Journal Afrika Statistika

ISSN 2316-090X

Journal Afrika Statistika
Vol. 5, N°2, 2010, page 197–209

An efficient locally asymptotic parametric test in nonlinear heteroscedastic time series models

Fateh Chebana[†] **and Naâmane Laïb**[‡]

[†]INRS-ETE, 490, de la Couronne, G1K 9A9 Qubec, QC Canada
[‡]LSTA, Université Paris VI. 175, rue du Chevaleret, 75013 PARIS, FRANCE

Received 29 May 2010; Accepted 27 September 2010

Abstract. In this paper we deal with a locally asymptotic stringent test for a general class of nonlinear time series heteroscedastic models. Based on the local asymptotic normality (LAN) property of these models, we propose a score-type test statistic for testing hypotheses on the parameters appearing in the mean and variance functions of the proposed statistical test with and without nuisance parameters. Its asymptotic null distribution is obtained as well as the local power of the test.

Résumé. Dans cet article, nous étudions les propriétés asymptotiques d'un test de score traitant simultanément des hypothèses portant sur des fonctions moyennes et variances conditionnelles dans une classe assez générale de modèles hétéroscédastiques non linéaires de séries chronologiques. La suite des alternatives locales considérée est paramétrique portant sur les paramètres intervenant dans les fonctions moyennes et variances du modèle. Nous établissons d'abord la normalité locale asymptotique (LAN) du modèle. En se basant sur ce résultat la loi limite de la statistique du test proposée a été obtenue sous l'hypothèse nulle et aussi sous des alternatives locales en présence ou non des paramètres de nuisance.

Key words: ARCH processes; Ergodic processes; LAN; Local power; Nonlinear processes; Score test; Time series.
AMS 2000 Mathematics Subject Classification : Primary 62G05, 62M20; Secondary 60J15.

1. Introduction

The present paper is concerned with the construction of asymptotically efficient test in a class of higher-order nonlinear time series models of the form

$$X_i = m(\mathbf{X}_{i-1}, \theta) + \sigma(\mathbf{X}_{i-1}, \rho)\epsilon_i, \quad i \geq d, \tag{1}$$

where $\mathbf{X}_{i-1} = (X_{i-1}, \ldots, X_{i-d})^\top$, $d \geq 1$, $m(\cdot, \theta)$ and $\sigma(\cdot, \rho)$ are given functions, $\nu = (\theta, \rho)^\top \in \Theta$ is a vector of parameters, Θ_1 and Θ_2 are open subsets of $\mathbb{R}^q$ ($q \geq 1$), $\Theta = \Theta_1 \times \Theta_2$, q and d are positive integers. The model specified through (1) is assumed to be identifiable, stationary and ergodic with finite second moment. The ϵ_i's are independent identically distributed (iid) with zero mean and variance one and a common known continuous positive Lebesgue density function f which admits third order derivative, and for any $i \geq d$, ϵ_i is independent of $\mathbf{X}_{i-1}$.

Model (1) is a d-order autoregressive process with ARCH errors (AR(d)-ARCH(d)). It has been considered in several research areas such as econometrics and control theory with specific assumptions on the innovations' distribution. Several papers have been devoted to the problem of testing simple and/or composite hypotheses on the parametric form of the conditional mean *or* the conditional variance functions. For more details and a literature review we refer for instance to Laïb [5] and Chebana and Laïb [1]. Note that most of these tests are usually derived for testing first-order autoregressive models. Furthermore, the study of the local power has attracted less attention. Hwang and Basawa [4]

Fateh Chebana: fateh_chebana@ete.inrs.ca
Naâmane Laïb: naamane.laib@upmc.fr

have constructed asymptotically efficient tests for testing the null hypothesis that the true model is a first-order linear autoregressive process against a sequence of local alternatives. Here we consider a more general class of processes defined by (1). Our goal is to propose a simultaneous testing procedure for testing the conditional mean and the conditional variance in parametric time series models. To be more precise, we consider testing the null hypothesis

$$H_0 : m(\cdot, \theta) = m(\cdot, \theta_0) \quad \text{and} \quad \sigma(\cdot, \rho) = \sigma(\cdot, \rho_0), \tag{2}$$

against the sequence of local alternatives

$$H_1^n : m(\cdot, \theta) = m(\cdot, \theta_0 + n^{-1/2}h_1) \quad \text{and} \quad \sigma(\cdot, \rho) = \sigma(\cdot, \rho_0 + n^{-1/2}h_2), \tag{3}$$

where θ_0 and ρ_0 are the true finite-dimensional parameters and h_1 and h_2 are given constant vectors of $\mathbb{R}^q$. For given functions $m(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$, the above hypotheses may be written as: $H_0 : (\theta, \rho) = (\theta_0, \rho_0)$ against $H_1^n : (\theta, \rho) = (\theta_0 + n^{-1/2}h_1,\ \rho_0 + n^{-1/2}h_2)$. In addition, we consider also the case with nuisance parameters.

The paper is organized as follows. In Section 2, we establish the LAN property for Model (1) via the quadratic mean differentiability. A score-type quadratic test based on this result is then obtained. Both the null and non-null limiting distributions of this statistical test are also derived. We also treated the case where nuisance parameters are present in the conditional mean and variance functions. A particular attention is given in Section 3 to a special class of models for which our results can be applied. The last section is devoted to proofs.

