{aligned}
s_{n,2} &= \sum_{j=-\infty}^0 |c_{n,jj}| \leq C \sum_{j=-\infty}^0 \sum_{t=1}^n a_{t-j}^2 \leq C \sum_{j=-\infty}^0 \frac{1}{1-j} \sum_{t=1}^n a_{t-j}^2(t-j) \\
&\leq \sum_{j=-\infty}^{-n} n^{-1} \sum_{t=1}^n a_{t-j}^2(t-j) + \left(\sum_{j=-n}^0 \frac{1}{1-j} \right) \left(\sum_{t \in \mathbb{Z}} a_t^2 |t| \right) \\
&\leq C n^{-1} \sum_{t=1}^n \left(\sum_{j \in \mathbb{Z}} a_j^2 |j| \right) + C \log n \leq C(1 + \log n) = o(n^{1/2}).
\end{aligned}$$

This completes the proof of (4.51) and (4.48).

Proof of (4.49). Recall notation $\gamma_k = \sum_{j \in \mathbb{Z}} a_j a_{j+k}$, and set $\theta_{t,u} = \sum_{j=-\infty}^0 a_{t-j} a_{u-j}$. Notice that $\sum_{j \in \mathbb{Z}} a_{t-j} a_{u-j} = \gamma_{t-u}$. Then,

$$i_n := \sum_{j=-\infty}^0 \sum_{k=0}^{\infty} c_{n,jk}^2 = \sum_{j=-\infty}^0 \sum_{k=0}^{\infty} \left(\sum_{t,s=1}^n g_{t-s} a_{t-j} a_{s-k} \right)^2 \leq \sum_{t,s,u,v=1}^n g_{t-s} g_{u-v} \theta_{t,u} \gamma_{s-v}.$$

Applying inequality (4.52), we get, for any $1 \leq t, u \leq n$,

$$\sum_{s,v=1}^n |\gamma_{s-v} g_{t-s} g_{u-v}| \leq \left(\sum_{u \in \mathbb{Z}} |\gamma_u| \right) \left(\sum_{s \in \mathbb{Z}} g_s^2 \right)^{1/2} \left(\sum_{v \in \mathbb{Z}} g_v^2 \right)^{1/2}$$

where the r.h.s. does not depend on t, u and n . Observe that $\sum_{u \in \mathbb{Z}} g_u^2 < \infty$ by (3.27) a), while $A := \sum_{j \in \mathbb{Z}} |a_j| < \infty$ holds by (3.27) b) and implies $\sum_{v \in \mathbb{Z}} |\gamma_v| < \infty$. Hence,

$$(4.53) \quad i_n \leq C \sum_{t,u=1}^n |\theta_{t,u}| = C \sum_{t,u=1}^n \left| \sum_{j=-\infty}^0 a_{t-j} a_{u-j} \right|.$$

Set $L = \log n$. Then,

$$\begin{aligned}
i_n &\leq \sum_{t,u=1}^n \left\{ \sum_{j=-L+1}^0 |a_{t-j} a_{u-j}| + \sum_{j=-\infty}^{-L} |a_{t-j} a_{u-j}| \right\} \\
&\leq \sum_{j=-L+1}^0 \left(\sum_{t \in \mathbb{Z}} |a_{t-j}| \right) \left(\sum_{u \in \mathbb{Z}} |a_{u-j}| \right) + \sum_{t=1}^n \left[\sum_{j=-\infty}^{-L} |a_{t-j}| \left(\sum_{u \in \mathbb{Z}} |a_{u-j}| \right) \right] \\
&\leq LA^2 + \sum_{t=1}^n \left(\sum_{s=L}^{\infty} |a_s| \right) A = o(n)
\end{aligned}$$

since $\sum_{s=L}^{\infty} |a_s| \rightarrow 0$ as $n \rightarrow \infty$. This proves (4.49).

