Asymptotic validity of a white noise test under weak conditions: a cumulants approach

Monika Bhattacharjee
Stat.-Math. Unit
Indian Statistical Institute
Kolkata, India
monaiidexp.gamma@gmail.com

Arup Bose [*]
Stat.-Math Unit
Indian Statistical Institute
Kolkata, India
bosearu@gmail.com

Radhenduhska Srivastava [†]
Department of Mathematics
Indian Institute of Technology
Mumbai, India
rsrivastava22@gmail.com

Submitted August 04, 2014

Abstract

[Shao, 2011] proposed a test for white noise under unknown dependence and the assumption of Geometric Moment Contraction of order 8 (GMC(8)). This assumption often makes the parameter space additionally restrictive. Moreover, the infinite order linear process with polynomially decaying coefficients does not satisfy the GMC condition of any order. [Hong, 1996] and [Hong and Lee, 2003] suggested the same test respectively for the IID process and stationary martingale differences, and respectively under the finiteness of the 4-th and 8-th moment of the process.

We take a cumulants approach. Under assumption of summability of the cumulants of the process, we prove that the above test remains valid. While this in itself requires all moments to be finite, we show, by an appropriate truncation argument, that for specific models, the test remains valid when the 4-th moment is finite. These include many non-linear models such as various GARCH models, stochastic autoregressive volatility model and autoregressive conditional duration model.

Under alternative hypotheses [Hong and Lee, 2003] and [Shao, 2011] showed probability convergence of the test statistic when the 8-th moment of the process is finite. We additionally prove distribution convergence of the suitably standardized test statistic. We also reduce the moment condition for some specific models such as infinite order moving average, non-linear moving average and bilinear models.

[*] Research supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India.

[†] Research supported by INSPIRE fellowship, Dept. of Science and Technology, Government of India.

Key words and phrases. Testing, white noise, strictly stationary process, α-mixing, β-mixing, cumulants, autocovariance, non-linear models, geometric moment contraction, kernel.

JEL Classifications System C12, C22.

1 Introduction

Tests for white noise in time series analysis are based on the sample autocovariances and use various degrees of assumptions so as to be asymptotically valid under the null hypothesis. Let $\{X_t\}$ be a zero-mean stationary process with finite variance $\mathrm{Var}(X_t) = \sigma^2$. Then the autocovariance and the autocorrelation function of the process are respectively defined as

$$\gamma(i) = \mathrm{Cov}(X_t, X_{t+i}), \quad \rho(i) = \sigma^{-2}\mathrm{Cov}(X_t, X_{t+i}), \ \ \forall i = 0, \pm 1, \pm 2, \dots. \tag{1.1}$$

Note that $\gamma(0) = \sigma^2$. Therefore, the null and alternative hypothesis to be tested are

$$H_0 : \gamma(i) = 0 \ \ \forall i \neq 0 \quad \text{against} \quad H_1 : \gamma(i) \neq 0 \ \text{ for some } i \neq 0. \tag{1.2}$$

Equivalently,

$$H_0 : \rho(i) = 0 \ \ \forall i \neq 0 \quad \text{against} \quad H_1 : \rho(i) \neq 0 \ \text{ for some } i \neq 0.$$

For a sample of size n, the i-th empirical autocovariance and autocorrelation are given by

$$\hat{\gamma}(i) = n^{-1} \sum_{t=|i|+1}^{n} X_t X_{t-|i|} \ \text{ and } \hat{\rho}(i) = \frac{\hat{\gamma}(i)}{\hat{\gamma}(0)}, \ \ \forall i = 0, \pm 1, \pm 2, \dots, \pm(n-1). \tag{1.3}$$

The widely used Box-Pierce ([Box and Pierce, 1970]) portmanteau test statistic is

$$Q_n = \sum_{j=1}^{m} \hat{\rho}^2(j),$$

where m is a fixed truncation number. When $\{X_t\}$ are independently and identically distributed (i.i.d.), the asymptotic distribution of nQ_n is $\chi^2(m)$. If X_t's are replaced by the residuals from a fitted model with m' parameters, then the limiting distribution is $\chi^2(m - m')$. Several modifications to this statistic have been suggested, see for example [Ljung and Box, 1978] and [Mc.Leod, 1978], to overcome its conservative nature. These statistics are well suited to deal with testing the adequacy of fitted standard linear time series models such as the ARMA. One crucial assumption often made in their application is that $\{X_t\}$ is Gaussian. Under Gaussianity, $\{X_t\}$ is a white noise if and only if it is i.i.d., and thus the null distributions for all the previous statistics are derived under independence. But in general, $\{X_t\}$ may not be Gaussian and also the true model may not be linear.

Since 1980, several classes of important non-linear models have been introduced. These include the ARCH and GARCH models ([Engle, 1982], [Bollerslev, 1986]), bi-linear models ([Subba Rao and Gabr, 1984], [Terdik, 1999]), RCA models ([Nicholls and Quinn, 1982]) and threshold models ([Tong, 1983]). The asymptotic null distributions of the portmanteau statistics obtained under i.i.d. assumptions on $\{X_t\}$ may not hold in the presence of non-linear dependence such as conditional heteroscedasticity. [Romano and Thombs, 1996] showed that the Box-Pierce statistic with χ^2 approximation may provide misleading inference for uncorrelated but dependent time series. [Francq et al., 2005] also showed that the Box-Pierce test, without suitable modification, applied to the residuals of an ARMA model with uncorrelated but dependent errors performs poorly.

[Hong, 1996] proposed a test statistic based on a distance between a kernel based spectral density estimator and the spectral density of the white noise. For the stationary process $\{X_t\}$ with autocorrelation function (1.1), the normalized spectral density function of X_t is given by (assuming $\sum_{k=0}^{\infty} |\rho(k)| < \infty$),

$$f(x) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ikx}, \quad \forall\ x \in [-\pi, \pi].$$

When X_t is a white noise, this reduces to

$$f(x) = \frac{1}{2\pi}, \quad \forall x \in [-\pi, \pi].$$

The lag window estimator of f ([Priestley, 1981]) is given by

$$\hat{f}_n(x) = \frac{1}{2\pi} \sum_{j=-n+1}^{n-1} K\left(\frac{j}{m_n}\right) \hat{\rho}(j)\, e^{ijx},$$

where $K(\cdot)$ is a non-negative symmetric kernel function and m_n is the bandwidth that depends on the sample size. With the quadratic distance, [Hong, 1996] proposed the test statistic $T_n = \pi \int_{-\pi}^{\pi} \left(\hat{f}_n(x) - \frac{1}{2\pi}\right)^2 dx$ which reduces to the equivalent form

$$T_n = \sum_{j=1}^{n} K^2\left(\frac{j}{m_n}\right) \hat{\rho}^2(j). \tag{1.4}$$

When $\{X_t\}$ is i.i.d. with finite 4-th order moment and $1/m_n + m_n/n \to 0$, [Hong, 1996] showed that

$$\frac{nT_n - C_n(K)}{\sqrt{2D_n(K)}} \xrightarrow{D} \mathcal{N}(0, 1) \tag{1.5}$$

where $C_n(K) = \sum_{j=1}^{n-1}(1-j/n)K^2(j/m_n)$, $D_n(K) = \sum_{j=1}^{n-2}(1-j/n)(1-(j+1)/n)K^4(j/m_n)$. Under some additional conditions on $K(\cdot)$ and m_n, (1.5) holds with $C_n(K)$ and $D_n(K)$ replaced by $m_nC(K)$ and $m_nD(K)$, respectively, i.e.,

$$\frac{nT_n - m_nC(K)}{\sqrt{2m_nD(K)}} \xrightarrow{D} \mathcal{N}(0,1) \tag{1.6}$$

where

$$C(K) = \int_0^\infty K^2(x)dx \quad \text{and} \quad D(K) = \int_0^\infty K^4(x)dx. \tag{1.7}$$

Later [Hong and Lee, 2003] established (1.5) assuming X_t to be a stationary martingale difference with finite 8-th order moment.

There are many white noise processes which are not martingale differences. For examples, one can consider

$$X_t = \varepsilon_{t-1}\varepsilon_{t-2} + \varepsilon_t \ \text{ where } \varepsilon_t \sim IID(0,1). \tag{1.8}$$

[Shao, 2011] showed that (1.6) is still valid under more general white noise assumption. He assumed that $\{X_t\}$ is a mean zero stationary causal process of the form

$$X_t = F(\ldots, \varepsilon_{t-1}, \varepsilon_t) \quad \forall t, \tag{1.9}$$

where ε_t are i.i.d. random variables and F is a measurable function for which X_t is well defined. Let $\{\varepsilon'_k\}_{k\in\mathbb{Z}}$ be an i.i.d. copy of $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ and

$$X'_t = F(\ldots, \varepsilon'_{-1}, \varepsilon'_0, \varepsilon_1, \ldots, \varepsilon_t) \tag{1.10}$$

be a coupled version of X_t. Say that $\{X_t\}$ satisfies the Geometric Moment Contraction (GMC(α)) condition ([Hsing and Wu, 2004], [Shao and Wu, 2007], [Wu and Shao, 2004]) for some $\alpha > 0$, if there exists $C > 0$ and $\eta = \eta(\alpha) \in (0,1)$ such that

$$E\left(|X_t - X'_t|^\alpha\right) \le C\eta^t, \quad n \in \mathbb{N}.$$

[Shao, 2011] established (1.6) under H_0 and the assumption that $\{X_t\}$ satisfies GMC(8) and $\log n = o(m_n)$, $m_n = o(n^{1/2})$. Further, under H_1, if GMC(8) holds and $1/m_n + m_n/n \to 0$, then [Shao, 2011] showed that

$$\frac{\sqrt{m_n}}{n}\left(\frac{nT_n - m_nC(K)}{\sqrt{2m_nD(K)}}\right) \xrightarrow{P} \frac{1}{\sqrt{8D(K)}}\sum_{j\neq 0}\rho^2(j). \tag{1.11}$$

[Hong and Lee, 2003] showed the same probability convergence under a different set of conditions but still assuming the finiteness of the 8-th moment of the process.

For many time series models, such as the bilinear model, various forms of GARCH models, the non-linear moving average model and the ARMA model (see [Hsing and Wu, 2005], [Shao and Wu, 2007], [Granger and Andersen, 1978], [Breidt et al., 2001], [Granger and Terasvirta, 1993]), the GMC(8) condition holds if the 8-the moment of X_t

is finite. As pointed out by a referee of [Shao, 2011], this condition is fairly strong and excludes some interesting GARCH models, such as the integrated GARCH model. In addition, the permissible parameter space for the regular GARCH(r, s) model becomes quite small under the 8-th moment assumption. Consider, for instance the ARCH model

$$X_t = \sigma_t \varepsilon_t \text{ and } \sigma_t^2 = c_0 + \sum_{i=1}^{p} b_i X_{t-i}^2, \tag{1.12}$$

where $c_0 \geq 0, b_i \geq 0, \ \forall i = 1, 2, \ldots, p$, are constants and $\{\varepsilon_t\} \sim IID(0, 1)$. In this case, for all i, the $2i$-th moment of X_t exists if and only if

$$\sum_{j=1}^{p} (b_j)^k < \frac{1}{E\,(\varepsilon_t)^{2k}}, \quad \forall k = 1, 2, \ldots, i. \tag{1.13}$$

Therefore, the finiteness of the 8-th moment imposes more restriction on the parameter space than finiteness of the lower order moments.

Moreover, finiteness of the 8-th moment may not guarantee GMC(8) in other situations. For example, suppose $\{X_t\}$ is a stationary linear process

$$X_t = \sum_{k=0}^{\infty} \theta_k \varepsilon_{t-k}, \quad \forall t \geq 1, \tag{1.14}$$

where $\{\varepsilon_t\}$ are i.i.d. with zero mean. Then (1.14) satisfies GMC(8) if and only if $\{\theta_k\}$ are exponentially decaying (see Subsection 4.3). Note that, any causal ARMA process can be expressed as (1.14) with exponentially decaying $\{\theta_k\}$ (see [Brockwell and Davis, 2009]) and hence GMC(8) holds provided $E(\varepsilon_t^8) < \infty$. But if $\{\theta_k\}$'s are polynomially decaying, then (1.14) does not satisfy GMC(8), even if $E(\varepsilon_t^8) < \infty$. Therefore, the results in [Shao, 2011] are not applicable for these linear processes.

In this paper, we take a different approach. Unlike [Hong, 1996], [Hong and Lee, 2003] and [Shao, 2011], we do not use any martingale arguments. We first assume the existence of all moments of X_t and summability of the cumulants of all orders and some mild conditions on the kernel. We use the fact that for any normal variable, all cumulants of order greater thantwo vanish. We establish the convergence (1.6) (see Theorem 2.1) by showing the convergence of all cumulants. In particular, the result is applicable to some non-linear moving average models (for example (1.8)), ARCH and various types of GARCH models and α-mixing stationary processes, with exponentially decaying mixing coefficient provided all moments are assumed to be finite.

We then focus on specific models and reduce the moment condition. By truncation arguments, we show that the existence of the 4-th moment is enough to obtain (1.6) for various types of GARCH and ARCH models, stochastic autoregressive volatility model and autoregressive conditional duration model. (1.6) also holds for the process (1.8) provided $E(\varepsilon_t^4) < \infty$.

Next we move to the alternate hypothesis. As mentioned earlier, [Hong and Lee, 2003] and [Shao, 2011] proved (1.11) under the alternative hypothesis H_1. We establish the joint distributional convergence of

$$\sqrt{n}\left(\hat{\gamma}^2(0) - \gamma^2(0)\right) \text{ and } \sqrt{n}(T_n - E(T_n))$$

to a bivariate normal distribution under the alternative hypothesis H_1 (see Theorem 2.2) and the same assumptions as in Theorem 2.1. Then again by truncating, we relax the moment conditions and prove that the finiteness of the 4-th moment is enough for the result to hold for ARMA, bilinear and non-linear moving average models. The result also holds for the polynomially decaying linear process, provided $E(\varepsilon_t^4) < \infty$.

Note that the convergence rates of T_n are different under null and alternative hypotheses. Moreover, the test statistic in [Shao, 2011] is still valid for most of the common processes under the existence of the 4-th moment only.

The main results of this paper are stated in Section 2 and the proofs of the theorems are given in Section 4. There are some simulation results in Section 3.

2 Main Results

In this section, we discuss joint convergence of suitably standardized $\hat{\gamma}^2(0)$ and T_n under the null and alternative hypothesis. First, we will discuss the necessary assumptions.

2.1 Assumptions

There are two main assumptions. One is on the cumulants of the process $\{X_t\}$ and the other is on the kernel $K(\cdot)$.