2. Assumptions and main results

2.1. Notations and technical assumptions

The statement of our results requires to introduce some notations and to impose some assumptions. For a vector $\mathbf{s} = (s_1, \ldots, s_q)$, $q \geq 1$, set $\|\mathbf{s}\| = \max_{1 \leq i \leq q} |s_i|$ and $\|\mathbf{s}\|_q$ the Euclidian norm.

Let U be an open set of $\mathbb{R}^q$ and $\psi : \mathbb{R}^q \times U \to \mathbb{R}$ which is assumed to be of class C^1 on U. For any $x \in \mathbb{R}^d$, we denote by

$$\nabla\psi(x, \mathbf{s}) = \left(\frac{\partial\psi}{\partial s_1}(x, \mathbf{s}), \ldots, \frac{\partial\psi}{\partial s_q}(x, \mathbf{s})\right)^\top \quad q \times 1 \quad \text{vector}$$

where

$$\frac{\partial\psi}{\partial s_k}(x, \mathbf{s}), \quad k = 1, \ldots, q,$$

are the partial derivative of $\psi(x, \mathbf{s})$ with respect to s_k.

To get simpler presentation, let us also put $\ell(x) := \log f(x)$ and denote by $\ell'(\cdot)$ its derivative. For $k = 0, 1, 2$, denote by I_k the quantity

$$I_k := E\left[\ell'(\epsilon_1)^2 \epsilon_1^k\right]. \tag{4}$$

Note that when $k = 0$, I_k represents the Fisher information. The notation $\stackrel{\mathcal{D}}{\to}$ stands for the convergence in distribution of random variables. The following assumptions are required to state the first results.

C1) There exist closed balls $\overline{B}_0 = \overline{B}(\theta_0, r_0)$ and $\overline{B}_1 = \overline{B}(\rho_0, r_1)$, with centers θ_0 and ρ_0 and radiuses r_0 and r_1, included in $\text{int}(\Theta_1)$ and $\text{int}(\Theta_2)$ respectively, and positive functions M_0 and M_1 such that

- 1

$$E\left(M_j^{2+\gamma}(X_{d-1})\right) < \infty \quad j = 0, 1 \text{ for a positive constant } \gamma$$

-2 For any $x \in \mathbb{R}^d$

$$\sup_{\theta \in \overline{B}_0} \left\| \frac{\nabla m(x, \theta)}{\sigma(x, \rho)} \right\|_q \leq M_0(x) \ \text{ and } \ \sup_{\rho \in \overline{B}_1} \left\| \frac{\nabla \sigma(x, \rho)}{\sigma(x, \rho)} \right\|_q \leq M_1(x).$$

C2) For all fixed x, the function $\theta \mapsto m(x, \theta)$ (resp. $\rho \mapsto \sigma(x, \rho)$) has continuous derivatives up to order 3 and for all θ (resp. ρ), the functions $x \mapsto m(x, \theta)$ and $\nabla m(x, \theta)$ (resp. $x \mapsto \sigma(x, \rho)$ and $\nabla\sigma(x, \rho)$) are continuous.

2.2. LAN of Model (1)

Our first aim in this subsection is to construct a score-type test to examine the hypothesis H_0 against the alternative H_1^n. We start by establishing a LAN property of Model (1) using the approach of quadratic mean differentiability (see definition below). To this end, let $g_\nu(X_i|\mathbf{X}_{i-1})$ be the conditional density function of X_i given $\mathbf{X}_{i-1}$ and $L_n(\nu)$ its conditional likelihood, which is given by

$$L_n(\nu) = \prod_{i=1}^{n} g_\nu(X_i|\mathbf{X}_{i-1}) = \prod_{i=1}^{n} \frac{1}{\sigma(\mathbf{X}_{i-1},\rho)} f\left(\frac{X_i - m(\mathbf{X}_{i-1},\theta)}{\sigma(\mathbf{X}_{i-1},\rho)}\right).$$

Therefore, the conditional log-likelihood ratio (for H_0 against H_1^n) is

$$\Lambda_n := \log\left[\frac{L_n(\nu_n)}{L_n(\nu_0)}\right] = 2\sum_{i=1}^{n} \log \phi_i(\nu_n,\nu_0), \quad \text{where} \quad \phi_i(\nu^*,\nu) := \frac{\sqrt{g_{\nu^*}(X_i|\mathbf{X}_{i-1})}}{\sqrt{g_\nu(X_i|\mathbf{X}_{i-1})}}$$

and $\nu_n = (\theta_0 + h_1/\sqrt{n},\ \rho_0 + h_2/\sqrt{n})$. Let $\dot\phi_i(\nu)$ be the quadratic mean derivative of $\phi_i(\nu^*,\nu)_{|\nu^*=\nu}$ given by the following $2q \times 1$ vector

$$\begin{aligned}\dot\phi_i(\nu) &= \frac{\nabla g_\nu(X_i|\mathbf{X}_{i-1})}{2g_\nu(X_i|\mathbf{X}_{i-1})} \qquad (5)\\ &= -\frac{1}{2}\left[\frac{\nabla m(\mathbf{X}_{i-1},\theta)}{\sigma(\mathbf{X}_{i-1},\rho)}\ell'(\epsilon_i); \frac{\nabla\sigma(\mathbf{X}_{i-1},\rho)}{\sigma(\mathbf{X}_{i-1},\rho)}\left(1+\epsilon_i\ell'(\epsilon_i)\right)\right]^\top.\end{aligned}$$