Proof of (4.50). Observe that

$$\sum_{t,s \in \mathbb{Z}} g_{t-s} a_{t-j} a_{s-k} = \sum_{u \in \mathbb{Z}} g_u \gamma_{u+k-j} = \sum_{u \in \mathbb{Z}} g_u \gamma_{u+j-l} = \beta_{j-k}.$$

Since $a_l = 0$ for $l \leq 0$, $\beta_{j-k} = \sum_{t,s=1}^{\infty} g_{t-s} a_{t-j} a_{s-k}$ for $1 \leq j, k \leq n$. Hence,

$$\begin{aligned} i'_n &:= \sum_{j,k=1}^n (c_{n,jk} - \beta_{j-k})^2 = \sum_{j,k=1}^n \left(\sum_{t,s=1}^n - \sum_{t,s=1}^{\infty} \right) g_{t-s} a_{t-j} a_{s-k} \Big)^2 \\ &\leq \sum_{j=1}^n \sum_{k \in \mathbb{Z}} \left(2 \sum_{t=n+1}^{\infty} \sum_{s=1}^{\infty} |g_{t-s} a_{t-j} a_{s-k}| \right)^2. \end{aligned}$$

Denote $\gamma_k^* = \sum_{j \in \mathbb{Z}} |a_j a_{j+k}|$ and $\theta_{n,t,u}^* = \sum_{j=1}^n |a_{t-j} a_{u-j}|$. Then,

$$i'_n \leq 4 \sum_{t,u=n+1}^{\infty} \sum_{s,v=1}^{\infty} |g_{t-s} g_{u-v}| \theta_{n,t,u}^* \gamma_{s-v}^*.$$

Clearly, as above in estimation of i_n , applying inequality (4.52) we obtain

$$\sum_{s,v=1}^{\infty} |\gamma_{s-v}^* g_{t-s} g_{u-v}| \leq \left(\sum_{u \in \mathbb{Z}} \gamma_u^* \right) \left(\sum_{s \in \mathbb{Z}} g_s^2 \right) \leq C < \infty$$

uniformly in t, u and n . Hence,

$$(4.54) \quad i'_n \leq C \sum_{t,u=n+1}^{\infty} \theta_{n,t,u}^* = C \sum_{t,u=n+1}^{\infty} \sum_{j=1}^n |a_{t-j} a_{u-j}|.$$

Similarly as in estimation i_n above, setting $L = \log n$ we obtain,

$$\begin{aligned} i'_n &\leq \sum_{t=n+1}^{n+L} \left(\sum_{j \in \mathbb{Z}} |a_{t-j}| \right) \left(\sum_{u \in \mathbb{Z}} |a_{u-j}| \right) + \sum_{j=1}^n \left[\sum_{t=n+L+1}^{\infty} |a_{t-j}| \left(\sum_{u \in \mathbb{Z}} |a_{u-j}| \right) \right] \\ &\leq LA^2 + \sum_{j=1}^n \left(\sum_{s=L}^{\infty} |a_s| \right) A = o(n) \end{aligned}$$

which proves (4.50). This completes the proof of the lemma. \square

5 Applications

We shall demonstrate the impact of the m.d. noise η_t on Whittle estimation using examples of AR(1) and MA(1) processes.

First we consider an AR(1) process

$$(5.55) \quad X_t = \phi X_{t-1} + \eta_t,$$

where $|\phi| < 1$ and η_t is a stationary ergodic m.d. noise satisfying (2.2). $\{X_t\}$ can be written as a stationary linear process

$$(5.56) \quad X_t = \sum_{j=0}^{\infty} \phi^j \eta_{t-j}.$$

It has the spectral density

$$(5.57) \quad f_{\sigma_0^2, \phi}(u) = \frac{\sigma_0^2}{2\pi} s_\phi(u), \quad s_\phi(u) = \left| \sum_{j=0}^{\infty} \phi^j e^{iju} \right|^2 = (1 - 2\phi \cos u + \phi^2)^{-1}$$

parametrized by the parameter ϕ . Since

$$(d/d\phi)s_\phi^{-1}(u) = -2 \cos u + 2\phi, \quad (d^2/d^2\phi)s_\phi^{-1}(u) = 2,$$

solving the equation $(d/d\phi)Q_n(\phi) = 0$ implies that Whittle estimate

$$(5.58) \quad \hat{\phi} = \frac{\int_{\Pi} \cos(u) I_n(u) du}{\int_{\Pi} I_n(u) du} = \frac{\sum_{k=2}^n X_k X_{k-1}}{\sum_{k=1}^n X_k^2}$$

is the sample correlation of $\{X_t\}$ at the lag 1.

Let $\Theta = [-a, a]$ where $0 < a < 1$. Clearly, the family of functions $s_\phi(u)$, $\phi \in \Theta$, $u \in \Pi$ satisfies Assumptions (a0) and (a1).

Theorem 2.2 implies the following result.