For a random vector $(Y_1, Y_2 \ldots, Y_l)$, its l-th order joint cumulant is given by

$$Cum(Y_1, Y_2, \ldots, Y_l) = \sum_V (-1)^{p-1}(p-1)!\Big(E\prod_{j\in V_1} Y_j\Big) \times \cdots \times \Big(E\prod_{j\in V_p} Y_j\Big), \tag{2.1}$$

where the summation is over all partitions $V = V_1 \cup V_2 \cup \ldots \cup V_p$ of size $p = 1, 2, \ldots, l$, of the index set $\{1, 2, \ldots, l\}$. Let us define, for all $k = 1, 2, \ldots$,

$$Q(t_1, t_2, \ldots t_k) = Cum\left(X_t, X_{t+t_1}, X_{t+t_2}, \ldots, X_{t+t_k}\right). \tag{2.2}$$

Note that, $Q(k) = \gamma(k)$ and we use this notation for $\gamma(k)$ for all $k = 0, \pm 1, \pm 2, \ldots$. We consider the following assumption on the process $\{X_t\}$.

Assumption 1. For any $k \geq 1$,

$$\sum_{\substack{t_i \in \mathbb{Z} \\ 1 \leq i \leq k}} |Q(t_1, t_2, \ldots, t_k)| < \infty. \tag{2.3}$$

Note that Assumption 1 implies $\sum_{j=1}^{\infty} |Q(j)| < \infty$. That in turn implies the existence of the spectral density. In the following remarks, we discuss models which satisfy Assumption 1 if all moments of the process are finite. In particular the linear process, defined in (1.14), always satisfies Assumption 1 provided all moments are finite. In Subsection 4.3 we will show that it does not necessarily satisfy the GMC condition.

Remark 2.1. *For a stationary Gaussian process, as all the cumulants of order greater than two are zero, Assumption* 1 *is trivially satisfied if* $\sum_{j=1}^{\infty} |Q(j)| < \infty$.

Remark 2.2. *If all moments of* $\{\varepsilon_t\}$ *are finite, then the linear process (1.14) satisfies Assumption* 1, *if*

$$\sum_{i=0}^{\infty} i|\theta_i| < \infty \tag{2.4}$$

To see this, note that for $0 \le t_1 \le t_2 \le \ldots \le t_k$, *we have*

$$\begin{aligned}
Q(t_1, t_2, \ldots, t_k) &= Cum(X_0, X_{t_1}, \ldots, X_{t_k}) \\
&= \Big(\frac{\partial}{\partial s_k \partial s_{k-1} \ldots \partial s_0} \log E\Big(e^{i\left(s_0 X_0 + s_1 X_{t_1} + \ldots + s_k X_{t_k}\right)}\Big)\Big)_{s_0 = s_1 = \ldots = s_k = 0} \\
&= \Big(\frac{\partial}{\partial s_k \partial s_{k-1} \ldots \partial s_0} \log E\Big(\prod_{r=0}^{\infty} e^{i\left(s_0 \theta_r + s_1 \theta_{t_1+r} + \ldots + s_k \theta_{t_k+r}\right)\varepsilon_r}\Big)\Big)_{s_0 = s_1 = \ldots = s_k = 0} \\
&= \Big(\frac{\partial}{\partial s_k \partial s_{k-1} \ldots \partial s_0} \sum_{r=0}^{\infty} \log E\Big(e^{i\left(s_0 \theta_r + s_1 \theta_{t_1+r} + \ldots + s_k \theta_{t_k+r}\right)\varepsilon_r}\Big)\Big)_{s_0 = s_1 = \ldots = s_k = 0} \\
&= \sum_{r=0}^{\infty} \Big(\frac{\partial}{\partial s_k \partial s_{k-1} \ldots \partial s_0} \log E\Big(e^{i\left(s_0 \theta_r + s_1 \theta_{t_1+r} + \ldots + s_k \theta_{t_k+r}\right)\varepsilon_r}\Big)\Big)_{s_0 = s_1 = \ldots = s_k = 0} \\
&= \sum_{r=0}^{\infty} Cum\left(\theta_r \varepsilon_r, \theta_{t_1+r}\varepsilon_r, \ldots, \theta_{t_k+r}\varepsilon_r\right) \\
&= K_r(\varepsilon_1)\Big(\sum_{r=0}^{\infty} \theta_r \theta_{t_1+r} \theta_{t_2+r} \ldots \theta_{t_k+r}\Big),
\end{aligned}$$

where $K_r(\varepsilon_1)$ *is the r-th order cumulant of* ε_1. *Hence, for some* $C > 0$, *we have*

$$|Q(t_1, t_2, \ldots, t_k)| \le C\Big(\sum_{r=0}^{\infty} |\theta_{t_1+r}|\Big) \ldots \Big(\sum_{r=0}^{\infty} |\theta_{t_k+r}|\Big).$$

Therefore, Assumption 1 *holds, since*

$$\sum_{u=0}^{\infty} \sum_{r=0}^{\infty} |\theta_{u+r}| = \sum_{r=0}^{\infty} (r+1)|\theta_r| < \infty.$$

Remark 2.3. *A strictly stationary process* $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$, *is said to be* α*-mixing if*

$$\alpha(n) \quad = \sup_{A \in \mathcal{F}_{-\infty}^{0},\ B \in \mathcal{F}_{n}^{\infty}} |P(A)P(B) - P(AB)| \to 0.$$

Here $\mathcal{F}_i^j$ *denotes the* σ*-algebra generated by* $\{X_t : i \le t \le j\}$ *and* $\alpha(n)$ *are called the* α*-mixing coefficients. We show later in Lemma 4.2 that, if a strictly stationary process* *is α-mixing with $\alpha(n) \le \alpha^n$, for some $0 < \alpha < 1$, then it satisfies Assumption 1 provided all moments are finite. An m-dependent process is α-mixing with finitely many non-zero $\alpha(n)$'s. In particular, the strictly stationary non-linear moving average processes,*

$$X_t = \beta\varepsilon_{t-1}\varepsilon_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim IID(0,\sigma^2), \;\; \forall t \ge 0, \;\; \beta \in \mathbb{R}, \tag{2.5}$$

$$X_t = \beta\varepsilon_{t-1}\varepsilon_t + \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim IID(0,\sigma^2), \;\; \forall t \ge 0, \;\; \beta \in \mathbb{R}, \tag{2.6}$$

satisfy Assumption 1 if all moments of $\{\varepsilon_t\}$ are finite.

Remark 2.4. *A strictly stationary process $\{X_t\}$ is called a β-mixing process if*

$$\beta(n) = E\Big(\sup_{B\in\mathcal{F}_n^\infty} |P(B) - P(B|X_0, X_{-1}, X_{-2}, \ldots)|\Big) \to 0.$$

$\beta(n)$ are known as the β-mixing coefficients. A β- mixing process is always an α-mixing process with $\alpha(n) \le \beta(n)$. The strictly stationary bilinear (p, q, P, Q) process,

$$X_t = \sum_{j=1}^{p} b_j X_{t-j} + \varepsilon_t + \sum_{k=1}^{q} a_k\varepsilon_{t-k} + \sum_{j=1}^{P}\sum_{k=1}^{Q} c_{jk}X_{t-j}\varepsilon_{t-k}, \;\; \varepsilon_t \sim IID(0,\sigma^2), \tag{2.7}$$

is a β-mixing process with exponentially decaying β-mixing coefficient (see [Fan and Yao, 2003]). Therefore, it satisfies Assumption 1 provided all its moments are finite.

Another important model in non-linear time series is the power GARCH (p, q, δ),

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \;\; t = 0, 1, 2, \ldots, \quad \varepsilon_t \sim IID(0,\sigma^2), \\ h_{t+1}^\delta &= \omega + \sum_{i=1}^{q} \alpha_i h_{t+1-i}^\delta \varepsilon_{t+1-i}^{2\delta} + \sum_{j=1}^{p} \beta_j h_{t+1-j}^\delta, \end{aligned} \tag{2.8}$$

where $\delta > 0$ fixed. Let us consider the following condition for this process.

Condition β: *(i) $\{\varepsilon_t\}$ are independent of h_0 with a positive continuous density,*

(ii) $E\,(\varepsilon_t)^{2\delta} < \infty$,

(iii) $\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$ and

(iv) h_0 is initialized from the invariant measure.

[Carrasco and Chen, 2002] showed that if Condition β holds, then $\{(X_t, h_t)\}$ is strictly stationary, β-mixing with exponentially decaying β-coefficients. Therefore, $\{X_t\}$ satisfies Assumption 1, provided all its moments are finite.

Now, we will discuss the assumptions on the kernel $K(\cdot)$.

Assumption 2. $K(\cdot)$ is continuous, even, non-increasing on $[0,\infty)$, square integrable with unique maximum value at 0 and $K(0) = 1$.

Many standard kernels, for example the well known kernels of Hamming, Hann and Bartlett, satisfy Assumption 2. [Shao, 2011] assumed differentiability of $K(\cdot)$ except at finite number of points on its support. Here, we do not need this assumption. Note that as $K(\cdot)$ is non-increasing, we have,

$$\int_0^{n/m_n} K^2(x)dx \;\le\; \frac{1}{m_n}\sum_{j=0}^{n} K^2\left(\frac{j}{m_n}\right) \;\le\; \int_0^{n/m_n} K^2(x)dx + K^2(0)\frac{1}{m_n}. \tag{2.9}$$

Therefore, if $m_n^{-1} + n^{-1}m_n \to 0$, then

$$\lim_{n\to\infty} \frac{1}{m_n}\sum_{j=0}^{n} K^2\left(j/m_n\right) = \int_0^{\infty} K^2(x)dx. \tag{2.10}$$

We will need this later in the proof of the Theorems.

2.2 Null Distribution

The following theorem provides us the joint convergence of suitably standardized $\hat{\gamma}^2(0)$ and T_n under the null hypothesis.

Theorem 2.1. *If Assumptions* 1 *and* 2 *hold and* $m_n^{-1} + n^{-1}m_n \to 0$ *then under* H_0*,*

$$\Big(\sqrt{n}(\hat{\gamma}^2(0) - E(\hat{\gamma}^2(0))), \frac{n}{\sqrt{m_n}}\Big(T_n - E(T_n)\Big)\Big) \xrightarrow{\mathcal{D}} \mathcal{N}_2(0, \Sigma_0), \tag{2.11}$$

where

$$\Sigma_0 = Q^2(0)\begin{bmatrix} 4\sum_{u=-\infty}^{\infty} Cov(X_0^2,\ X_u^2) & 0 \\ 0 & 2Q^{-2}(0)D(K) \end{bmatrix} \tag{2.12}$$

and $C(K)$, $D(K)$ *are as in (1.7).*

Remark 2.5. *By (2.9), if the kernel* $K(\cdot)$ *is compactly supported, then we have*

$$\begin{aligned} & m_n^{-1/2} nE\Big(\sum_{j=1}^{\infty} K^2(\frac{j}{m_n})\hat{\gamma}^2(j)\Big) = (n\sqrt{m_n})^{-1}\sum_{t_1,t_2}\sum_{j} K^2(\frac{j}{m_n})E(X_{t_1}X_{t_1+j}X_{t_2}X_{t_2+j}) \\ &= \sqrt{m_n}Q^2(0)(C(K) + O(m_n^{-1})) = \sqrt{m_n}Q^2(0)C(K) + O(m_n^{-1/2}), \end{aligned}$$

where $C(K)$ *is given by (1.7). Also, by Assumption* 1*, we have,*

$$\begin{aligned} \sqrt{n}E(\hat{\gamma}^2(0)) &= n^{-3/2}\sum_{t_1,t_2} Cov(X_{t_1}^2, X_{t_2}^2) + \sqrt{n}Q^2(0) \\ &= n^{-1/2}\sum_{u\neq 0}(\frac{n-u}{n})Q(0,u,u) + n^{-3/2}\sum_{u} Var(X_u^2) + \sqrt{n}Q^2(0) \\ &= O(n^{-1/2}) + \sqrt{n}Q^2(0). \end{aligned}$$

Therefore, by (2.11), we have

$$\left(\sqrt{n}(\hat{\gamma}^2(0) - Q^2(0)), \frac{1}{\sqrt{m_n}}\left(nT_n - m_n C(K)\right)\right) \xrightarrow{\mathcal{D}} \mathcal{N}_2(0, \Sigma_0). \tag{2.13}$$

The asymptotic distribution of the second component agrees with that in [Shao, 2011].

Remark 2.6. *Consider the power GARCH model (2.8) for which Condition β holds. If all moments are finite then Theorem 2.1 applies and (2.13) holds. [Shao, 2011] required only the* 8*-th moment to be finite for (2.13) to hold. However, in Subsection 4.4 (a), we show that (2.13) continues to hold if only* 4*-th moment is assumed to be finite. This reduction in moment is important, since as discussed in Section 1, finiteness of* 8*-th moment restricts the parameter space.*

Remark 2.7. *Consider any process $\{X_t\}$ of the form*

$$X_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma^2), \quad t \geq 0 \tag{2.14}$$

where σ_t is $\mathcal{F}_{-\infty}^{t-1} = \sigma\{X_i : i \leq t-1\}$ measurable. The truncation argument in Subsection 4.4 (a), used for power GARCH model also works here. Therefore, (2.13) holds if $\{X_t\}$ is exponentially decaying α-mixing and has finite 4*-th moment. There are widely used non-linear time series models, such as other GARCH models, stochastic volatility models and autoregressive conditional duration models, which are of the form (2.14). We discuss a few such models and their mixing properties below, following [Carrasco and Chen, 2002].*

Example 1. Stochastic Autoregressive Volatility Model: Cnsider a log AR(1) version of the stochastic autoregressive volatility model introduced by [Anderson, 1994]:

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \log \sigma_t &= \omega + \beta \log \sigma_{t-1} + (\gamma + \alpha \log \sigma_{t-1})\, u_t, \quad \forall t \geq 0. \end{aligned} \tag{2.15}$$

where $\{\varepsilon_t\}$ and $\{u_t\}$ are mutually independent, i.i.d. random variables with zero means and unit variances, $\alpha + \beta > 0$ and $\alpha + \gamma > 0$. If $\alpha = 0$ and if ε_t and u_t are i.i.d. normally distributed, then (2.15) becomes the lognormal stochastic volatility model. Assume that

(a) (i) of *Condition* β holds,
(b) $|\beta| < 1$,
(c) there is an integer $s \geq 1$ such that $E|u_t|^s < \infty$, $E|\beta + \alpha u_t|^s < 1$,
(d) $\{\sigma_t\}$ is initialized from its stationary distribution.

Then $\{(X_t, \sigma_t)\}$ is strictly stationary and exponential β-mixing.

Example 2. (Autoregressive Conditional Duration) This model was proposed by [Engle and Russell, 1998] for high frequency data from financial markets:

$$X_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma^2), \quad \forall t \geq 0,$$

$$\sigma_t = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i} + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \tag{2.16}$$

with $\omega > 0$, $\alpha_i \geq 0$, $i = 1, 2, \ldots, q$, $\beta_j \geq 0$, $j = 1, 2, \ldots, p$. Consider the following conditions:

(i) ε_t has density with support $[0, \infty)$,

(ii) $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$,

(iii) (X_0, σ_0) is initialized from its stationary distribution.

Then $\{(X_t, \sigma_t)\}$ is strictly stationary and exponential β-mixing.