Recall that a random function $\zeta(\nu)$ is differentiable in quadratic mean at ν if $\frac{1}{t}\{\zeta(\nu + t\mathbf{h}) - \zeta(\nu)\} \xrightarrow{L_2} \mathbf{h}^\top\dot\zeta(\nu)$ as $t \to 0$, uniformly in bounded $\mathbf{h}$, where $\xrightarrow{L_2}$ means the convergence in quadratic mean.

Proposition 1 below states that, under H_0, the function $\dot\phi_i(\nu)$ is the derivative in quadratic mean of $\phi_i(\nu^*,\nu)$ with respect to ν^* at $\nu^* = \nu$.

Proposition 1. *Assuming conditions (C1)-(C2) hold, then we have under* H_0

$$\frac{1}{t}\{\phi_i(\nu_0 + t\mathbf{h};\nu_0) - 1\} \xrightarrow{L_2} \mathbf{h}^\top\dot\phi_i(\nu_0) \quad \text{as } t \to 0, \quad \text{uniformly in bounded } \mathbf{h}.$$

The LAN property of Model (1), stated in Theorem 1 below, is a consequence of Proposition 1 (see Roussas [8, pp. 53-54]).

Theorem 1. *Assuming satisfied the conditions of Proposition 1, then we have under* H_0

$$\begin{aligned}\Lambda_n &= \mathbf{h}^\top S_n(\nu_0) - \frac{1}{2}\mathbf{h}^\top\Gamma(\nu_0)\mathbf{h} + o_P(1), \text{ and}\\ S_n(\nu_0) &\xrightarrow{\mathcal{D}} \mathcal{N}(0,\Gamma(\nu_0)). \qquad (6)\end{aligned}$$

The score function $S_n(\nu)$ *and its covariance matrix* $\Gamma(\nu)$ *are given by*

$$\begin{aligned}S_n(\nu) &= \frac{2}{\sqrt{n}}\sum_{i=d}^{n}\dot\phi_i(\nu)\\ \Gamma(\nu) &= 4E\left[\dot\phi_d(\nu)\dot\phi_1^\top(\nu)\right].\end{aligned}$$

We deduce from Theorem 1 that, under H_0

$$(S_n(\nu_0),\Lambda_n) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix}0\\ -\frac{1}{2}\mathbf{h}^\top\Gamma(\nu_0)\mathbf{h}\end{pmatrix}, \begin{pmatrix}\Gamma(\nu_0) & \Gamma(\nu_0)\mathbf{h}\\ \mathbf{h}^\top\Gamma(\nu_0) & \mathbf{h}^\top\Gamma(\nu_0)\mathbf{h}\end{pmatrix}\right).$$

Consequently, since the hypotheses H_0 and H_1^n are contiguous, Le Cam's third lemma leads to

Corollary 1. *Under assumptions of Theorem 1, we have*

$$S_n(\nu_0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\Gamma(\nu_0)\mathbf{h}, \Gamma(\nu_0)\right) \quad \textit{under } H_1^n. \tag{7}$$

The quantities given in Theorem 1 can be given explicitly as follows:

$$S_n(\nu) = -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\nabla m(\mathbf{X}_{i-1},\theta)}{\sigma(\mathbf{X}_{i-1},\rho)}\ell'(\epsilon_i); \frac{\nabla \sigma(\mathbf{X}_{i-1},\rho)}{\sigma(\mathbf{X}_{i-1},\rho)}\left(1+\epsilon_i\ell'(\epsilon_i)\right)\right]^\top \tag{8}$$

and the $2q \times 2q$ matrix

$$\Gamma(\nu) = \begin{pmatrix} I_0 E\frac{\nabla m(\mathbf{X}_d,\theta)}{\sigma(\mathbf{X}_d,\rho)}\left(\frac{\nabla m(\mathbf{X}_d,\theta)}{\sigma(\mathbf{X}_d,\rho)}\right)^\top & I_1 E\frac{\nabla m(\mathbf{X}_d,\theta)}{\sigma(\mathbf{X}_d,\rho)}\left(\frac{\nabla \sigma(\mathbf{X}_d,\rho)}{\sigma(\mathbf{X}_d,\rho)}\right)^\top \\ I_1 E\frac{\nabla \sigma(\mathbf{X}_d,\rho)}{\sigma(\mathbf{X}_d,\rho)}\left(\frac{\nabla m(\mathbf{X}_d,\theta)}{\sigma(\mathbf{X}_d,\rho)}\right)^\top & (I_2-1) E\frac{\nabla \sigma(\mathbf{X}_d,\rho)}{\sigma(\mathbf{X}_d,\rho)}\left(\frac{\nabla \sigma(\mathbf{X}_d,\rho)}{\sigma(\mathbf{X}_d,\rho)}\right)^\top \end{pmatrix}$$

where the I_k's and $\dot{\phi}_1(\nu)$ are defined in (4) and (5) respectively.