COROLLARY 5.1 *The Whittle estimator $\hat{\phi}$ given by (5.58) has the following properties:*

$$(5.59) \quad n^{1/2}(\hat{\phi} - \phi_0) \rightarrow \mathcal{N}(0, v_{\phi_0, \eta}^2),$$

$$v_{\phi_0, \eta}^2 := \frac{E[X_{-1}^2 \eta_0^2]}{\text{var}^2(X_0)} = (1 - \phi_0^2) + \frac{\text{cov}(X_{-1}^2, \eta_0^2)}{\text{var}^2(X_0)}.$$

(i) *If the m.d. noise $\{\eta_t\}$ is an i.i.d. sequence, then*

$$(5.60) \quad v_{\phi_0, \eta}^2 = 1 - \phi_0^2.$$

(ii) If the m.d. noise $\{\eta_t = \varepsilon_t \varepsilon_{t-1}\}$ is as in (2.14), then

$$(5.61) \quad v_{\phi_0, \eta}^2 = (1 - \phi_0^2) + 2(1 - \phi_0^2)^2.$$

REMARK 5.1 Relations (5.59) and (5.61) show that m.d. noise $\{\eta_t\}$ may have strong impact on the variance $v_{\phi_0, \eta}^2$ of the estimate $\hat{\phi}$ and thus on the confidence intervals for ϕ_0 .

The unknown variance $v_{\phi_0, \eta}^2$ in (5.59) can be estimated as follows. Recall that $EX_t = 0$. Since $\{\eta_t^2 X_{t-1}^2\}$ and $\{X_{t-1}^2\}$ are stationary ergodic sequences, then

$$\frac{n^{-1} \sum_{k=2}^n \eta_k^2 X_{k-1}^2}{(n^{-1} \sum_{k=1}^n X_k^2)^2} \rightarrow_p \frac{E[\eta_0^2 X_{-1}^2]}{(E[X_0^2])^2} = v_{\phi_0, \eta}^2.$$

Hence, $v_{\phi_0, \eta}^2$ can be estimated by

$$(5.62) \quad \hat{v}_{\phi_0, \eta}^2 := \frac{n^{-1} \sum_{k=2}^n \hat{\eta}_k^2 X_{k-1}^2}{(n^{-1} \sum_{k=1}^n X_k^2)^2} \rightarrow_p v_{\phi_0, \eta}^2, \quad \hat{\eta}_k := X_k - \hat{\phi} X_{k-1}$$

which implies

$$\frac{n^{1/2}}{\sqrt{\hat{v}_{\phi_0, \eta}^2}} (\hat{\phi} - \phi_0) \rightarrow \mathcal{N}(0, 1).$$

Proof of Corollary 5.1. Recall that

$$\gamma_{k, \phi} := (2\pi)^{-1} \int_{\Pi} e^{iku} s_{\phi}(u) du = \phi^k \gamma_{0, \phi}, \quad k \geq 1; \quad \gamma_{0, \phi} = (1 - \phi^2)^{-1}.$$

Thus, $\beta_{k, \phi}$'s in (2.9) can be written as

$$\begin{aligned} \beta_{k, \phi} &= -(2\pi)^{-1} \int_{\Pi} e^{iku} \nabla_{\phi} s_{\phi}^{-1}(u) s_{\phi}(u) du \\ &= (2\pi)^{-1} \int_{\Pi} e^{iku} (2 \cos u - 2\phi) s_{\phi}(u) du \\ &= (2\pi)^{-1} \int_{\Pi} e^{iku} (e^{iu} + e^{-iu} - 2\phi) s_{\phi}(u) du \\ &= \gamma_0 (\phi^{k+1} + \phi^{k-1} - 2\phi^{k+1}) \\ (5.63) \quad &= \phi^{k-1}, \quad k \geq 1. \end{aligned}$$

So, $\sum_{j=1}^{\infty} \beta_{j,\phi_0}^2 = (1 - \phi_0^2)^{-1} =: B_{\phi_0}$, and by (2.9) and (2.10), we obtain

$$\begin{aligned} W_{\phi_0} &= 4\pi B_{\phi_0}, \\ V_{\phi_0,\eta} &= 4E\left[\left(\sum_{k=1}^{\infty} \beta_{k,\phi_0} \eta_{-k}\right)^2 \eta_0^2\right] = 4E\left[\left(\sum_{k=0}^{\infty} \phi_0^k \eta_{-k-1}\right)^2 \eta_0^2\right] = 4E[X_{-1}^2 \eta_0^2], \\ \Omega_{\phi_0,\eta} &= \sigma_0^{-4} B_{\phi_0}^{-2} E[X_{-1}^2 \eta_0^2] = B_{\phi_0}^{-1} + \sigma_0^{-4} B_{\phi_0}^{-2} \text{cov}(X_{-1}^2, \eta_0^2). \end{aligned}$$