Example 3. (LGARCH and NGARCH) Linear GARCH model of [Bollerslev, 1986] and non-linear asymmetric GARCH model of [Engle and Ng, 1993] are given by

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma^2), \quad t \geq 0, \\ h_t &= \omega + \beta h_{t-1} + \alpha(\varepsilon_{t-1} - c)^2 h_{t-1}, \end{aligned} \tag{2.17}$$

with $\omega > 0$, $\beta \geq 0$ and $\alpha \geq 0$. Here $c = 0$ corresponds to the linear GARCH(1, 1) model. Suppose (i) of *Condition* β holds and there is an integer $s \geq 1$ such that

$$\begin{array}{ll} \text{either} & E\left(\beta + \alpha(\varepsilon_t - c)^2\right)^s < 1 \\ \text{or} & E|\varepsilon_t|^{2s} < \infty, \text{ and } \beta + \alpha < \dfrac{1}{\left(E[(\varepsilon_t - c)^{2s}]\right)^{1/s}}. \end{array}$$

Further suppose that h_0 is initialized from its stationary distribution. Then $\{(X_t, h_t)\}$ is strictly stationary and exponential β-mixing.

Example 4. (MGARCH) This model of [Geweke, 1986] is defined as

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma^2), \quad t \geq 0, \\ \log h_t &= \omega + \beta \log h_{t-1} + \alpha \log \varepsilon_{t-1}^2. \end{aligned} \tag{2.18}$$

Consider (i) of *Condition* β and suppose that there is an integer $s \geq 1$ such that $|\beta| < 1$ and $E|\log \varepsilon_t^2|^s < \infty$. Moreover, h_0 is initialized from its invariant measure. Then $\{(X_t, h_t)\}$ is strictly stationary and exponential β-mixing.

Example 5. (EGARCH model) This model of [Nelson, 1991] is defined as

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma^2), \quad t \geq 0, \\ \log h_t &= \omega + \beta \log h_{t-1} + \alpha\left(|\varepsilon_{t-1}| + \gamma \varepsilon_{t-1}\right), \end{aligned} \tag{2.19}$$

where $\gamma \neq 0$. Consider (i) of *Condition* β and suppose that there is an integer $s \geq 1$ such that $|\beta| < 1$ and $E|\varepsilon_t|^s < \infty$. Moreover, h_0 is initialized from its invariant measure. Then $\{(X_t, h_t)\}$ is strictly stationary and exponential β-mixing.

Example 6. (VGARCH) This model of [Engle and Ng, 1993] is defined as

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \;\; \varepsilon_t \sim IID(0,\sigma^2), \;\; t \geq 0, \\ h_t &= \omega + \beta h_{t-1} + \alpha\left(|\varepsilon_{t-1}| - c\right)^2, \end{aligned} \tag{2.20}$$

where $\omega > 0, \beta \geq 0$ and $\alpha \geq 0$. Consider (i) of *Condition* β and suppose that there is an integer $s \geq 1$ such that $|\beta| < 1$ and $E|\varepsilon_t|^{2s} < \infty$. Moreover, h_0 is initialized from its invariant measure. Then $\{(X_t, h_t)\}$ is strictly stationary and exponential β-mixing.

Example 7. (GJR-GARCH) This model of [Glosten et al., 1993] is defined as

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, \;\; \varepsilon_t \sim IID(0,\sigma^2), \;\; t \geq 0, \\ h_t &= \omega + \beta h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 h_{t-1} + \alpha_2\left(\max 0, -\varepsilon_{t-1}\right)^2 h_{t-1}, \end{aligned} \tag{2.21}$$

where $\omega > 0, \beta \geq 0,\ \alpha_1 > 0$ and $\alpha_1 + \alpha_2 \geq 0$. Consider (i) of *Condition* β and suppose that there is an integer $s \geq 1$ such that

$$E\left(\beta + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2\left(\max 0, -\varepsilon_{t-1}\right)^2\right)^s < 1.$$

Moreover, h_0 is initialized from its invariant measure. Then $\{(X_t, h_t)\}$ is strictly stationary and exponential β-mixing.

2.3 Alternative distribution

Next, we will discuss the joint convergence of suitably standardized $\hat{\gamma}^2(0)$ and T_n under alternative hypothesis. Let us define,

$$\begin{aligned} Q * Q(u) &= \sum_{j=1}^{\infty} Q(j)Q(u+j), \\ \sigma_{00} &= 4Q^2(0)\Big(2\sum_{u=-\infty}^{\infty} Q^2(u) + \sum_{u=-\infty}^{\infty} Q(0,u,u)\Big), \\ \sigma_{01} &= 4Q(0)\Big(3\sum_{u=-\infty}^{\infty} Q(u)Q*Q(u) + \sum_{u=-\infty}^{\infty} Q(u)Q*Q(-u) \\ &\quad + \sum_{j=1}^{\infty}\sum_{u=-\infty}^{\infty} Q(j)Q(0,u,u+j) + \sum_{j=1}^{\infty}\sum_{u=-\infty}^{\infty} Q(j)Q(j,u,u)\Big), \\ \sigma_{11} &= 2\sum_{u=-\infty}^{\infty}\left[Q*Q(u) + Q*Q(-u)\right]^2 + 4\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\sum_{u=-\infty}^{\infty} Q(i)Q(j)Q(i,u,u+j), \end{aligned}$$

$$\Sigma_1 = \begin{bmatrix} \sigma_{00} & Q^{-2}(0)\sigma_{01} \\ Q^{-2}(0)\sigma_{01} & Q^{-4}(0)\sigma_{11} \end{bmatrix}. \tag{2.22}$$

Theorem 2.2. *Suppose Assumptions* 1 *and* 2 *hold and* $m_n^{-1} + n^{-1}m_n \to 0$*. Then*

$$\Big(\sqrt{n}\Big(\hat{\gamma}^2(0) - E(\hat{\gamma}^2(0))\Big),\ \sqrt{n}\Big(T_n - E\Big(T_n\Big)\Big)\Big) \xrightarrow{\mathcal{D}} \mathcal{N}_2(0, \Sigma_1). \tag{2.23}$$

Remark 2.8. *Theorem 2.2 implies (1.11), given in [Shao, 2011] since*

$$\begin{aligned} E(T_n) &= n^{-2}Q^{-2}(0) \sum_{j,t_1,t_2} K^2(\frac{j}{m_n})\Big(Q(j, t_2 - t_1, t_2 - t_1 + j) + Q^2(j) + Q^2(t_2 - t_1) \\ &\qquad + Q(t_2 - t_1 + j)Q(t_2 - t_1 - j)\Big) \\ &= O(n^{-1}) + \sum_{j=1}^{n} K^2(\frac{j}{m_n})\rho^2(j) + O(n^{-1}m_n) + O(n^{-1}) \to \frac{1}{2}\sum_{j\neq 0}\rho^2(j) \end{aligned} \tag{2.24}$$

and hence,

$$\frac{\sqrt{m_n}}{n}\left(\frac{nT_n - m_nC(K)}{\sqrt{m_n}}\right) = \frac{1}{\sqrt{n}}\Big(\sqrt{n}(T_n - E(T_n))\Big) + E(T_n) - \frac{m_n}{n}C(K) \to \frac{1}{2}\sum_{j\neq 0}\rho^2(j). \tag{2.25}$$

Remark 2.9. *Also, observe that*

$$\begin{aligned} E(\hat{\gamma}^2(0)) &= n^{-2}\sum_{t_1,t_2} E(X_{t_1}^2 X_{t_2}^2) = n^{-2}\sum_{t_1,t_2}\Big(Q(0, t_2 - t_1, t_2 - t_1) + Q^2(0) + 2Q^2(t_2 - t_1)\Big) \\ &= Q^2(0) + O(n^{-1}). \end{aligned} \tag{2.26}$$

Hence, from (2.24) and (2.26), if $m_n = o(\sqrt{n})$*, then we have*

$$\Big(\sqrt{n}\Big(\hat{\gamma}^2(0) - Q^2(0)\Big),\ \sqrt{n}\Big(T_n - \sum_{j=1}^{n} K^2(\frac{j}{m_n})\rho^2(j)\Big)\Big) \xrightarrow{\mathcal{D}} \mathcal{N}_2(0, \Sigma_1). \tag{2.27}$$

Remark 2.10. *We have already shown that the process (1.14), satisfies Assumption* 1 *if all moments of* ε_t *are finite and condition (2.4) holds. Hence (2.23) holds. In Subsection 4.4 (b), we show that if* $E(\varepsilon_1^4) < \infty$, $m_n = O(\sqrt{n})$ *and condition (2.4) holds, then (2.23) continues to hold for the process (1.14). In Section 4.4 (c) we show that, under the same conditions, (2.23) holds for the non-linear moving average processes (2.5) and (2.6).*

Remark 2.11. *As discussed in Remark 2.4, a strictly stationary bilinear model is a* β*-mixing process with exponentially decaying mixing coefficient (see [Fan and Yao, 2003]). In general, all processes of the form (2.7) are not strictly stationary. For the process (2.7), it is known that ([Pham, 1985], [Fan and Yao, 2003])*

$$X_t = h^T Z_{t-1} + \varepsilon_t, \quad Z_t = A(\varepsilon_t)Z_{t-1} + c(\varepsilon_t),$$

where Z_t *is a random vector,* h *is a constant vector, and* $A(\cdot)$ *and* $c(\cdot)$ *are constant matrix and vector functions, respectively. Then (2.7) has a unique strictly stationary solution if (see [Fan and Yao, 2003])*

(i) *the upper Lyapunov exponent of the sequence* $\{A(\varepsilon_t)\}$ *is negative,*

(ii) $E\left(\|c(\varepsilon_t)\|^2\right) < \infty$, *where* $\|\cdot\|$ *is the usual Euclidean norm of a vector.*

Therefore, if (i) *and* (ii) *hold and all order moments of the process (2.7) are finite, then (2.23) holds for the process (2.7). Also, under strict stationary,* $\{X_t\}$ *can be expressed as an ARMA*(p, q') *process (see [Fan and Yao, 2003]), where* $q' = \max(q, Q)$*. Therefore, by Remark* 2.10*, if* (i) *and* (ii) *holds and* $E(X_t^4) < \infty$*, then (2.23) holds for the process (2.7).*

3 Simulation

We report simulations for six different time series models, whose 4-th moment is finite but 8-th moment does not exist. In all these cases, simulations strongly support the joint asymptotic normality stated in Theorems 2.1 and 2.2. We also present empirical powers for the test statistic $\frac{nT_n - m_n C(K)}{\sqrt{2m_n D(K)}}$ at 5% level of significance which demonstrate the consistency of the test.

Asymptotic Normality: We choose the Hamming kernel (see [Harris, 1978])

$$K(x) = \frac{1}{2}\left(1 + \cos(\pi x)\right) I_{[-1,1]}(x)$$

and consider 5000 runs each with sample size $n = 1000$ and relatively small $m_n = \sqrt{n}$. In each case, we present the bivariate kernel density estimate of the bivariate statistic and its contours.

Null hypothesis. We consider two models as follows.

(1) **IID process**: $X_t = \varepsilon_t^{-1}$, where $\varepsilon_t \sim \text{Beta}(5, 1)$. Note that $E(\varepsilon_t^r) < \infty$ if and only if $r \leq 4$. Figure 4.1 (a) and (b) respectively represent the bivariate kernel density estimate of the bivariate statistic given in (2.11) and its contours for model (1). The eilliptic contours and the bell shaped density estimate clearly indicate asymptotic normality. Moreover, the axes of the elliptical contours are parallel to the X and Y-axes, indicating the asymptotic independence between $\hat{\gamma}^2(0)$ and T_n. The empirical product moment correlation coefficient was 0.002 and the empirical level at 5% level of significance turned out to be 0.0489.

(2) **ARCH process**: $X_t = \sigma_t \varepsilon_t,\ \sigma_t^2 = 1 + 0.5X_{t-1}^2$, where $\varepsilon_t \sim \mathcal{N}(0, 1)$.

Note that in this case, by (1.13), as $\left(E(\varepsilon_t^8)\right)^{-1/4} < 0.5 < \left(E(\varepsilon_t^4)\right)^{-1/2}$, the 4-th moment of the process is finite whereas the 8-th moment is not. Figure 4.2 (a) and (b) respectively represent the bivariate kernel density estimate of the bivariate statistic given in (2.11) and its contour for model (2). Similar to (1), these plots also indicate the asymptotic normality and asymptotic independence of $\hat{\gamma}^2(0)$ and T_n. The empirical correlation coefficient was 0.006 and the empirical level at 5% level of significance was 5.8%.

Alternative hypothesis. We consider the following four models:

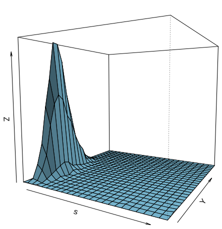
(3) **ARMA process**: $X_t - 0.89X_{t-1} - 0.49X_{t-2} + 0.45X_{t-3} = \varepsilon_t - 0.23\varepsilon_{t-1} + 0.25\varepsilon_{t-2}$,

(4) **Non-linear Moving Average process**: $X_t = 0.8\varepsilon_{t-1}\varepsilon_t + \varepsilon_{t-1} + \varepsilon_t$,

(5) **Bilinear process**: $X_t = 0.2X_{t-1}\varepsilon_{t-1} + \varepsilon_t$,

(6) **AR**(1)**-ARCH**(1) **process**: $X_t = 0.8X_{t-1} + \sigma_t\varepsilon_t$, $\sigma_t^2 = 1 + 0.7X_{t-1}^2$,

where in all cases, $\{\varepsilon_t\}$ is as (1). In all cases, the elliptical contours and bell shaped density estimates indicate asymptotic normality. However, unlike (1) and (2), the axes of the elliptical contours are not parallel to the X and Y axes. This indicates that $\hat{\gamma}^2(0)$ and T_n are not asymptotically independent. Empirical correlation coefficients for these models were 0.34, 0.2, 0.29 and 0.38 respectively.