2.3. Locally asymptotic statistical test in presence of nuisance parameters

In this subsection we construct a locally asymptotically most stringent test (in the Le Cam sense theory) in presence of nuisance parameters for testing simultaneously the linearity and the heteroscedasticity in Model (1). The notion of most stringency is a concept of optimality (see e.g. Wald [11]). A test ϕ^* is most stringent in the class $\mathcal{C}_\alpha := \{\phi \mid E_\nu(\phi) \le \alpha,\ \forall \nu \in H_0\}$, if $\phi^* \in \mathcal{C}_\alpha$ and its maximum regret

$$r(\phi) := \sup_{\nu \in H_1}\left(\sup_{\phi' \in \mathcal{C}_\alpha} E_\nu(\phi') - E_\nu(\phi)\right) \tag{9}$$

achieves a minimum over $\mathcal{C}_\alpha$, i.e. $r(\phi^*) \le r(\phi),\ \forall \phi \in \mathcal{C}_\alpha$.

To treat now the problem of the presence of some nuisance parameters in the model, let us start with a simple case. Suppose for instance that the space of parameters Θ is an open subset of $\mathbb{R}^4$ and the parameter $\nu = (\theta_1, \theta_2, \rho_1, \rho_2)$ is partitioned into $\nu_1 = (\theta_1, \rho_1)$ and $\nu_2 = (\theta_2, \rho_2)$. Assume also that ν_1 is the parameter of interest and ν_2 is a nuisance parameter. The null hypothesis K_0 and the local alternative K_1^n can then be formulated as follows

$$K_0 : (\theta_1,\ \rho_1) = (\theta_{01},\ \rho_{01}) \quad \text{and} \quad K_1^n : (\theta_1,\ \rho_1) = \left(\theta_{01} + h_{11}/\sqrt{n},\ \rho_{01} + h_{21}/\sqrt{n}\right) \tag{10}$$

with $\mathbf{h} = (h_{11}, h_{12}, h_{21}, h_{22}) \in \mathbb{R}^4$. Define the 4×2-dimensional matrix $\Omega = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^\top$. One can observe that the hypothesis K_0 is equivalent to $\nu - \nu_0 \in \mathcal{M}(\Omega)$, where $\mathcal{M}(\Omega)$ stands for the linear subspace of $\mathbb{R}^4$ spanned by the columns of Ω.

To deal with a general framework, we assume that Θ is an open subset of $\mathbb{R}^K$ with $K = 2q$, Ω is a $K \times r$-dimensional matrix with rank r ($r < K$), $\mathcal{M}(\Omega)$ is a linear subspace of $\mathbb{R}^K$ spanned by the columns of Ω. Let $\mathcal{M}(\Omega)_\perp$ be the linear subspace of $\mathbb{R}^K$ orthogonal to $\mathcal{M}(\Omega)$ and $P_{\mathcal{M}(\Omega)}$ be the orthogonal projection on $\mathcal{M}(\Omega)$ which is characterized by

$$P_{\mathcal{M}(\Omega)} = \Omega(\Omega^\top\Omega)^{-1}\Omega^\top = Id - P_{\mathcal{M}(\Omega)_\perp},$$

where Id is the identity matrix. The vector ν of parameters is partitioned into ν_1 and ν_2 vectors. Here ν_1 is considered as r-column matrix and ν_2 as $(K-r)$-column matrix. In this case, the null hypothesis K_0 and the local alternative K_1^n can be expressed as

$$K_0 : n^{-1/2}\mathbf{h} \in \mathcal{M}(\Omega) \quad \text{against} \quad K_1^n : n^{-1/2}\mathbf{h} \notin \mathcal{M}(\Omega).$$

They can be also written as

$$K_0 : \Gamma^{\frac{1}{2}}(\nu_0)\mathbf{h} \in \mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\sqrt{n}\ \Omega) \quad \text{against} \quad K_1^n : \Gamma^{\frac{1}{2}}(\nu_0)\mathbf{h} \notin \mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\sqrt{n}\ \Omega)$$

or equivalently

$$K_0 : \left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{\frac{1}{2}}(\nu_0)\mathbf{h} = 0 \quad \text{against} \quad K_1^n : \left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{\frac{1}{2}}(\nu_0)\mathbf{h} \neq 0. \tag{11}$$

Notice that the case of absence of nuisance parameters corresponds to $\Omega \equiv 0$ (the null matrix). The hypotheses treated by Hwang and Basawa [4] can be obtained as a particular case by choosing $\Omega = (0,1)^\top$.

According to the bilateral form of the alternative hypothesis in (11), it is convenient to use a statistical test based on a quadratic form of the score statistic $S_n(\nu_0)$ such as

$$\begin{aligned}\zeta^S_{n,\Omega}(\nu_0) &= S_n(\nu_0)^\top \left[\Gamma^{-1}(\nu_0) - \Omega(\Omega^\top\Gamma(\nu_0)\Omega)^{-1}\Omega^\top\right] S_n(\nu_0) \\ &= \left\| \left[Id - P_{\mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\Omega)}\right] \Gamma^{-\frac{1}{2}}(\nu_0) S_n(\nu_0)\right\|^2_K . \end{aligned} \tag{12}$$

This statistical test can be viewed as a projection of the full score statistic $S_n(\nu_0)$ on the subspace $\mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\Omega)_\perp$. In the following theorem we drive its asymptotic properties.