Notice that $\sigma_0^2 B_{\phi_0} = \text{var}(X_0)$. This together with (2.11) of Theorem 2.2 proves (5.59). Clearly, (5.59) implies (i) and (ii). \square

Next we consider the example of MA(1) process

$$(5.64) \quad X_t = \eta_t - \theta \eta_{t-1}$$

where $|\theta| < 1$ and η_t is as in (5.55). This process has the spectral density

$$(5.65) \quad f_{\sigma_0^2,\theta}(u) = \frac{\sigma_0^2}{2\pi} s_{\theta}(u), \quad s_{\theta}(u) = |1 - \theta e^{iu}|^2 = 1 - 2\theta \cos u + \theta^2$$

parametrized by the parameter θ .

Let $\Theta = [-a, a]$ where $0 < a < 1$. Since $(d/d\theta)s_{\theta}(u) = -2\cos u + 2\theta$, $(d^2/d^2\theta)s_{\theta} = 2$, functions $s_{\theta}(u)$, $\theta \in \Theta$, $u \in \Pi$ satisfy Assumptions (a0) and (a1). Moreover, (5.63) implies that the weights β_{k,θ_0} in (2.9) satisfy

$$\beta_{k,\theta_0} := (2\pi)^{-1} \int_{\Pi} e^{iku} \nabla_{\theta} \log s_{\theta_0}(u) du = -(2\pi)^{-1} \int_{\Pi} e^{iku} \nabla_{\theta} \log s_{\theta_0}^{-1}(u) du = -\theta^{k-1}, \quad k \geq 1.$$

Hence, by (2.9) and (2.10), setting $Z_t = \sum_{k=0}^{\infty} \theta_0^k \eta_{t-k}$, we obtain

$$\begin{aligned} W_{\theta_0} &= 4\pi \sum_{j=1}^{\infty} \beta_{j,\theta_0}^2 = 4\pi(1 - \theta_0^2)^{-1} =: 4\pi B_{\theta_0}, \\ V_{\theta_0,\eta} &= 4E\left[\left(\sum_{k=1}^{\infty} \beta_{k,\theta_0} \eta_{-k}\right)^2 \eta_0^2\right] = 4E\left[\left(\sum_{k=0}^{\infty} \theta_0^k \eta_{-k-1}\right)^2 \eta_0^2\right] = 4E[Z_{-1}^2 \eta_0^2], \\ \Omega_{\theta_0,\eta} &= \sigma_0^{-4} B_{\theta_0}^{-2} E[Z_{-1}^2 \eta_0^2] = B_{\theta_0}^{-1} + \sigma_0^{-4} B_{\theta_0}^{-2} \text{cov}(Z_{-1}^2, \eta_0^2). \end{aligned}$$

Since $\sigma_0^2 B_{\theta_0} = \text{var}(Z_0)$, (2.11) of Theorem 2.2 implies the following result.

COROLLARY 5.2 *The Whittle estimator $\hat{\theta}$ for MA(1) process (5.64) has the following*

properties:

$$(5.66) \quad n^{1/2}(\hat{\theta} - \theta_0) \rightarrow \mathcal{N}(0, v_{\theta_0, \eta}^2),$$

$$v_{\theta_0, \eta}^2 := \frac{E[Z_{-1}^2 \eta_0^2]}{\text{var}^2(Z_0)} = (1 - \theta_0^2) + \frac{\text{cov}(Z_{-1}^2, \eta_0^2)}{\text{var}^2(Z_0)}.$$

- (i) If the m.d. noise $\{\eta_t\}$ is an i.i.d. sequence, then $v_{\theta_0, \eta}^2 = 1 - \theta_0^2$.
(ii) If the m.d. noise $\{\eta_t = \varepsilon_t \varepsilon_{t-1}\}$ is as in (2.14), then $v_{\theta_0, \eta}^2 = (1 - \theta_0^2) + 2(1 - \theta_0^2)^2$.

REMARK 5.2 Expressions of asymptotic variances $v_{\phi_0, \eta}^2$ in (5.59) and $v_{\theta_0, \eta}^2$ in (5.66) for parametric Whittle estimators of AR(1) and MA(1) models are remarkably similar. For AR(1) model, $v_{\phi_0, \eta}^2$ can be estimated by (5.62).