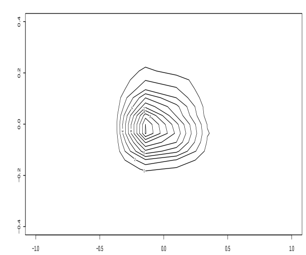
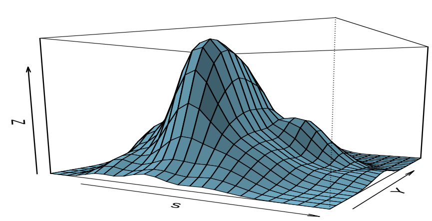


Figure 4.1 (*a*): *IID process: density estimate* (*b*) *IID process: contour plot*



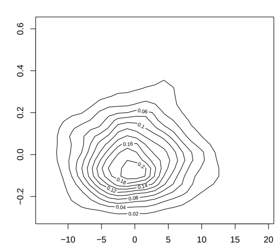
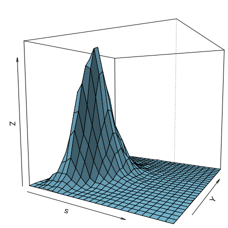


Figure 4.2 (*a*): *ARCH process: density estimate* (*b*) *ARCH process: contour plot*



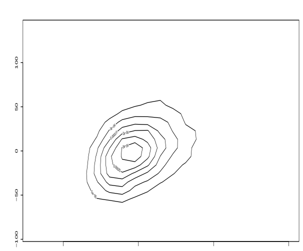


Figure 4.3 (*a*): *ARMA process: density estimate* (*b*) *ARMA process: contour plot*

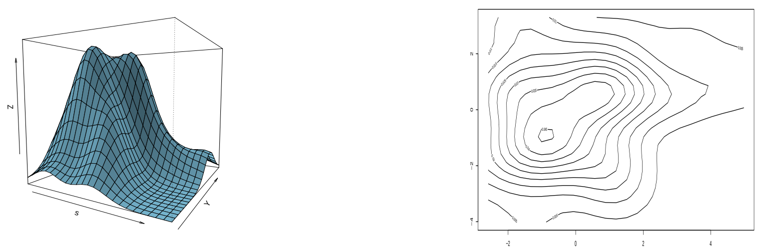


Figure 4.4 (*a*)*: Non-linear MA: density estimate* (*b*) *Non-linear: contour plot*

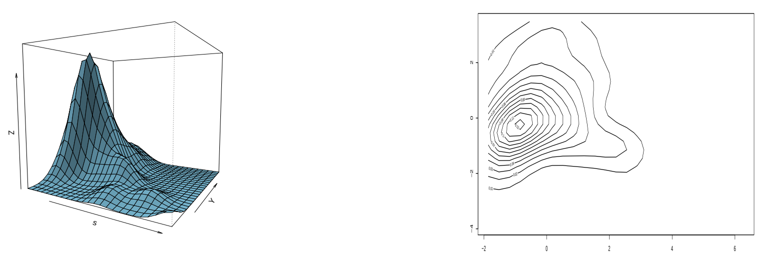


Figure 4.5 (*a*)*: Bilinear process: density estimate* (*b*) *Bilinear process: contour plot*

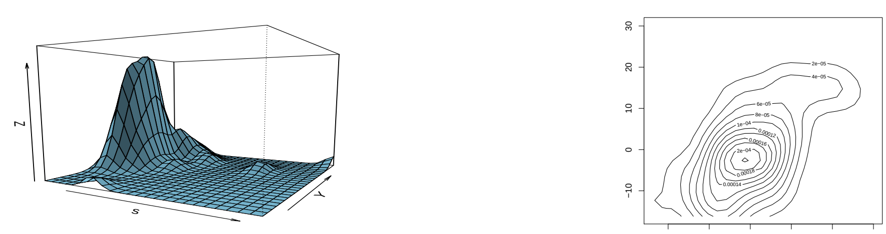


Figure 4.6 (*a*) *AR*(1)-*ARCH*(1)*: density estimate* (*b*) *AR*(1)-*ARCH*(1)*: contour plot*

Empirical power: We now present the empirical powers of the test statistic $\frac{nT_n - m_nC(K)}{\sqrt{2m_nD(K)}}$ at 5% level of significance for models (3)–(6). For each model, we consider two cases: (a) ε_t as in (1) and (b) $\varepsilon_t \sim \mathcal{N}(0,1)$. Note that in ($a$), moments of order greater than 4 do not exist but in (b), all moments of the process exist except in model (6). We choose the same kernel as before. We consider 10000 runs each with sample size $n = 100, 500, 1000, 5000$ and choose $m_n = n^{1/3}$. The following table presents the empirical powers. Figures in brackets are for (b).

Table 1: Empirical power of the test based on 10000 simulation runs

Model	$n = 100$	$n = 500$	$n = 1000$	$n = 5000$
(3)	0.41 (0.38)	0.84 (0.79)	0.99 (0.96)	1 (1)
(4)	0.49 (0.54)	0.89 (0.79)	0.95 (0.97)	1 (1)
(5)	0.33 (0.29)	0.75 (0.79)	0.97 (0.94)	1 (1)
(6)	0.38 (0.32)	0.96 (0.92)	0.97 (0.98)	1 (1)

For each model the power gets close to 1 as the sample size increases. This indicates the consistency of the test. Also, the test performs equally good in cases (a) and (b).

Remark 3.1. *While our results are asymptotically valid, the finite sample accuracy of the test is influenced by the choice of the kernel K and the bandwidth $\{m_n\}$. In probability density estimation, generally, the flat-top kernels e.g. trapezoid lag window, have the best rate of convergence and also appropriate bandwidths are determined in [Politis, 2003]. These choices may continue to perform well in the present case. This needs detailed investigation.*

Acknowledgement. We thank Yoon-Jin Lee for providing us with a copy of their unpublished technical report [Hong and Lee, 2003] and Dimitris Politis for his comments.

4 Appendix

In this section, we will discuss all the necessary lemmas and proofs. In Section 4.1, we prove Theorems 2.1 and 2.2. In Section 4.2, we discuss a lemma which is used to show that Assumption 1 holds for many time series models. In Section 4.3 we show that $\{X_t\}$ defined in (1.14), has the GMC property only if $\{\theta_k\}$ is exponentially decreasing, as pointed out in Section 1. Finally, in Section 4.4 we present the truncation arguments for various models to reduce the moment condition.

4.1 Proofs of Theorem 2.1 and 2.2:

We will prove the theorems using Cramer-Wold device. Let us define

$$\sigma_0^2 = 4t_1^2 Q^2(0) \sum_{u=-\infty}^{\infty} Cov(X_0^2, X_u^2) + 2t_2^2 Q^4(0) D(K), \tag{4.1}$$

$$\sigma_1^2 = t_1^2 \sigma_{00} + 2t_1 t_2 \sigma_{01} + t_2^2 \sigma_{11}, \tag{4.2}$$

where σ_{00}, σ_{01} and σ_{11} are as in (2.22). Note that (4.3) and (4.4) imply $\hat{\gamma}(0) \xrightarrow{P} Q(0)$ under null and alternative hypothesis respectively. Therefore, by Slutsky's theorem, it is enough to prove

$$A_n = t_1 \sqrt{n}\Big(\hat{\gamma}^2(0) - E(\hat{\gamma}^2(0))\Big) + t_2 \frac{n}{\sqrt{m_n}}\Big(\hat{\gamma}^2(0) T_n - E(\hat{\gamma}^2(0) T_n)\Big) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_0^2), \tag{4.3}$$

$$B_n = t_1 \sqrt{n}\Big(\hat{\gamma}^2(0) - E(\hat{\gamma}^2(0))\Big) + t_2 \sqrt{n}\Big(\hat{\gamma}^2(0) T_n - E(\hat{\gamma}^2(0) T_n)\Big) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_1^2), \tag{4.4}$$

respectively for Theorems 2.1 and 2.2. Let us define, for all $k = 2, 3, \ldots,$

$$C_k = Cum(X_{i_1} X_{i_1+j_1} X_{i_2} X_{i_2+j_1},\ X_{i_3} X_{i_3+j_2} X_{i_4} X_{i_4+j_2},\ \ldots,\ X_{i_{2k-1}} X_{i_{2k-1}+j_k} X_{i_{2k}} X_{i_{2k}+j_k}),$$

and

$$\delta_{0n}(x) = \begin{cases} t_1 & \text{if} \quad x = 0, \\ t_2 \sqrt{\frac{n}{m_n}} & \text{if} \quad x > 0, \end{cases} \qquad \delta_0(x) = \begin{cases} t_1, & x = 0 \\ t_2, & x > 0. \end{cases}$$

By (2.1), for $k \geq 2$, the k-th order cumulant of A_n and B_n respectively as,

$$\kappa_{n,k} = \frac{1}{n^{3k/2}} \sum_{j_l=0}^{n} \sum_{\substack{i_{2l-1}=1 \\ 1 \leq l \leq k}}^{n-j_l} \sum_{i_{2l}=1}^{n-j_l} \Big[\prod_{l=1}^{k} K^2(\frac{j_l}{m_n}) \delta_{0n}(j_l)\Big] C_k, \tag{4.5}$$

$$\kappa'_{n,k} = \frac{1}{n^{3k/2}} \sum_{j_l=0}^{n} \sum_{\substack{i_{2l-1}=1 \\ 1 \leq l \leq k}}^{n-j_l} \sum_{i_{2l}=1}^{n-j_l} \Big[\prod_{l=1}^{k} K^2\Big(\frac{j_l}{m_n}\Big) \delta_0(j_l)\Big] C_k. \tag{4.6}$$

As all Gaussian cumulants of order greater than two vanish, to prove the theorems, it is enough to show

$$\begin{aligned} &(a) \lim_{n\to\infty} \kappa_{n,2} = \sigma_0^2, \qquad (b) \lim_{n\to\infty} \kappa_{n,k} = 0, \quad \forall k \geq 3, \\ &(c) \lim_{n\to\infty} \kappa'_{n,2} = \sigma_1^2, \qquad (d) \lim_{n\to\infty} \kappa'_{n,k} = 0, \quad \forall k \geq 3. \end{aligned} \tag{4.7}$$

To calculate these cumulants, we need some analysis on indecomposable partitions.

Indecomposable partitions: Consider a two-way table

$$\begin{array}{cccc} (1,1) & (1,2) & \dots & (1,J) \\ (2,1) & (2,2) & \dots & (2,J) \\ \vdots & & & \vdots \\ (I,1) & (I,2) & \dots & (I,J) \end{array} \tag{A}$$

and a partition $P_1 \cup P_2 \cup \ldots \cup P_M$ of its entries. Two sets $P_{m'}$ and $P_{m''}$, of the partition, are called *hooked* if there exist $(i_1, j_1) \in P_{m'}$ and $(i_2, j_2) \in P_{m''}$ such that $i_1 = i_2$. Two sets $P_{m'}$ and $P_{m''}$ *commmunicate* if there exists a sequence of sets $P_{m_1} = P_{m'}, P_{m_2}, P_{m_3}, \ldots, P_{m_N} = P_{m''}$, $N \leq M$, such that P_{m_i} and $P_{m_{i+1}}$ are hooked for $i = 1, 2, \ldots, N-1$. A partition is called *indecomposable* if all sets communicate.

We need the following lemma to compute joint cumulants of polynomials of random variables.

Lemma 4.1. *[Brillinger, 2001] Consider a two-way array of random variables X_{ij}, $j = 1, 2, \ldots, J$ and $i = 1, 2, \ldots, I$. Consider I random variables,*

$$Y_i = \prod_{j=1}^{J} X_{ij}, \quad i = 1, 2, \ldots, I.$$

The joint cumulant $Cum\,(Y_1, Y_2, \ldots, Y_I)$ is then given by

$$\sum_{V} \Big[\prod_{k=1}^{p} Cum\Big(X_{ij} : (i, j) \in V_k\Big) \Big]$$

where the sum is over all indecomposable partitions $V = \cup_{i=1}^{p} V_i$ of Table (A).

$\{V_i : i = 1, 2, \ldots, p\}$ are called blocks of the partition V. Now to compute cumulants (4.5) and (4.6), Table (A) reduces to

$$\begin{array}{cccc} (1, i_1) & (1, i_1 + j_1) & (1, i_2) & (1, i_2 + j_1) \\ (2, i_3) & (2, i_3 + j_2) & (2, i_4) & (2, i_4 + j_2) \\ \vdots & & & \vdots \\ (k, i_{2k-1}) & (k, i_{2k-1} + j_k) & (k, i_{2k}) & (k, i_{2k} + j_k). \end{array} \tag{B}$$

Let $\mathcal{D}_0$ be the set of all indecomposable partitions of Table (B) and let $\mathcal{D}$ be the set of derived partitions after dropping the first index. Then by Lemma 4.1,

$$C_k = \sum_{\substack{V = V_1 \cup V_2 \cup \ldots \cup V_p \\ V \in \mathcal{D}}} \Big[\prod_{i=1}^{p} Cum\,(X_t : t \in V_i) \Big]. \tag{4.8}$$

Provided the limits exist, for all $V \in \mathcal{D}$, let us define,

$$\tau_{V,k} = \lim_{n\to\infty} \frac{1}{n^{3k/2}} \sum_{j_l=0}^{n} \sum_{\substack{i_{2l-1}, i_{2l}=1 \\ 1\leq l\leq k}}^{n-j_l} \Big[\prod_{l=1}^{k} K^2(\frac{j_l}{m_n})\delta_{0n}(j_l) \Big] \Big[\prod_{i=1}^{p} Cum\,(X_t : t \in V_i) \Big], \tag{4.9}$$

$$\tau'_{V,k} = \lim_{n\to\infty} \frac{1}{n^{3k/2}} \sum_{j_l=0}^{n} \sum_{\substack{i_{2l-1},i_{2l}=1 \\ 1\le l\le k}}^{n-j_l} \Big[\prod_{l=1}^{k} K^2(\frac{j_l}{m_n})\delta_0(j_l)\Big]\Big[\prod_{i=1}^{p} Cum\,(X_t : t \in V_i)\Big]. \quad (4.10)$$

Let, $\mathcal{D}'$ be the set of all partitions having at least one singleton block. Since $\{X_t\}$ has zero mean, $\tau_{V,k} = 0,\ \forall V \in \mathcal{D}'$. Then, by (4.5), (4.6) and (4.8), for all $k \ge 2$, we have

$$\lim_{n\to\infty} \kappa_{n,k} = \sum_{\substack{V=V_1\cup V_2\cup\ldots\cup V_p \\ V\in\mathcal{D}-\mathcal{D}'}} \tau_{V,k}, \quad (4.11)$$

$$\lim_{n\to\infty} \kappa'_{n,k} = \sum_{\substack{V=V_1\cup V_2\cup\ldots\cup V_p \\ V\in\mathcal{D}-\mathcal{D}'}} \tau'_{V,k}. \quad (4.12)$$

Next, our goal is to find out $V \in \mathcal{D} - \mathcal{D}'$ such that $\tau_{V,k} \neq 0$.

First, let us consider an example. For $k = 2$, consider the pair partition from $\mathcal{D}-\mathcal{D}'$,

$$(P1)\ \ (i_1, i_3)\ \ (i_1 + j_1, i_3 + j_2)\ \ (i_2, i_4)\ \ (i_2 + j_1, i_4 + j_2) = (V_1, V_2, V_3, V_4),\ \ \text{(say)}.$$

Then for $(P1)$,

$$\begin{aligned}
|\tau'_{V,2}| &= \lim_{n\to\infty} \Big|\, \frac{1}{n^3} \sum_{j_1,j_2} \sum_{i_1,i_2,i_3,i_4} \Big(\prod_{l=1}^{2} K^2(\frac{j_l}{m_n})\delta_0(j_l)\Big) \prod_i Cum(X_t : t \in V_i) \,\Big| \\
&= \lim_{n\to\infty} \Big|\, \frac{1}{n^3} \sum_{j_1,j_2} \sum_{i_1,i_2,i_3,i_4} \Big(\prod_{l=1}^{2} K^2(\frac{j_l}{m_n})\delta_0(j_l)\Big) Q(i_3 - i_1)Q(i_3 - i_1 + j_2 - j_1) \\
&\qquad\qquad Q(i_4 - i_2)Q(i_4 - i_2 + j_2 - j_1) \,\Big| \\
&\le \lim_{n\to\infty} \frac{m_n}{n}\Big(m_n^{-1} \sum_{j_1} K^2(\frac{j_1}{m_n})\Big)\Big(\sum_{u_1,u_2,u} |\, Q(u_1)Q(u_1 + u)Q(u_2)Q(u_2 + u) \,|\,\Big) = 0.
\end{aligned}$$

Note that, $\prod_{i=1}^{4} Cum(X_t : t \in V_i)$ is independent of i_1, i_2 and j_1. An index x is called *free* in a partition $V = \cup_{i=1}^{p} V_i$, if $\prod_{i=1}^{p} Cum(X_t : t \in V_i)$ does not depend on x. In the partition $(P1)$, i_1, i_2 and j_1 are free indices. The number of free indices vary from partition to partition. Let $f(V)$ be the total number of free indices in the partition V. For example, $f(P1) = 3$. The above calculations for partition $(P1)$ clearly indicate that $|\tau_{V,k}|$ may be ∞, 0 or non-zero finite constant depending on the value of $f(V)$. Therefore, it is important to know $f(V)$ for different partitions in $V \in \mathcal{D} - \mathcal{D}'$. We will actually find an upper bound for $f(V)$. Later, we will see that this upper bound serves our purpose.