Theorem 2. *Under conditions of Theorem 1, we have*
(i) the limiting distribution of the score-type statistic (12) is given by

$$\zeta^S_{n,\Omega}(\nu_0) \xrightarrow{\mathcal{D}} \begin{cases} \chi^2_{K-r}, & \text{under } K_0; \\ \chi^2_{K-r}(\lambda^2), & \text{under } K_1^n, \end{cases}$$

where

$$\lambda^2 = \mathbf{h}^\top[\Gamma(\nu_0) - \Gamma(\nu_0)\Omega(\Omega^\top\Gamma(\nu_0)\Omega)^{-1}\Omega^\top\Gamma(\nu_0)]\mathbf{h} \tag{13}$$

is the non-centrality parameter of the chi-square r.v. χ^2_{K-r},
(ii) the locally asymptotically α*-level test for testing* K_0 *against* K^n*, with rejection region*

$$R := \left\{\zeta^S_{n,\Omega}(\nu_0) \geq \chi^2_{K-r,1-\alpha}\right\} \tag{14}$$

is most stringent
(iii) its asymptotic power is given by

$$1 - \Upsilon_{K-r}\left(\chi^2_{K-r,1-\alpha} - \lambda^2\right),$$

where Υ_{K-r} *stands for the distribution function of a* χ^2_{K-r} *r.v. and* $\chi^2_{K-r,1-\alpha}$ *represents its* $(1-\alpha)$*-quantile.*

To be operational, the unknown parameter ν in $\zeta^S_{n,\Omega}(\nu)$ should be estimated. We have then to deal with the problem of plug-in an estimator $\hat\nu_n$ of ν_0 in both $S_n(\nu_0)$ and $\Gamma^{-\frac{1}{2}}(\nu_0)$. The test statistic given in (12) becomes

$$\zeta^S_{n,\Omega}(\hat\nu_n) := S_n(\hat\nu_n)^\top[\Gamma_n^{-1}(\hat\nu_n) - \Omega(\Omega^\top\Gamma_n(\hat\nu_n)\Omega)^{-1}\Omega^\top]S_n(\hat\nu_n) \tag{15}$$

where $\Gamma_n(\nu)$ is the empirical version of the $\Gamma(\nu)$.

To obtain the asymptotic distribution of the test statistics given in (15), we first establish some asymptotic properties of the matrix Γ_n in Proposition 2 below. Then, this proposition will be used in the derivation of the uniform local asymptotic normality (ULAN) and the regularity of the score function. Our results are obtained under the following additional assumptions :

C3) -1 $E\left[\ell'(\epsilon_d)\epsilon_d\right] = -1$
-2 $E\left[\left(\ell''(\epsilon_d) + \ell'(\epsilon_d)^2\right)\epsilon_d^2\right] = 2$
-3 $E\left[\ell'(\epsilon_d)\right] = 0$
-4 $E\left[\ell''(\epsilon_d) + \ell'(\epsilon_d)^2\right] = 0$
-5 $E\left[\left(\ell''(\epsilon_d) - \ell'(\epsilon_d)^2\right)\epsilon_d\right] = 0$
-6 $E(|\epsilon_d|^{2(2+\gamma)}) < \infty$ for some constant $\gamma > 0$.

C4) There exists an integrable function $L(\cdot)$ such that $\sigma(x,\rho) \geq L(x)$ for all x and for ρ in a neighborhood of ρ_0.

C5) The estimators $(\hat\theta_n, \hat\rho_n)$ of the parameters (θ_0, ρ_0) are such that

$$\sqrt{n}(\hat\theta_n - \theta_0) = O_P(1) \qquad \sqrt{n}(\hat\rho_n - \rho_0) = O_P(1).$$

C6) **-1** There exist positive square integrable random functions $\gamma(\mathbf{X}_{i-1},\theta)$ and $\gamma'(\mathbf{X}_{i-1},\rho)$ and positive constants c_1 and c_2 such that for any θ^* and ρ^* with $||\theta^* - \theta_0||_q \leq c_1$ and $||\rho^* - \rho_0||_q \leq c_2$, we have

$$\left|\frac{\partial^k m(\mathbf{X}_{i-1},\theta^*)}{\partial\theta_s\partial\theta_t\partial\theta_u}\right| \leq \gamma(\mathbf{X}_{i-1},\theta), \qquad \left|\frac{\partial^k \sigma(\mathbf{X}_{i-1},\rho^*)}{\partial\rho_s\partial\rho_t\partial\rho_u}\right| \leq \gamma'(\mathbf{X}_{i-1},\rho)$$

for $k = 1,2,3$ and $s,t,u = 1,2,\ldots,q$.
-2 For the functions M_0, M_1 and L in C1 and C4, we have:

$$E\left(\frac{\gamma(X_{d-1},\theta)^k M_j(X_{d-1})^l}{L(X_{d-1})^k}\right) < \infty \text{ and } E\left(\frac{\gamma'(X_{d-1},\rho)^k M_j(X_{d-1})^l}{L(X_{d-1})^k}\right) < \infty \quad k=1,2; l=0,1; j=0,1.$$