For MA(1) model, a consistent estimate of $v_{\theta_0, \eta}^2$ can be constructed as follows. By inverting MA(1) process (5.64) we obtain $\eta_t = (1 - \theta L)^{-1} X_t = \sum_{j=0}^{\infty} \theta^j X_{t-j}$, where L is the backshift operator. Similarly,

$$Z_t = (1 - \theta L)^{-1} \eta_t = (1 - \theta L)^{-2} X_t = \left(\sum_{s=0}^{\infty} (s+1) \theta^s L^s \right) X_t = \sum_{s=0}^{\infty} (s+1) \theta^s X_{t-s}.$$

Since $\{Z_{t-1}^2 \eta_t^2\}$ and $\{Z_t^2\}$ are stationary ergodic processes, then

$$\frac{n^{-1} \sum_{j=2}^n Z_{j-1}^2 \eta_j^2}{(n^{-1} \sum_{j=1}^n Z_j^2)^2} \rightarrow_p \frac{E[Z_{-1}^2 \eta_0^2]}{(E[Z_0^2])^2} = v_{\theta_0, \eta}^2.$$

Setting $\hat{Z}_t = \sum_{s=0}^t (s+1) \hat{\theta}^s X_{t-s}$, $\hat{\eta}_t = \sum_{s=0}^t \hat{\theta}^s X_{t-s}$ we obtain the required estimate:

$$\hat{v}_{\theta_0, \eta}^2 = \frac{n^{-1} \sum_{j=2}^n \hat{Z}_{j-1}^2 \hat{\eta}_j^2}{(n^{-1} \sum_{j=1}^n \hat{Z}_j^2)^2} \rightarrow_p v_{\theta_0, \eta}^2.$$

Acknowledgments. The authors would like to thank the two anonymous referees for valuable comments and suggestions. Liudas Giraitis and Murad S. Taqqu would like to thank Masanobu Taniguchi for his hospitality in Japan. M. Taniguchi was supported by JSPS Kiban grant A-15H02061 at Waseda University.

The corresponding author states that there is no conflict of interest.

References

- Bhansali, R.J., Giraitis, L. and Kokoszka, P. (2007). Approximations and limit theory for quadratic forms of linear processes. *Stoch. Process. Appl.* **117**, 71–95.

- De Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Th. Rel. Fields* **75**, 261–277.
- Fox, R. and Taqqu, M.S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14**, 517-532.
- Fox, R. and Taqqu, M.S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Th. Rel. Fields* **74**, 213–240.
- Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle’s estimate. *Probab. Th. Rel. Fields.* **86**, 87-104.
- Giraitis, L., Hidalgo, J. and Robinson, P.M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *Ann. Statist.* **29**, 987-1023.
- Giraitis, R. Leipus and D. Surgailis. (2007). Recent advances in ARCH modelling. In: G. Teyssiere and A.P. Kirman, Editors, *Long Memory in Economics*, Springer, 3-38.
- Giraitis, L., Koul, H.L. and Surgailis, D. (2012) *Large Sample Inference for Long Memory Processes*, Imperial College Press, London.
- Giraitis, Taniguchi, M., and Taqqu, M.S. (2016). Asymptotic normality of quadratic forms of martingale differences. *Statistical Inference of Stochastic Processes.* **20**, 315–327.
- Guttorp, P. and Lockhart, R. A. (1988). On the asymptotic distribution of quadratic forms in uniform order statistics. *Ann. Statist.* **16**, 433–449.
- Hannan, E.J. (1973). The asymptotic theory of linear time series models. *J. Appl. Probab.*, **10**, 130-145.
- Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.* **10**, 132-153.
- Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23**, 1630–1661.
- Rotar, V. I. (1973). Certain limit theorems for polynomials of degree two. *Teoria Veroyatnosti i Primenenia.* **18**, 527–534; in Russian.
- Stout, W. (1974). *Almost Sure Convergence*. Academic Press, New York.
- Taniguchi, M. (1982). On estimation of the integrals of the fourth order cumulant spectral density. *Biometrika* **69**, 117–122.

Wu, P.T. and Shieh, S.J. (2007). Value-at-Risk analysis for long-term interest rate futures: Fat-tail and long memory in return innovations. *J. of Empirical Finance*. **14**, 248-259.

Whittle, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2**, 423-443.