Consider the partition

$$(P2)\ \ (i_1, i_3, i_1 + j_1)\ \ (i_2, i_4, i_3 + j_2)\ \ (i_2 + j_1, i_4 + j_2) = (V_1, V_2, V_3),\ \ \text{(say)}.$$

Therefore, $\displaystyle\prod_{i=1}^{3} Cum(X_t : t \in V_i) = Q(u_1, j_1)Q(u_3, u_1 + u_2 + j_2)Q(u_3 + j_2 - j_1)$,

which is independent of i_1. Hence, $f(P2) = 1$. Note that in $(P1)$, it is enough to know $i_3 - i_1$ and $i_4 - i_2$ instead of individual i_1, i_2, i_3 and i_4. But in $(P2)$ as block sizes increase, we also need to know $i_1 - i_2$ together with $i_3 - i_1$ and $i_4 - i_2$. Therefore only one index can be free among i_1, i_2, i_3 and i_4. Again in $(P1)$, it is enough to know $j_2 - j_1$ instead of individual j_1 and j_2. But in $(P2)$, as the block $(i_1 + j_1, i_3 + j_2)$ splits up, we need to know j_1 and j_2 individually. Therefore if we consider any partition and split a block to increase other block sizes, then the indices which are free in the previous partition, may not remain free in the new partition and as a consequence the number of free indices gets reduced. In general, let $V = \cup_{i=1}^{p} V_i$ and $V' = \cup_{i=1} \cup V'_i$ be two partitions from $\mathcal{D}$ such that $p' < p$. V is called a *generator* of the partition V', if for each V'_i there exist $V_{k_1}, V_{k_2}, \ldots V_{k_{l_i}}, 1 \leq l_i < p, i = 1, 2, \ldots, p'$ such that

$$V'_i \subseteq V_{k_1} \cup V_{k_2} \cup \ldots \cup V_{k_{l_i}}, \ \ \forall i = 1, 2, \ldots, p'.$$

For example the partition $(P1)$ is a generator of the partion $(P2)$. In general,

$$f(V') \leq f(V). \tag{4.13}$$

Now, as pair partitions are generators of all the partitions in $\mathcal{D} - \mathcal{D}'$, to get an upper bound for $f(V)$, by (4.13), it is enough only to consider the indecomposable pair partitions from $\mathcal{D}$. Consider any pair partition V. Clearly, in this case, $p = 2k$ and cardinality of each V_i is 2. For any subset of blocks $\{V_{l_1}, V_{l_2}, \ldots, V_{l_s}\}$, $1 \leq s \leq 2k$, we say that it forms an *i-chain* of length s if each $(V_{l_N}, V_{l_{(N+1)(mod\ s)}}), N = 1, 2, \ldots, s$, has exactly one i-index in common. For example consider the following partition from $\mathcal{D}$ for $k = 4$:

$(P3)$ (i_1, i_3) $(i_1{+}j_1, i_3{+}j_2)$ (i_2, i_4) $(i_4{+}j_2, i_6{+}j_3)$ (i_6, i_8) $(i_8{+}j_4, i_5)$ $(i_5{+}j_3, i_2{+}j_1)$ $(i_7, i_7{+}j_4)$.

In this case, we have three i-chains: $\{(i_1, i_3), (i_1 + j_1, i_3 + j_2)\}$ of length 2, $\{(i_2, i_4), (i_4 + j_2, i_6 + j_3), (i_6, i_8), (i_8 + j_4, i_5), (i_5, i_2 + j_1)\}$ of length 5 and $\{(i_7, i_7 + j_4)\}$ of length 1.

For general k, each i-chain involves different i-indices. Moreover, only one i-index can be free from each i-chain and, other i indices involved in that chain can never be free for the entire partition. Similarly, one can define j-chains and the same facts hold for j-chains also. Let $f_i(V)$ and $f_j(V)$ be respectively the number of free i and j-indices in the partition V. Then clearly we have

$$f_i(V) = \text{total number of } i\text{−chains in } V, \qquad f_j(V) = \text{total number of } j\text{−chains in } V.$$

Hence, $f(V) = f_i(V) + f_j(V) =$ total number of chains in V. We will actually find an upper bound for $f_i(V)$ and then use it to get upper bounds for $f(V)$.

We call a chain *communicating* if it is an i-chain and each row of Table B communicates with other rows in this chain. In $(P1)$, there are two i-chains and both are communicating. In $(P3)$, the second i-chain with length 5 is communicating, whereas the first and third i-chains are not communicating. As we are concerned with indecomposable partitions, there must be at least one communicating chain in the partition. Also, for k-many rows in Table B, as the communicating chain involves at least k i-indices, the least possible length of a communicating chain is k. Therefore, $(k - 1)$ i-indices can never be free. As we have $2k$ many i-indices, $f_i(V) \leq k + 1$ for all pair partitions $V \in \mathcal{D}$. Also any j-index occurs twice with different i-indices. Therefore to make a j-index free, we have to lose at least one i-index from being free. Also, if we have r many j-indices free in some pair partition $V \in \mathcal{D}$, then at most $(k+1) - r$ many i-indices can be free in V. Hence, for all pair partitions $V \in \mathcal{D}$, $f(V) \leq k+1$. As pair partitions are generators of other partitions in $\mathcal{D} - \mathcal{D}'$, by (4.13)

$$f(V) \leq k+1 \quad \forall V \in \mathcal{D} - \mathcal{D}'. \tag{4.14}$$

Now, we will use this upper bound of $f(V)$ to prove (4.7). Suppose, for a $V \in \mathcal{D} - \mathcal{D}'$, we have r many j-indices free. Therefore by (4.14), there are at most $(k+1-r)$ many i-indices that can be free. Also, let J and I be the set of all j and i-indices respectively which are not free in V. By Assumption 1, for all $V \in \mathcal{D} - \mathcal{D}'$, we have,

$$\lim_{n\to\infty} \sum_{I,J} |\prod_{k=1}^{p} Cum(X_t : t \in V_k)| < \infty. \tag{4.15}$$

Also by (2.10), we have

$$\lim_{n\to\infty} \frac{1}{m_n^r} \sum_{J^c} \prod_{l:j_l \text{ free}} K^2(\frac{j_l}{m_n}) = \Big(\lim_{n\to\infty} \frac{1}{m_n} \sum_{j_l} K^2(\frac{j_l}{m_n})\Big)^r = \Big(\int_0^\infty K^2(x)dx\Big)^r. \tag{4.16}$$

Let us denote the number of zero j-indices by u. Therefore, by Assumption 2, as $K(\cdot)$ is bounded, we have, by (4.9)

$$\begin{aligned}
\big|\tau_{V,k}\big| \;&\leq\; \lim_{n\to\infty} \sum_{u=0}^{k} \frac{n^{k+1+k/2-u/2}}{n^{3k/2} m_n^{k/2-u/2}} \Big(\frac{m_n}{n}\Big)^r \Big\{\frac{1}{m_n^r} \sum_{J^c} \prod_{l:j_l \text{ free}} K^2(\frac{j_l}{m_n}) \delta_0(j_l)\Big\} \\
&\qquad \Big\{\sum_{I,J} \Big(\prod_{l:j_l \text{ not free}} K^2(\frac{j_l}{m_n}) \delta_0(j_l)\Big) \prod_{i=0}^{P} |Cum\,(X_t : t \in V_i)|\Big\} \\
&\leq\; \lim_{n\to\infty} \sum_{u=0}^{k} \frac{n^{k+1+k/2-u/2}}{n^{3k/2} m_n^{k/2-u/2}} \Big(\frac{m_n}{n}\Big)^r \Big\{\frac{1}{m_n^r} \sum_{J^c} \Big(\prod_{l:j_l \text{ free}} K^2(\frac{j_l}{m_n})\Big) \delta_0(j_l)\Big\} \\
&\qquad \Big\{\sum_{I,J} \prod_{i=0}^{p} |Cum\,(X_t : t \in V_i)|\Big\}.
\end{aligned}$$

Hence, by (4.15) and (4.16),

$$\tau_{V,k} = 0 \quad \textit{unless} \quad r = 1, \;\; k = 2, \;\; u = 0 \qquad \text{or} \qquad r = 0, \;\; k = 2 \;\; u = k \tag{4.17}$$

Therefore, (4.7)(b) is proved.

Now we will prove (4.7)(a). Let us first consider the case $r = 1$, $k = 2$, and $u = 0$. We need to find out partitions with one free j-index and two free i-indices. There are two such partitions as follows.

$$(P4) \qquad (i_1, i_3) \;\; (i_1 + j_1, i_3 + j_2) \;\; (i_2, i_4) \;\; (i_2 + j_1, i_4 + j_2),$$

$$(P5) \qquad (i_1, i_4)\ (i_1 + j_1, i_4 + j_2)\ (i_2, i_3)\ (i_2 + j_1, i_3 + j_2).$$

Note that contribution of these two partitions are same and the total contribution is

$$2t_2^2 Q^4(0) \lim_{n\to\infty} \frac{1}{m_n} \sum_{j=1}^{\infty} K^4(\frac{j}{m_n}) = 2t_2^2 Q^4(0) D(K),$$

where $D(K)$ is given by (1.7). Next, for $r = 0$, $k = 2$ and $u = k$, the contributing partitions are

$(P6)$	$(i_1, i_1)\ (i_3, i_3)\ (i_2, i_4)\ (i_2, i_4)$,	$(P7)$	$(i_1, i_1)\ (i_4, i_4)\ (i_2, i_3)\ (i_2, i_3)$,
$(P8)$	$(i_2, i_2)\ (i_3, i_3)\ (i_1, i_4)\ (i_1, i_4)$,	$(P9)$	$(i_2, i_2)\ (i_4, i_4)\ (i_1, i_3)\ (i_1, i_3)$,
$(P10)$	$(i_1, i_1)\ (i_3, i_3)\ (i_2, i_2, i_4, i_4)$,	$(P11)$	$(i_1, i_1)\ (i_4, i_4)\ (i_2, i_2, i_3, i_3)$,
$(P12)$	$(i_2, i_2)\ (i_3, i_3)\ (i_1, i_1, i_4, i_4)$,	$(P13)$	$(i_2, i_2)\ (i_4, i_4)\ (i_1, i_1, i_3, i_3)$,

and the contribution is $4t_1^2 Q^2(0) \sum_{u=-\infty}^{\infty} Cov(X_0^2, X_u^2)$. Hence,

$$\lim_{n\to\infty} \kappa_{n,2} = 4t_1^2 Q^2(0) \sum_{u=-\infty}^{\infty} Cov(X_0^2, X_u^2) + 2t_2^2 Q^4(0) D(K) = \sigma_0^2.$$

Therefore, (4.7)(a) holds. This completes the proof of (4.3) and hence of Theorem 2.1.

Using the same arguments used to prove Theorem 2.1, we have

$$\begin{aligned} \left|\tau'_{V,k}\right| &\leq \lim_{n\to\infty} \frac{n^{k+1}}{n^{3k/2}} \left(\frac{m_n}{n}\right)^r \Big\{\frac{1}{m_n^r} \sum_{J^c} \Big(\prod_{l:j_l\ free} K^2(\frac{j_l}{m_n}) \delta_0(j_l)\Big)\Big\} \qquad (4.18) \\ &\quad \Big\{ \sum_{I,J} \Big(\prod_{J} K^2(\frac{j_l}{m_n}) \delta_0(j_l)\Big) \prod_{i=0}^{p} |Cum\,(X_t : t \in V_i)| \Big\} \\ &\leq \lim_{n\to\infty} C \frac{n^{k+1}}{n^{3k/2}} \left(\frac{m_n}{n}\right)^r \frac{1}{m_n^r} \sum_{J^c} \Big(\prod_{l:j_l\ free} K^2(\frac{j_l}{m_n}) \delta_0(j_l)\Big) \sum_{I,J} \prod_{i=0}^{p} |Cum(X_t : t \in V_i)|\,. \end{aligned}$$

Hence, by (4.15) and (4.16),

$$\tau_{V,k} = 0 \quad unless\ k = 2\ \ and\ \ r = 0. \qquad (4.19)$$

Therefore, (4.7)(d) is proved.