Remark 1. Some of the previous assumptions have already been employed in Chebana and Laïb [1]. For instance, it can be easily seen that condition (C3) holds true whenever $\lim_{|x|\to\infty} x^{l-1} f^{(k)}(x) = 0$ for k and l taking values 1 or 2. It is satisfied, for instance, by standard centered normal distribution with variance σ^2, and t-distribution with degree of freedom greater than 3. Condition (C5) assumes $\sqrt{n}$-convergence of the estimators $\hat{\theta}_n$ and $\hat{\rho}_n$ which is satisfied by most estimators.

Proposition 2. *Assume that Conditions (C1)-(C6) are satisfied. Then we have, under K_0, that*

1. $\Gamma_n(\nu_0) \xrightarrow{P} \Gamma(\nu_0)$,
2. $\Gamma_n(\nu_n) - \Gamma_n(\nu_0) \xrightarrow{P} 0$ *uniformly in bounded* $\mathbf{h}$,
3. $\Gamma_n(\nu_n) \xrightarrow{P} \Gamma(\nu_0)$,

where $\nu_n = \nu + \mathbf{h}/\sqrt{n}$.

In Theorem 3 below, we drive the ULAN of the log-likelihood ratio Λ_n and the regularity of the score S_n. As a consequence of theses results, we obtain both the null and the non-null asymptotic law of the test statistics defined by (15), which are formulated in Corollary 2 below. Notice that, the ULAN result allows to replace the unknown parameter ν_0 by its consistent estimator $\hat{\nu}_n$ without any effect on the asymptotic behavior of the proposed test.

The log-likelihood ratio Λ_n is said ULAN if (i) for each $\nu \in \Theta$, the LAN property holds true for Λ_n, and (ii) $\sup_{\mathbf{h}} |\Lambda_n - \mathbf{h}^\top S_n(\nu) + \frac{1}{2}\mathbf{h}^\top \Gamma(\nu)\mathbf{h}| = o_P(1)$ under H_0, where the sup is taken over the set $\{\mathbf{h} : ||\mathbf{h}|| \le \tilde{C}\}$, for some fixed constant $0 < \tilde{C} < \infty$. The score $S_n(\nu)$ is said to be regular if

$$\forall \nu \in \Theta \quad S_n(\nu_n) = S_n(\nu) - \Gamma(\nu)\mathbf{h} + o_P(1) \quad \text{uniformly in bounded } \mathbf{h},$$

where $\nu_n = \nu + \mathbf{h}/\sqrt{n}$ (see Hall and Mathiason, 1990 for more details).

Theorem 3. *Under conditions (C1)-(C6), we have*

i) Λ_n *is ULAN.*
ii) $S_n(\cdot)$ *is regular.*
iii) For any $\sqrt{n}$*-consistent estimator* $\hat{\nu}_n$ *of* ν_0*, we have* $S_n(\hat{\nu}_n) = S_n(\nu_0) - \Gamma(\nu_0)\sqrt{n}(\hat{\nu}_n - \nu_0) + o_P(1)$ *under* H_0.

Corollary 2. *Under the conditions of Proposition 2, we have*

$$\zeta^S_{n,\Omega}(\hat{\nu}_n) = \zeta^S_{n,\Omega}(\nu_0) + o_P(1).$$

Remark 2. Comparing the statistic $\zeta^S_{n,\Omega}(\hat{\nu}_n)$ with $\hat{V}_n$ defined in Theorem 2.1 in Chebana and Laïb [1] for testing the nonparametric form of the functions $m(\cdot,\theta)$ and $\sigma(\cdot,\rho)$, one may observe that: (i) the estimation of the parameters does not affect the limiting distribution of $\zeta^S_{n,\Omega}(\hat{\nu}_n)$, however it reduces the asymptotic power of the test based on $\hat{V}_n$, (ii) there is a connection between $\zeta^S_{n,\Omega}(\hat{\nu}_n)$ and $\hat{V}_n$ in the case where $G \equiv \mathbf{h}_1^\top \dot{m}(\cdot,\theta_0)$ and $S \equiv \mathbf{h}_2^\top \dot{\sigma}(\cdot,\rho_0)$ (the functions G and S are related to the deviation of the local alternatives from the null, see Chebana and Laïb [1]). In such case, we get, under K_0, $\hat{V}_n = \mathbf{h}^\top S_n(\hat{\nu}_n)$ and $\tau_0^2 = \mathbf{h}^\top \Gamma(\nu_0)\mathbf{h}$, where τ_0 stands for the limiting variance of $\hat{V}_n$ and $\mathbf{h} = (1,\ldots,1)$.

3. Applications

In this section we consider a particular class of models for which the results of Section 2 can be applied. From a specified example of this class we formulate the corresponding statistics and the asymptotic quantities in Section 2.