To prove (4.7)(c), by (4.19), we need to find out the contribution of the partition $V \in \mathcal{D} - \mathcal{D}'$ for $k = 2$ and having three i-indices free. First consider pair partitions from $\mathcal{D} - \mathcal{D}'$. To make three i-indices free, we need to construct three i-chains. As the partitions are indicomposable, there is a communicating chain of length at least two. Therefore, other two i-chains have to be of length one each. These partitions are

$$(P14) \qquad (i_1, i_1 + j_1)\ (i_3, i_3 + j_2)\ (i_2, i_4)\ (i_2 + j_1, i_4 + j_2),$$

$(P15)$ $\quad (i_1, i_1 + j_1)\ (i_4, i_4 + j_2)\ (i_2, i_3)\ (i_2 + j_1, i_3 + j_2),$
$(P16)$ $\quad (i_2, i_2 + j_1)\ (i_3, i_3 + j_2)\ (i_1, i_4)\ (i_1 + j_1, i_4 + j_2),$
$(P17)$ $\quad (i_2, i_2 + j_1)\ (i_4, i_4 + j_2)\ (i_1, i_3)\ (i_1 + j_1, i_3 + j_2),$
$(P18)$ $\quad (i_1, i_1 + j_1)\ (i_3, i_3 + j_2)\ (i_2, i_4 + j_2)\ (i_4, i_2 + j_1),$
$(P19)$ $\quad (i_1, i_1 + j_1)\ (i_4, i_4 + j_2)\ (i_2, i_3 + j_2)\ (i_3, i_2 + j_1),$
$(P20)$ $\quad (i_2, i_2 + j_1)\ (i_3, i_3 + j_2)\ (i_1, i_1 + j_1)\ (i_4, i_4 + j_2),$
$(P21)$ $\quad (i_2, i_2 + j_1)\ (i_4, i_4 + j_2)\ (i_1, i_3 + j_2)\ (i_3, i_1 + j_1).$

Then by Assumption 1 and DCT, for $V \in \{(P14), (P15), (P16), (P17)\}$, the total contribution is

$$\sum_{u=-\infty}^{\infty} \Big(4t_1^2 Q^2(0) Q^2(u) + 4t_2^2 (Q * Q(u))^2 + 8t_1t_2 Q(0)Q(u)Q * Q(u)\Big). \quad (4.20)$$

Similarly, for all $V \in \{(P18), (P19), (P20), (P21)\}$, the total contribution is,

$$\sum_{u=-\infty}^{\infty} \Big(4t_1^2 Q^2(0) Q^2(u) + 4t_2^2 (Q * Q(u))(Q * Q(-u)) + 4t_1t_2 Q(0)Q(u)Q * Q(-u) + 4t_1t_2 Q(0)Q(u)Q * Q(u)\Big). \quad (4.21)$$

Next consider the non-pair partitions from $\mathcal{D} - \mathcal{D}'$ for $k = 2$ such that three i-indices are free. Note that those partitions have to be generated from $(P14)$ to $(P21)$. The only possible way to achieve this is to merge the last two blocks of $(P14)$ to $(P21)$ yielding the partitions:

$(P22)$ $\quad (i_1, i_1 + j_1)\ (i_3, i_3 + j_2)\ (i_2, i_2 + j_1, i_4, i_4 + j_2),$
$(P23)$ $\quad (i_1, i_1 + j_1)\ (i_4, i_4 + j_2)\ (i_2, i_2 + j_1, i_3, i_3 + j_2),$
$(P24)$ $\quad (i_2, i_2 + j_1)\ (i_3, i_3 + j_2)\ (i_1, i_1 + j_1, i_4, i_4 + j_2),$
$(P25)$ $\quad (i_2, i_2 + j_1)\ (i_4, i_4 + j_2)\ (i_1, i_1 + j_1, i_3, i_3 + j_2).$

The total contribution of the partitions $(P22)$ to $(P25)$ is

$$\sum_{u=-\infty}^{\infty} \Big(4t_1^2 Q^2(0) Q(0, u, u) + 4t_2^2 \sum_{j_1, j_2 = 1}^{\infty} Q(j_1)Q(j_2)Q(j_1, u, u + j_2)\Big) + 4t_1t_2 Q(0) \sum_{j=1}^{\infty} \sum_{u=-\infty}^{\infty} Q(j)(Q(0, u, u + j) + Q(j, u, u)). \quad (4.22)$$

Hence, by (4.20), (4.21) and (4.22), $\lim_{n\to\infty} \kappa'_{n,2}$ equals

$$4t_1^2 Q^2(0) \sum_{u=-\infty}^{\infty} \Big(2Q^2(u) + Q(0, u, u)\Big)$$

$$
\begin{aligned}
&+4t_2^2 \sum_{u=-\infty}^{\infty} \Big[(Q*Q(u))^2 + (Q*Q(u))(Q*Q(-u)) + \sum_{j_1=1}^{\infty}\sum_{j_2=1}^{\infty} Q(j_1)Q(j_2)Q(j_1,u,u+j_2)\Big] \\
&+4t_1t_2Q(0) \sum_{u=-\infty}^{\infty} \Big[3Q(u)Q*Q(u) + Q(u)Q*Q(-u) + \sum_{j=1}^{\infty} Q(j)\left(Q(0,u,u+j) + Q(j,u,u)\right)\Big] \\
= \;& 4t_1^2Q^2(0) \sum_{u=-\infty}^{\infty} \Big(2Q^2(u) + Q(0,u,u)\Big) \\
&+2t_2^2 \sum_{u=-\infty}^{\infty} \Big[Q*Q(u) + Q*Q(-u)\Big]^2 + 4t_2^2 \sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\sum_{u=-\infty}^{\infty} Q(i)Q(j)Q(i,u,u+j) \\
&+4t_1t_2Q(0) \sum_{u=-\infty}^{\infty} \Big[3Q(u)Q*Q(u) + Q(u)Q*Q(-u) + \sum_{j=1}^{\infty} Q(j)\left(Q(0,u,u+j) + Q(j,u,u)\right)\Big] \\
= \;& \sigma_1^2.
\end{aligned}
$$

Therefore, (4.7)(c) is proved and the proof of Theorem 2.2 is complete.

4.2 Verification of Assumption 1

We show how Assumption 1 in Section 2 may be verified for various non-linear time series models and more generally for α-mixing strictly stationary processes with exponentially decaying mixing coefficients. The following lemma is needed.

Lemma 4.2. (a) *[Ibragimov, 1962] Suppose $\{X_t\}$ is a strictly stationary α-mixing process. If ξ and η are measurable with respect to $\mathcal{F}_{-\infty}^{k}$ and $\mathcal{F}_{k+n}^{\infty}$ respectively, and if $|\xi| < C_1$, $|\eta| < C_2$, then*

$$
|\; E(\xi\eta) - E(\xi)E(\eta) \;| \leq \; 4C_1C_2\alpha(n).
$$

(b) *If $\{X_t\}$ is a uniformly bounded α-mixing stationary process with $\alpha(n) \leq \alpha^n$ for some $0 < \alpha < 1$, then, for $t_1 \leq t_2 \leq \ldots \leq t_k$, there exists $C > 0$ such that*

$$
|Q(t_1, t_2, \ldots, t_k)| \leq C\alpha^{t_k/k} \;\; \forall k \geq 1. \tag{4.23}
$$

Proof. We will show that

$$
|Q(t_1, t_2, \ldots, t_k)| \leq C\alpha^{t_j - t_{j-1}}, \quad \forall 1 \leq j \leq k, \;\; t_0 = 0. \tag{4.24}
$$

Then (4.23) follows by taking the geometric mean on both sides of (4.24).

We will use induction on k to prove (4.24). Let $\mathcal{P}(k)$ be the set of all possible partitions of $\{1, 2, \ldots, k\}$. We have to deal with $\mathcal{P}(k+1)$ to compute $Q(t_1, t_2, \ldots, t_k)$. Let $\mathcal{P}'(k+1)$ be the set of all possible partitions of $\{1, 2, \ldots, (k+1)\}$ such that there is no partition with one block as $\{k+1\}$. Therefore, $|Q(t_1, t_2, \ldots, t_k)|$ equals

$$
|\; Cum(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_k}) \;|=|\; Cum(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_k} - E(X_{t_k})) \;|
$$

$$= \;|\sum_{\sigma\in\mathcal{P}'(k+1)} (-1)^{r-1}(r-1)!E\Big((X_{t_k}-E(X_{t_k}))\prod_{\substack{t_j\in V^{\sigma t_k}\\ t_j\neq t_k}} X_{t_j}\Big)\prod_{\substack{i=1\\ V_i\neq V^{\sigma t_k}}}^{r} E\Big(\prod_{t_j\in V_i} X_{t_j}\Big)|$$

$$= \;|\sum_{\sigma\in\mathcal{P}'(k+1)} (-1)^{r-1}(r-1)!\Big(E\Big(\prod_{t_j\in V^{\sigma t_k}} X_{t_j}\Big)-E(X_{t_k})E\Big(\prod_{\substack{t_j\in V^{\sigma t_k}\\ t_j\neq t_k}} X_{t_j}\Big)\Big)\prod_{\substack{i=1\\ V_i\neq V^{\sigma t_k}}}^{r} E\Big(\prod_{t_j\in V_i} X_{t_j}\Big)|$$

$$= \;|\sum_{\sigma\in\mathcal{P}'(k+1)} \mathcal{S}_\sigma\;|, \text{ (say).}$$

where $\sigma = \cup_{i=1}^r V_i$. By Lemma 4.2 ($a$), for some $C_1, C_2 > 0$ and for all $\sigma \in \mathcal{P}'(k+1)$,

$$\mathcal{S}_\sigma \;\leq\; C_1 \mid Cov(\prod_{\substack{t_j\in V^{\sigma t_k}\\ t_j\neq t_k}} X_{t_j}, X_{t_k})| \leq C_2\alpha^{(t_k-t_{k-1})}. \tag{4.25}$$

Therefore, (4.24) holds for $j = k$. To prove other inequalities, we consider the disjoint decomposition of $\mathcal{P}'(k+1)$: $\mathcal{P}'(k+1) = \cup_{i=1}^k S_i^{(k)}$, where

$$\begin{aligned}
S_1^{(k)} &= \Big\{\sigma\in\mathcal{P}'(k+1) : (k+1)\in V^{\sigma k}\Big\}\\
S_2^{(k)} &= \Big\{\sigma\in(S_1^{(k)})^c : (k+1)\in V^{\sigma(k-1)}\Big\}\\
&\;\;\vdots\\
S_{k-1}^{(k)} &= \Big\{\sigma\in\Big(\cup_{i=1}^{k-2}S_i^{(k)}\Big)^c : (k+1)\in V^{\sigma 2}\Big\}\\
S_k^{(k)} &= \Big\{\sigma\in\Big(\cup_{i=1}^{k-1}S_i^{(k)}\Big)^c : (k+1)\in V^{\sigma 1}\Big\}
\end{aligned}$$

and for all $a = 1, 2, 3, \ldots k$, $V^{\sigma a}$ is the block of the partition σ with the element a. For example,

$$\begin{aligned}
S_1^{(3)} &= \{\{1,2,3,4\},\{1,2\}\{3,4\},\{1\}\{2,3,4\},\{1,3,4\}\{2\}\{1\}\{2\}\{3,4\}\},\\
S_2^{(3)} &= \{\{1,3\}\{2,4\},\{1,2,4\}\{3\},\{1\}\{3\}\{2,4\}\},\\
S_3^{(3)} &= \{\{1,4\}\{2,3\},\{1,4\}\{2\}\{3\}\}
\end{aligned}$$

and $\mathcal{P}'(4) = S_1^{(3)}\cup S_2^{(3)}\cup S_3^{(3)}$. Let us also define, for $1 \leq s \leq k$,

$$\begin{aligned}
T_{s1}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}) &= \sum_{\substack{\sigma\in S_s^{(k)}\\ \sigma=V_1\cup V_2\cup\ldots\cup V_r}} (-1)^{r-1}(r-1)!\prod_{i=1}^{r} E\Big(\prod_{t_j\in V_i} X_{t_j}\Big),\\
T_{s2}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}) &= -E(X_{t_k})T_{s1}^{(k)}(X_0, X_{t_1},\ldots,1),\\
T_{s}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}) &= T_{s1}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}) + T_{s2}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}).
\end{aligned}$$

Then it is easy to see that $|Q(t_1, t_2, \ldots, t_k)|$ equals

$$= \;|\sum_{i=1}^{k}\Big(T_{i1}^{(k)}(X_0, X_{t_1},\ldots,X_{t_k}) - E(X_{t_k})T_{i1}^{(k)}(X_0, X_{t_1}, X_{t_2},\ldots,X_{t_{k-1}},1)\Big)|$$

$$
\begin{aligned}
&= |\sum_{i=1}^{k}\Big(T_{i1}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_k}) + T_{i2}^{(k)}(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_{k-1}}, X_{t_k})\Big)| \\
&= |\sum_{i=1}^{k} T_i^{(k)}(X_0, X_{t_1}, \ldots, X_{t_k})|. \qquad (4.26)
\end{aligned}
$$

Now, note that, for some $C_1, C_2 > 0$,

$$
|T_k^{(k)}(X_0, X_{t_1}, \ldots, X_{t_k})| \leq C_1 |Cov(X_0, X_{t_k})| \leq C_2 \alpha^{t_k} \leq C_2 \alpha^{t_j - t_{j-1}}, \quad \forall 1 \leq j \leq k, \ t_0 = 0. \tag{4.27}
$$

Therefore, the proof will be complete if we show that, there exists C_{iju}, $u = 1, 2$, $1 \leq i \leq k-1$ and $1 \leq j \leq k-1$ such that

$$
|T_{iu}^{(k)}(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_k})| \leq C_{iju} \alpha^{(t_j - t_{j-1})}, \quad \forall 1 \leq j \leq k-1, \ t_0 = 0. \tag{4.28}
$$

To prove (4.28), we use induction on k. First, let us see this for the lower order cumulants. For $k = 2$,

$$
\begin{aligned}
T_{11}^{(2)}(X_0, X_{t_1}, X_{t_2}) &= E(X_0 X_{t_1} X_{t_2}) - E(X_0) E(X_{t_1} X_{t_2}) = Cov(X_0, X_{t_1} X_{t_2}), \\
T_{12}^{(2)}(X_0, X_{t_1}, X_{t_2}) &= -E(X_{t_2}) E(X_0 X_{t_1}) + E(X_0) E(X_{t_1}) E(X_{t_2}) = E(X_{t_2}) Cov(X_0, X_{t_1}).
\end{aligned}
$$

Hence, by Lemma 4.2 (a) and for some $C > 0$,

$$
|T_{11}^{(2)}(X_0, X_{t_1}, X_{t_2})| \leq C\alpha^{t_1}, \quad |T_{12}^{(2)}(X_0, X_{t_1}, X_{t_2})| \leq C\alpha^{t_1}.
$$

Therefore, (4.28) is proved for $k = 2$.

Let us assume (4.28) is true for k. Also, note that, for all $1 \leq j \leq k$,

$$
\begin{aligned}
T_{j1}^{(k+1)}(X_0, X_{t_1}, \ldots, X_{t_{k+1}}) &= \sum_{i=1}^{j-1} T_{i2}(X_0, X_{t_1}, \ldots, X_{t_{k-j}}, X_{t_{k-j+2}}, \ldots, X_{t_k}, X_{t_{k-j+1}} X_{t_{k+1}}) \\
&\quad + \sum_{i=j}^{k} T_i^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{k-j}}, X_{t_{k-j+2}}, \ldots, X_{t_k}, X_{t_{k-j+1}} X_{t_{k+1}}). \qquad (4.29)
\end{aligned}
$$

Therefore, to prove for $k+1$, it suffices to show that, for all $1 \leq i \leq k$, $1 \leq j \leq k$ and $1 \leq l \leq k+1$, there exists $C > 0$ such that

$$
\begin{aligned}
&|T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j} X_{t_{k+1}}) - E(X_{t_{k+1}}) T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j}| \\
&\leq C\alpha^{t_l - t_{l-1}}, \ \ t_0 = 0. \qquad (4.30)
\end{aligned}
$$

Now, note that as $X_{t_{k+1}}$ and X_{t_j} are in the same block, for some $C_1 \in \mathbb{R}$,

$$
\begin{aligned}
&T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j} X_{t_{k+1}}) - E(X_{t_{k+1}}) T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j}) \\
&= C_1 Cov(X_{t_j} \prod_{\substack{t_i \in V^{\sigma t_{k+1}} \\ t_i \neq t_j, t_{k+1}}} X_{t_i}, X_{t_{k+1}}).
\end{aligned}
$$

Also note that, if $t_i \in V^{\sigma t_{k+1}}$ and $t_i \neq t_j, t_{k+1}$, then $0 \leq i \leq j-1$. Therefore, by Lemma 4.2 (a), for all $j+1 \leq l \leq k+1$ and some $C_2 > 0$,

$$
\begin{aligned}
& |T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j}X_{t_{k+1}}) - E(X_{t_{k+1}})T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j})| \\
\leq \;\; & C_2\alpha^{t_{k+1}-t_j} \leq C_2\alpha^{t_l - t_{l-1}}.
\end{aligned}
$$

Hence, (4.30) is proved for $j+1 \leq l \leq k+1$.