3.1. Particular class of models and its properties

Let $g_1,\ldots,g_r$ and $v_1,\ldots,v_r$ be given real-valued functions on $\mathbb{R}$. Consider the special case of model (1) for which

$$m(x,\theta) = g_1(x)\theta_1 + \cdots + g_r(x)\theta_r \quad \text{and} \quad \sigma^2(x,\rho) = v_1(x)\rho_1^2 + \cdots + v_r(x)\rho_r^2. \tag{16}$$

This class, where the functions $m(.,.)$ and $\sigma(.,.)$ have an additive form, includes some known examples of nonlinear time series models given in Tong [10] and Taniguchi and Kakizawa [9] such as AR, EXPAR, ARCH and β-ARCH. In order to establish our results for this class of models, the following assumptions are required:

AP1) The ϵ_i's are iid with common nonnegative density function f such that $E|\epsilon_1| < \infty$.
AP2) For $k = 1 \ldots r$, g_k and v_k are Lipschitzian functions, and there exists a positive constant κ such that $\sigma(\cdot, \rho) \geq \kappa > 0$ for a neighborhood of ρ_0.
AP3) For all $x \in \mathbb{R}$ there exist $\alpha_k \geq 0, \alpha'_k \geq 0$ and $\beta_k \geq 0, k = 1 \ldots r$ such that $|g_k(x)| \leq \alpha_k|x|$ and $|v_k(x)| \leq \alpha'_k|x|^2+\beta_k$ with

$$\sum_{k=1}^{r} \alpha_k|\theta_k| + \max(1, E|\epsilon_1|) \left(\sum_{k=1}^{r} \alpha'_k \rho_k^2\right)^{1/2} < 1. \tag{17}$$

AP4) g_k and v_k are non null real-valued functions such that $Eg_k^4(X_d) < \infty$ and $Ev_k^2(X_d) < \infty$ for $k = 1 \ldots r$.

In (AP3) the constant β_k serves to bound the function on compact subsets while the power function to control the growth of the function on the tails.

The proposition below summarizes the statistical and probabilistic properties of the class of models specified by (16).

Proposition 3.

1. *Under (AP1)-(AP3), the model defined through (16) is stationary and ergodic.*
2. *If the parameter ν is estimated by the conditional least squares estimator $\hat{\nu}_n$, then, under (AP2) and (AP4), the class of models (16) satisfies the conditions (C1), (C2) and (C4)-(C6).*

Note that (C3) is not related to the functions $m(.,.)$ and $\sigma(.,.)$.

3.2. Example

The aim of this example is to check the required assumptions and to explicit all the asymptotic quantities defined in Section 2. Let us take in (16)

$$g_1(x) = x, \quad g_2(x) = xe^{-\varsigma x^2}, \quad v_1(x) = 1 \quad \text{and} \quad v_2(x) = x^2 e^{-\eta x^2}$$

with ς, η are positive constants and $\rho_1 > 0$. Then the model specified by (16) reduces to

$$X_i = \theta_1 X_{i-1} + \theta_2 X_{i-1}\, e^{-\varsigma\, X_{i-1}^2} + \sqrt{\rho_1^2 + \rho_2^2\, X_{i-1}^2\, e^{-\eta\, X_{i-1}^2}}\; \epsilon_i. \tag{18}$$

Suppose that the ϵ_i's are iid with standard normal distribution, $E|X_1|^4 < \infty$ and $|\theta_1| + |\theta_2| + \rho_2 < 1$.
It is clear that (AP1) is satisfied since the ϵ_i's are iid standard normally distributed. Moreover, (AP2) is fulfilled since the function $x \mapsto xe^{-\varsigma x^2}$ is Lipschitzian. The condition (AP3) is also satisfied by taking $\alpha_1 = \alpha_2 = 1, \alpha'_1 = 0, \alpha'_2 = 1, \beta'_1 \geq 1, \beta'_2 \geq 0$, in this case the condition (17) being $|\theta_1| + |\theta_2| + \rho_2 < 1$ since $E|\epsilon_1| \leq 1$. Therefore, model (18) is stationary and ergodic by Proposition 3. The assumption (AP4) is fulfilled since the ϵ_i's normally distributed. The assumption (C3) is concerned with the regularity of the density f, it is satisfied in our setting since f is supposed to be a Gaussian density function whereas (C1) and (C6) are trivially satisfied by taking $M_0(x) = |x|/\rho_1$, $M_1(x) = \max(1, x^2e^{-\eta x^2})/\rho_1$, $\gamma(x, \theta) = x, \gamma'(x, \rho) = 1/\sigma(x, \rho)^3$ and $L(x) = x^2$. Finally, (C2) and (C4) follow from (AP2) and (AP4) whereas the assumption (C5) is satisfied with the conditional least squares estimators $\hat{\theta}_n$ and $\hat{\rho}_n$ defined by

$$\hat{\theta}_n = \text{argmin}_\theta \sum_{i=1}^{n} (X_i - m(X_{i-1}, \theta))^2 \text{ and } \hat{\rho}_n = \text{argmin}_\rho \sum_{i=1}^{n} \left(\left[X_i - m(X_{i-1}, \hat{\theta}_n)\right]^2 - \sigma^2(X_{i-1}, \rho)\right)^2.$$

Furthermore, $I_0 = E\epsilon_1^2 = 1, I_1 = E\epsilon_1^3 = 0$ and $I_2 = E\epsilon_1^4 = 3$.