Now, as $T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_j}, \ldots, X_{t_k})$, $T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j}X_{t_{k+1}})$ and $T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j})$ are same in their first j arguments and last the $k+1-j$ arguments are all $\mathcal{F}_{t_j}^{\infty}$ measurable, by Lemma 4.2 (a) and the statement (4.28) for k, for all $u = 1, 2$, $1 \leq i \leq k$, $1 \leq l \leq j$ and some $C > 0$,

$$
\begin{aligned}
|T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j}X_{t_{k+1}})| &\leq C\alpha^{t_l - t_{l-1}}, \\
|T_{iu}^{(k)}(X_0, X_{t_1}, \ldots, X_{t_{j-1}}, X_{t_{j+1}}, \ldots, X_{t_k}, X_{t_j})| &\leq C\alpha^{t_l - t_{l-1}}, \;\; t_0 = 0.
\end{aligned}
$$

Therefore, (4.30) is proved for $1 \leq l \leq j$. Therefore, by (4.26) and (4.30), (4.28) is proved for $k+1$ and the proof is complete. $\square$

4.3 GMC for linear processes: Assumption 1, Section 2

Consider the linear process $\{X_t\}$ defined in (1.14) where the $\{\varepsilon_t\}$ have finite eighth moment. Let $\{\varepsilon_t'\}$ be an i.i.d. copy of $\{\varepsilon_t\}$. Then from (1.10),

$$
X_t' = \sum_{k=0}^{t-1} \theta_k \varepsilon_{t-k} + \sum_{k=t}^{\infty} \theta_k \varepsilon_{t-k}'.
$$

Note that, $\{(\varepsilon_t - \varepsilon_t')\}_{t=-\infty}^{\infty}$ are i.i.d. and symmetric about zero. Hence, all their odd moments are zero. Therefore, for some $C_1 > 0$,

$$
\begin{aligned}
& E(X_t - X_t')^8 = E\Big(\sum_{k=t}^{\infty} \theta_k(\varepsilon_{t-k} - \varepsilon_{t-k}')\Big)^8 \\
= \;\; & \Big(\sum_{k=t}^{\infty} \theta_k^8\Big) E\Big(\varepsilon_t - \varepsilon_t'\Big)^8 + \Big(\sum_{k_1=t}^{\infty} \sum_{\substack{k_2=t \\ k_1 \neq k_2}}^{\infty} \theta_{k_1}^6 \theta_{k_2}^2\Big) E\,(\varepsilon_t - \varepsilon_t')^6\, E\,(\varepsilon_t - \varepsilon_t')^2 \\
& + \Big(\sum_{k_1=t}^{\infty} \sum_{\substack{k_2=t \\ k_1 \neq k_2}}^{\infty} \theta_{k_1}^4 \theta_{k_2}^4\Big) \Big(E\,(\varepsilon_t - \varepsilon_t')^4\Big)^2 + \Big(\sum_{k_1=t}^{\infty} \sum_{\substack{k_2=t \\ k_1 \neq k_2 \neq k_3}}^{\infty} \sum_{k_3=t}^{\infty} \theta_{k_1}^2 \theta_{k_2}^2 \theta_{k_3}^2\Big) \Big(E\,(\varepsilon_t - \varepsilon_t')^2\Big)^3 \\
\geq \;\; & C_1 \Big(\sum_{k=t}^{\infty} \theta_k^2\Big)^4.
\end{aligned}
$$

Similarly, for some $C_2 > 0$, $E\,(X_t - X_t')^8 \leq C_2 \Big(\sum_{k=t}^{\infty} \theta_k^2\Big)^4$. Therefore, (1.14) will have GMC(8) property if and only if $\{\theta_k\}$ are exponentially decaying.

4.4 Truncation arguments for Section 2

In this subsection, we will discuss how to reduce the moment condition for some specific models (see Remarks 2.6, 2.7, 2.10 and 2.11). The following lemma is needed.

Lemma 4.3. *[Brockwell and Davis, 2009] Let $\{X_n\}$ and $\{Y_{nj}\}$, $n \geq 1, j \geq 1$, be random variables such that*

(i) $Y_{nj} \xrightarrow{\mathcal{D}} Y_j$ *as* $n \to \infty$ *for each* $j = 1, 2, \ldots$.
(ii) $Y_j \xrightarrow{\mathcal{D}} Y$ *as* $j \to \infty$, *and*
(iii) $\lim_{j\to\infty} \lim_{n\to\infty} P\Big(|X_n - Y_{nj}| > \epsilon\Big) = 0$ *for every* $\epsilon > 0$.

Then $X_n \xrightarrow{\mathcal{D}} Y$ *as* $n \to \infty$.

Now, we will discuss truncation arguments for specific models.

(a) Truncation for power GARCH model (see Remark 2.6). We show that (2.13) holds for power GARCH model if Assumption 1 is replaced by $\mathrm{E}\, X_t^4 < \infty$ and *Condition* β. We do this in three steps.

Step 1*:* First assume that all moments of the IID process $\{\varepsilon_t\}$ exist. Define

$$\tilde{X}_t = \sqrt{h_t}\, I(|h_t| \leq c)\, \varepsilon_t \ \ \forall t. \tag{4.31}$$

We first show that Theorem 2.1 holds for $\{\tilde{X}_t\}$. Note that, all moments of the new process $\{\tilde{X}_t\}$ exist. As discussed before, [Carrasco and Chen, 2002] showed that $\{(X_t, \sqrt{h_t})\}$ is β mixing with exponentially decaying mixing coefficient under *Condition* β. Therefore, $\{\tilde{X}_t\}$ is also a β-mixing process with exponentially decaying mixing coefficient under the same assumptions as before. Moreover, $\{\tilde{X}_t\}$ is a white noise. Therefore, Theorem 2.1 holds for $\{\tilde{X}_t\}$.

Step 2*:* Now we show that (2.13) holds for $\{X_t\}$ if all moments of $\{\varepsilon_t\}$ exist. Note that it is enough to show

$$\begin{aligned}
&\lim_{c\to\infty}\lim_{n\to\infty} nE\Big(\frac{\hat{\gamma}^2(0) - E(\hat{\gamma}^2(0))}{\Big[4Q^2(0)\sum_{u=-\infty}^{\infty} Cov(X_0^2, X_u^2)\Big]^{1/2}} - \frac{\hat{\gamma}_c^2(0) - E(\hat{\gamma}_c^2(0))}{\Big[4Q_c^2(0)\sum_{u=-\infty}^{\infty} Cov(\tilde{X}_0^2, \tilde{X}_u^2)\Big]^{1/2}}\Big)^2 = 0,\\
&\lim_{c\to\infty}\lim_{n\to\infty} \frac{n^2}{m_n}E\Bigg[\frac{\sum_{j=0}^{\infty} K^2(\frac{j}{m_n})\Big(\hat{\gamma}^2(j) - E(\hat{\gamma}^2(j))\Big)}{\sqrt{2Q^4(0)D(K)}} - \frac{\sum_{j=0}^{\infty} K^2(\frac{j}{m_n})\Big(\hat{\gamma}_c^2(j) - E(\hat{\gamma}_c^2(j))\Big)}{\sqrt{2Q_c^4(0)D(K)}}\Bigg]^2 = 0,
\end{aligned} \tag{4.32}$$

where $\hat{\gamma}_c(\cdot)$ is the sample autocovariances of the process $\{\tilde{X}_t\}$ and $Q_c(0) = Var(\tilde{X}_t)$. Now, for the first term $T_{1,n,c}$, say, since all limits below exist, by (4.7)(a) and Step 1,

$\lim_{n\to\infty} T_{1,n,c}$ equals

$$= \lim_{n\to\infty}\Big(nE\Big(\frac{\hat\gamma^2(0)-Q^2(0)}{[4Q^2(0)\sum_{u=-\infty}^{\infty}Cov(X_0^2,X_u^2)]^{1/2}}\Big)^2 + nE\Big(\frac{\hat\gamma_c^2(0)-Q_c^2(0)}{[4Q_c^2(0)\sum_{u=-\infty}^{\infty}Cov(\tilde X_0^2,\tilde X_u^2)]^{1/2}}\Big)^2\Big)$$

$$-\lim_{n\to\infty} 2n\Big(\frac{Cov\left(\hat\gamma^2(0),\hat\gamma_c^2(0)\right)}{[16Q^2(0)Q_c^2(0)\sum_{u=-\infty}^{\infty}Cov(X_0^2,X_u^2)\sum_{u=-\infty}^{\infty}Cov(\tilde X_0^2,\tilde X_u^2)]^{1/2}}\Big)$$

$$= 1+1-\frac{2\lim_{n\to\infty} n^{-3}\sum_{t_1,t_2,t_3,t_4} Cov(X_{t_1}^2X_{t_2}^2,\tilde X_{t_3}^2\tilde X_{t_4}^2)}{[16Q^2(0)Q_c^2(0)\sum_{u=-\infty}^{\infty}Cov(X_0^2,X_u^2)\sum_{u=-\infty}^{\infty}Cov(\tilde X_0^2,\tilde X_u^2)]^{1/2}}$$

$$= 2-\frac{8Q(0)Q_c(0)\sum_{u=-\infty}^{\infty}Cov(X_0^2,\tilde X_u^2)}{[16Q^2(0)Q_c^2(0)\sum_{u=-\infty}^{\infty}Cov(X_0^2,X_u^2)\sum_{u=-\infty}^{\infty}Cov(\tilde X_0^2,\tilde X_u^2)]^{1/2}} = T_{1,c},\ \text{(say).}$$

By the existence of 4-th moment of the process and DCT, as $c\to\infty$, we have

$$Cov(X_0^2,\tilde X_u^2),\ Cov(\tilde X_0^2,\tilde X_u^2)\to Cov(X_0^2,X_u^2)\ \ \forall u.$$

Hence, $T_{1,c}=2-2=0$. Therefore, (4.32) is proved.

Similarly, for the second term $T_{2,n,c}$, say, $\lim_{n\to\infty} T_{2,n,c}$ equals

$$= 2-\frac{2\lim_{n\to\infty} n^{-2}m_n^{-1}\sum_{\substack{j_1,j_2\\t_1,t_2,t_3,t_4}} K^2(\frac{j_1}{m_n})K^2(\frac{j_2}{m_n})Cov\left(X_{t_1}X_{t_1+j_1}X_{t_2}X_{t_2+j_1},\tilde X_{t_3}\tilde X_{t_3+j_2}\tilde X_{t_4}\tilde X_{t_4+j_2}\right)}{2Q^2(0)Q_c^2(0)D(K)}$$

$$= 2-\frac{4}{2Q^2(0)Q_c^2(0)D(K)}\lim_{n\to\infty}\Big(\frac{1}{m_n}\sum_j K^4(\frac{j}{m_n})\Big)\Big(Cov(X_t,\tilde X_t)\Big)^4$$

$$= 2-\frac{2\lim_{n\to\infty}\Big(m_n^{-1}\sum_j K^4(\frac{j}{m_n})\Big)(E(\varepsilon_t^2))^4(E\,(h_tI(|h_t|\le c)))^4}{(E(\varepsilon_t^2))^4(E(h_t))^2(E(h_tI(|h_t|\le c)))^2D(K)}$$

$$= 2\Big(1-\frac{(E(h_tI(|h_t|\le c)))^2}{(E\,(h_t))^2}\Big)\to 0\ \text{ as } c\to\infty,$$

since $\mathrm{E}\,X_t^4<\infty$ implies $\mathrm{E}\,h_t^2<\infty$. Therefore, (4.32) is proved. Hence, (2.13) holds for $\{X_t\}$ if all moments of $\{\varepsilon_t\}$ are finite.

Step 3*:* Now we finally show that (2.13) holds if $\mathrm{E}\,X_t^4<\infty$ and *Condition* β holds. We define another IID process $\{\varepsilon_t'\}$ as $\varepsilon_t'=\varepsilon_tI\,(|\varepsilon_t|\le c)$ and $X_t'=\sqrt{h_t}\varepsilon_t'$ for all t . Note that

$\{X'_t\}$ is not a power GARCH model but using the arguments as given in Step 1 and Step 2, we can say that (2.13) holds for $\{X'_t\}$. Also, one can similarly show that (4.32) holds with $\hat{\gamma}_c(\cdot)$ and $Q_c(0)$ replaced by $\hat{\gamma}'_c(\cdot)$ and $Q'_c(0)$ respectively, where $\hat{\gamma}'_c(\cdot)$ is the sample autocovariances of the process $\{X'_t\}$ and $Q'_c(0) = Var(X'_t)$. Hence, the result follows.

(b) Truncation for linear process (see Remark 2.10). For some $c > 0$ and all $i = 0, \pm 1, \pm 2, \ldots$, we define

$$\begin{aligned} \varepsilon_i^* &= \varepsilon_i I\left(|\varepsilon_i| \le c\right) - E\left(\varepsilon_i I\left(|\varepsilon_i| \le c\right)\right), \\ \varepsilon_i^{**} &= \varepsilon_i I\left(|\varepsilon_i| > c\right) - E\left(\varepsilon_i I\left(|\varepsilon_i| > c\right)\right). \end{aligned}$$

Then, it is easy to see that $\varepsilon_i = \varepsilon_i^* + \varepsilon_i^{**}, \ \ \forall i = 0, \pm 1, \pm 2, \ldots$. As all the moments of ε_i^* is finite, Theorem 2.2 holds for

$$Y_{tc} = \sum_{k=0}^{\infty} \theta_k \varepsilon_{t-k}^*, \ \ t \ge 1, \ c > 0, \tag{4.33}$$

with limiting variance Σ_{1c}, obtained from Σ_1 by replacing each $E\left(\varepsilon_1^r\right)$ by $E\left(\varepsilon_1^{*r}\right)$ for all $r = 1, 2, 3, 4$. Also, as $E\left(\varepsilon_1^4\right) < \infty$, by DCT, it is easy to see that as $c \to \infty$,

$$E\left(\varepsilon_1^{*r}\right) \to E\left(\varepsilon_1^r\right), \ \ \forall r = 1, 2, 3, 4. \tag{4.34}$$

Hence, $\lim\limits_{c\to\infty} \Sigma_{1c} = \Sigma_1$ (componentwise). Therefore, (i) and (ii) of Lemma 4.3 are satisfied. To show (iii) of Lemma 4.3, we consider

$$\hat{\gamma}_c^*(i) = \frac{1}{n} \sum_{t=0}^{n-i} Y_{tc} Y_{(t+i)c}, \ \ i = 0, \pm 1, \pm 2, \ldots,$$

where $\{Y_{tc}\}$ is given by (4.33). Let,

$$g\left(t_1, t_2, k_j, u_j : \ j = 1, 2, 3, 4\right) = \varepsilon_{t_1-k_1}^{u_1} \varepsilon_{t_1-k_2+i}^{u_2} \varepsilon_{t_2-k_3}^{u_3} \varepsilon_{t_2-k_4+i}^{u_4} - E\left(\varepsilon_{t_1-k_1}^{u_1} \varepsilon_{t_1-k_2+i}^{u_2} \varepsilon_{t_2-k_3}^{u_3} \varepsilon_{t_2-k_4+i}^{u_4}\right),$$

where $u_j = *$ or $**$ and s is the number of j's such that $u_j = **$. Therefore, on simplification, we have

$$\begin{aligned} A_n &= \sqrt{n} \sum_{i=1}^{n-i} K^2\left(\frac{i}{m_n}\right) \left(\hat{\gamma}^2(i) - E\left(\hat{\gamma}^2(i)\right)\right) - \sqrt{n} \sum_{i=1}^{n-i} K^2\left(\frac{i}{m_n}\right) \left(\hat{\gamma}^{*2}(i) - E\left(\hat{\gamma}^{*2}(i)\right)\right) \\ &= n^{-3/2} \sum_{i=1}^{n-1} \sum_{t_1, t_2} \sum_{k_1, k_2, k_3, k_4} K^2\left(\frac{i}{m_n}\right) \theta_{k_1} \theta_{k_2} \theta_{k_3} \theta_{k_4} \sum_{\substack{u_1, u_2, u_3, u_4 \\ s \ge 1}} g\left(t_1, t_2, k_j, u_j : j = 1, 2, 3, 4\right). \end{aligned}$$

We split the above sum into several cases and deal with only three major cases. Similar calculations hold for the other cases.

Case 1: $t_1 - k_1 = t_1 - k_2 + i = t_2 - k_3 = t_2 - k_4 + i = v$, (say). Then, by (2.4), for some $C > 0$,

$$|A_n| \leq Cn^{-1/2}\Big(\sum_{i=1}^{\infty}|\theta_i|\Big)^3 \sum_{\substack{u_1,u_2,u_3,u_4\\ s\geq 1}} \frac{1}{n}\sum_{v} \mid \Big(\prod_{j,s\geq 1}\varepsilon_v^{u_j}\Big) - E\Big(\prod_{j,s\geq 1}\varepsilon_v^{u_j}\Big) \mid \xrightarrow{P} 0,$$

as $E \mid \Big(\prod_{j,s\geq 1}\varepsilon_v^{u_j}\Big) - E\Big(\prod_{j,s\geq 1}\varepsilon_v^{u_j}\Big) \mid < \infty$.

Case 2: $t_1 - k_1 = t_1 - k_2 + i = t_2 - k_3 = v \neq t_2 - k_4 + i$, (say). Then, by (2.4), for some $C > 0$,

$$|A_n| \leq Cn^{-1/2}\sum_{i,k_1,k_2,k_3,k_4}\theta_{k_1}\theta_{k_1+i}\theta_{k_3}\theta_{k_4}\sum_{\substack{u_1,u_2,u_3,u_4\\ s\geq 1}}\frac{1}{n}\sum_{v} \mid \varepsilon_v^{u_1}\varepsilon_v^{u_2}\varepsilon_v^{u_3}\varepsilon_{v+k_3-k_4+i}^{u_4} \mid \xrightarrow{P} 0,$$

as $E \mid \varepsilon_v^{u_1}\varepsilon_v^{u_2}\varepsilon_v^{u_3}\varepsilon_{v+k_3-k_4+i}^{u_4} \mid < \infty$.

Case 3: $t_1 - k_1 \neq t_1 - k_2 + i \neq t_2 - k_3 \neq t_2 - k_4 + i$. Then

$$\begin{aligned} E\left(A_n^2\right) &= \frac{m_n^2}{n}\Big(\frac{1}{m_n}\sum_{i=1}^{n}K^2(\frac{i}{m_n})\Big)\Big(\frac{1}{m_n}\sum_{j=1}^{n}K^2(\frac{j}{m_n})\Big)\sum_{k_i,1\leq i\leq 8}\prod\theta_{k_i} \\ &\quad \sum_{\substack{u_i,1\leq i\leq 8,\\ s\geq 1}}\frac{1}{n^2}\sum_{t_1,t_2,t_3,t_4}E\left(\varepsilon_{t_1-k_1}^{u_1}\varepsilon_{t_1-k_2+i}^{u_2}\varepsilon_{t_2-k_3}^{u_3}\varepsilon_{t_2-k_4+i}^{u_4}\varepsilon_{t_3-k_5}^{u_5}\varepsilon_{t_3-k_6+j}^{u_6}\varepsilon_{t_4-k_7}^{u_7}\varepsilon_{t_4-k_8+j}^{u_8}\right). \end{aligned}$$

Again, we can split the above sum into several cases. For example, consider $t_1 - k_1 = t_3 - k_5,\ t_1 - k_2 + i = t_3 - k_6 + j,\ t_2 - k_3 = t_4 - k_7,\ t_2 - k_4 + i = t_4 - k_8 + j$. Then,

$$\begin{aligned} E\left(A_n^2\right) &= \frac{m_n^2}{n}\Big(\frac{1}{m_n}\sum_{i=1}^{n}K^2(\frac{i}{m_n})\Big)\Big(\frac{1}{m_n}\sum_{j=1}^{n}K^2(\frac{j}{m_n})\Big)\sum_{k_1,k_2,k_3,k_4}\prod\theta_{k_i} \\ &\quad \sum_{\substack{u_i,1\leq i\leq 8,\\ s\geq 1}}\frac{1}{n^2}\sum_{t_1,t_2}E\left(\varepsilon_{t_1-k_1}^{u_1}\varepsilon_{t_1-k_2+i}^{u_2}\varepsilon_{t_2-k_3}^{u_3}\varepsilon_{t_2-k_4+i}^{u_4}\varepsilon_{t_1-k_1}^{u_5}\varepsilon_{t_1-k_2+i}^{u_6}\varepsilon_{t_2-k_3}^{u_7}\varepsilon_{t_2-k_4+i}^{u_8}\right). \end{aligned}$$

Now, as $s \geq 1$, by (4.34) and (2.4), we have

$$\lim_{c\to\infty}\lim_{n\to\infty}\sum_{\substack{u_i,1\leq i\leq 8\\ s\geq 1}}\frac{1}{n^2}\sum_{t_1,t_2}E\left(\varepsilon_{t_1-k_1}^{u_1}\varepsilon_{t_1-k_2+i}^{u_2}\varepsilon_{t_2-k_3}^{u_3}\varepsilon_{t_2-k_4+i}^{u_4}\varepsilon_{t_1-k_1}^{u_5}\varepsilon_{t_1-k_2+i}^{u_6}\varepsilon_{t_2-k_3}^{u_7}\varepsilon_{t_2-k_4+i}^{u_8}\right) = 0.$$

Same argument holds for other terms. Hence, as $m_n = O(\sqrt{n})$, we have $\lim_{c\to\infty}\lim_{n\to\infty}E\left(A_n^2\right) = 0$. Therefore, for all $\epsilon > 0$, $\lim_{c\to\infty}\lim_{n\to\infty}P\left(|A_n| > \epsilon\right) = 0$.

Also, consider

$$B_n = \sqrt{n}\left(\hat{\gamma}^2(0) - E\left(\hat{\gamma}^2(0)\right)\right) - \sqrt{n}\left(\hat{\gamma}^{*2}(0) - E\left(\hat{\gamma}^{*2}(0)\right)\right).$$

Then similar calculation as A_n implies that, for all $\epsilon > 0$, $\lim_{c\to\infty}\lim_{n\to\infty} P(|B_n| > \epsilon) = 0$. Hence, condition (*iii*) of Lemma 4.3 is satisfied and (2.13) holds for linear process defined in (1.14), under the finiteness of the 4-th moment of the process and if $m_n = O(\sqrt{n})$ and (2.4) holds.

(c) Truncation for non-linear moving average process (see Remark 2.10). As discussed in Remark 2.3, non-linear moving average processes (2.5) and (2.6)are α-mixing with finitely many non-zero mixing coefficient. Therefore, if all moments of $\{\varepsilon_t\}$ are finite, Theorems 2.1 and 2.2 hold respectively for the processes (2.5) and (2.6). Here, we will show truncation only for the process (2.5). Similar arguments hold for the process (2.6). We define $\{\varepsilon_t^*\}$ and $\{\varepsilon_t^{**}\}$ as in Section 4.4 (b). Then Theorem 2.2 holds for

$$Y_{tc} = \beta\varepsilon_{t-1}^*\varepsilon_{t-2}^* + \varepsilon_t^*, \ \ t \geq 1, \ c > 0. \tag{4.35}$$

Using the same arguments as in Section 4.4 (b), (i) and (ii) of Lemma 4.3 are satisfied for the process (2.5). To show (iii) of Lemma 4.3, we consider $\hat{\gamma}_c^*(i)$ as in Section 4.4 (b) for the process (4.35). Then on simplification we have

$$\frac{n}{\sqrt{m_n}}\sum_{j=1}^{n} K^2(\frac{j}{m_n})\Big(\hat{\gamma}^2(j)-E(\hat{\gamma}^2(j))\Big)-\frac{n}{\sqrt{m_n}}\sum_{j=1}^{n} K^2(\frac{j}{m_n})\Big(\hat{\gamma}_c^{*2}(j)-E(\hat{\gamma}_c^{*2}(j))\Big) = A_n - E(A_n),$$

where

$$\begin{aligned} A_n &= (\sqrt{m_n}n)^{-1}\sum_{j=1}^{n} K^2(\frac{j}{m_n})\sum_{t_1,t_2}\Big(\sum_{\substack{u_1,u_2,\ldots,u_8\\ s\geq 1}} \beta^4\varepsilon_{t_1-2}^{u_1}\varepsilon_{t_1-1}^{u_2}\varepsilon_{t_1-2+j}^{u_3}\varepsilon_{t_1-1+j}^{u_4}\varepsilon_{t_2-2}^{u_5}\varepsilon_{t_2-1}^{u_6}\varepsilon_{t_2-2+j}^{u_7}\varepsilon_{t_2-1+j}^{u_8} \\ &\quad + \sum_{\substack{u_1,u_2,\ldots,u_4\\ s\geq 1}} \varepsilon_{t_1}^{u_1}\varepsilon_{t_1+j}^{u_2}\varepsilon_{t_2}^{u_3}\varepsilon_{t_2+j}^{u_4} + \text{other similar terms}\Big) \end{aligned}$$

where $u_i = *$ or $**$ for all i, and s is the number of i's such that $u_i = **$. Now, using same arguments as in Section 4.4 (b), (iii) of Lemma 4.3 holds for (2.5). Hence, $E(X_t^4) < \infty$ and $m_n = O(\sqrt{n})$ is enough for Theorem 2.1 to hold for this process (2.5).

References

T. Anderson. Stochastic autoregressive volatility: a framework for volatility clustering. *Journal of Mathematical Finance.*, 4(2):75–102, 1994.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

G. E. P. Box and D. A. Pierce. Distribution of residual autocorrelations in autoregressive integrated moving average time series moddels. *J. Amer. Statist. Assoc.*, 65(1):1509–1526, 1970.

F. J. Breidt, R. A. Davis, and A. A. Trindade. Least absolute deviation estimation for all-pass time-series models. *Ann. of Statist.*, 29:919–946, 2001.

D. R. Brillinger. *Time series: data analysis and theory*, volume 36. Siam, 2001.

P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer, 2009.

M. Carrasco and X. Chen. Mixing and moment properties of various GARCH and stocahstic volatility models. *Econometric Theory*, 18:17–39, 2002.

R. Engle. Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.

R. Engle and V. Ng. Measuring and testing the impact of news on volatility. *Journal of Finance.*, 48:1749–1778, 1993.

R. F. Engle and J. R. Russell. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pages 1127–1162, 1998.

J. Fan and Q. Yao. *Nonlinear time Series: nonparametric and parametric methods.* Springer, New York, 2003.

C. Francq, R. Roy, and J. M. Zakoïan. Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100:532–544, 2005.

J. Geweke. Modelling the persistence of conditional variances: a comment. *Econometric Reviews.*, 5:57–61, 1986.

L. Glosten, R. Jagannathan, and D. Runke. Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance.*, 48: 1779–1801, 1993.

C. W. Granger and T. Terasvirta. *Modelling non-linear economic relationships*. Oxford University Press, 1993.

C. W. J. Granger and A. P. Andersen. *An introduction to bilinear time series models*. Vandenhoeck und Ruprecht Göttingen, 1978.

F. J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proc. IEEE*, 66:51–83, 1978.

Y. Hong. Consistent testing for serial correlation of unknown form. *Econometrica*, 64: 837–864, 1996.

Y. Hong and Y. J. Lee. Consistent testing for serial uncorrelation of unknown form under general conditional heteroscedasticity. *Preprint, Cornell University, Department of Economics*, 2003.

T. Hsing and W. B. Wu. On weighted U-statistics for stationary processes. *Ann. Probab.*, 32:1600–1631, 2004.

T. Hsing and W. B. Wu. On linear processes with dependent innovations. *Stoch. Proc. Appl.*, 115:939–958, 2005.

I. A. Ibragimov. Mixing and moment properties of various GARCH and stocahstic volatility models. *Theory Probab. Appl.*, VII(4):349–382, 1962.

G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65:297–303, 1978.

A. Mc.Leod. On the distribution of residual autocorrelations in Box-Jenkins models. *J. Roy. Statist. Soc. B*, 40:296–302, 1978.

D. Nelson. Conditional heteroskedasticity on asset returns: A new approach. *Econometrica.*, 59:347–370, 1991.

D. F. Nicholls and B. G. Quinn. *Random coefficient autoregressive models: An Introduction.* Springer Verlag, New York, 1982.

D. T. Pham. Bilinear Markovian representation and bilinear model. *Stoch. Proc. Appl.*, 20:295–306, 1985.

D. N. Politis. Adaptive bandwidth choice. *J. Nonparam. Statist.*, 15(4-5):517–533, 2003.

M. B. Priestley. *Spectral analysis and Time series. Vol I.* Academic, New York, 1981.

J. L. Romano and L. A. Thombs. Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.*, 91:590–600, 1996.

X. Shao. Testing for White Noise under unknown dependence and its application to diagnostic checking for time series models. *Econometric Theory*, pages 1–32, 2011.

X. Shao and W. B. Wu. Asymptotic spectral theory for non-linear time series. *Ann. Statist.*, 4:1773–1801, 2007.

T. Subba Rao and M. M. Gabr. *An introduction to bispectral analysis and bilinear time series*. Springer Verlag, New York, 1984.

G. Terdik. *Bilinear stochastic models and related problems of nonlinear time series analysis: A frequency domain approach.* Springer Verlag, New York, 1999.

H. Tong. *Threshold models in non-linear time series analysis.* Springer Verlag, New York, 1983.

W. B. Wu and X. Shao. Limit theorems for iterated random functions. *J. Appl. Probab.*, 41:425–436, 2004.