The score function $S_n(\nu)$ and the matrix $\Gamma(\nu)$ defined in Theorem 1 are then

$$S_n(\nu) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{X_{i-1}\epsilon_i}{\sigma(X_{i-1}, \rho)}, \frac{X_{i-1}e^{-\varsigma X_{i-1}^2}\epsilon_i}{\sigma(X_{i-1}, \rho)}, \frac{\rho_1(\epsilon_i^2 - 1)}{\sigma^2(X_{i-1}, \rho)}, \frac{\rho_2 X_{i-1}^2 e^{-\eta X_{i-1}^2}}{\sigma^2(X_{i-1}, \rho)}(\epsilon_i^2 - 1)\right]^\top$$

$$\Gamma(\nu) = \begin{pmatrix} E\frac{X_1^2}{\sigma^2(X_1,\rho)} & E\frac{X_1^2 e^{-\varsigma X_1^2}}{\sigma^2(X_1,\rho)} & 0 & 0 \\ E\frac{X_1^2 e^{-\varsigma X_1^2}}{\sigma^2(X_1,\rho)} & E\frac{X_1^2 e^{-2\varsigma X_1^2}}{\sigma^2(X_1,\rho)} & 0 & 0 \\ 0 & 0 & E\frac{\rho_1^2}{\sigma_\rho^4(X_1)} & E\frac{\rho_1\rho_2 X_1^2 e^{-\eta X_1^2}}{\sigma^4(X_1,\rho)} \\ 0 & 0 & E\frac{\rho_1\rho_2 X_1^2 e^{-\eta X_1^2}}{\sigma^4(X_1,\rho)} & E\frac{\rho_2^2 X_1^4 e^{-2\eta X_1^2}}{\sigma^4(X_1,\rho)} \end{pmatrix}$$

In the following, we treat two situations, the first one concerned with simple hypotheses and the second one concerns hypotheses with presence of nuisance parameters.

(i) In the first case we test the hypotheses $H_0 : \nu = \nu_0$ versus $H_1^n : \nu = \nu_n$, where $\nu_0 = (\theta_{1,0}, \theta_{2,0}, \rho_{1,0}, \rho_{2,0})$ is known, $\nu_n = \nu_0 + \mathbf{h}/\sqrt{n}$ and $\mathbf{h} = (h_1, h_2, h_3, h_4)$. Here, we have $\Omega \equiv 0$ as described in Section 2.2. Consequently, we get

$$D_{n,0} := \zeta_{n,\Omega}^S(\nu_0) = S_n(\nu_0)^\top \Gamma(\nu_0) S_n(\nu_0)$$

and Theorem 2 leads to

$$D_{n,0} \xrightarrow{\mathcal{D}} \begin{cases} \chi_4^2, & \text{under } H_0; \\ \chi_4^2(\lambda_0^2), & \text{under } H_1^n, \end{cases}$$

where

$$\lambda_0^2 = \mathbf{h}^\top \Gamma_n(\nu_0) \mathbf{h}.$$

The corresponding rejection region is $\{D_{n,0} \geq \chi_{4,1-\alpha}^2\}$. At a level $\alpha = 0.05$, we have $\chi_{4,0.95}^2 = 9.49$. Therefore, the asymptotic power is

$$1 - \Upsilon_4\left(\chi_{4,0.95}^2 - \lambda_0^2\right) \text{ for } \lambda_0^2 \leq 9.49, \text{ and } = 1 \text{ elsewhere.}$$

(ii) The hypotheses to be tested in the second case are

$$K_0 : \theta_1 = \theta_{1,0}, \quad \rho_1 = \rho_{1,0} \qquad \text{versus} \qquad K_1^n : \theta_1 = \theta_{1,n}, \quad \rho_1 = \rho_{1,n}$$

where $\theta_{1,n} = \theta_{1,0} + h_1/\sqrt{n}$ and $\rho_{1,n} = \rho_{1,0} + h_4/\sqrt{n}$. In this case, θ_1 and ρ_1 represent the parameters of interest and the others are nuisance parameters. The corresponding matrix Ω takes the form:

$$\Omega = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^\top,$$

and the statistic $\zeta_{n,\Omega}^S(\hat{\nu}_n)$ given in (15) can be written explicitly as :

$$\hat{D}_n = \zeta_{n,\Omega}^S(\hat{\nu}_n) = \frac{\hat{\Gamma}_{22}}{\hat{A}} \hat{S}_{n1}^2 - 2\frac{\hat{\Gamma}_{12}}{\hat{A}} \hat{S}_{n1}\hat{S}_{n2} + \frac{\hat{\Gamma}_{12}^2}{\hat{A}\hat{\Gamma}_{22}} \hat{S}_{n2}^2 + \frac{\hat{\Gamma}_{44}}{\hat{B}} \hat{S}_{n3}^2 - 2\frac{\hat{\Gamma}_{34}}{\hat{B}} \hat{S}_{n3}\hat{S}_{n4} + \frac{\hat{\Gamma}_{34}^2}{\hat{B}\hat{\Gamma}_{44}} \hat{S}_{n4}^2$$

in which $\left(\hat{S}_{nk}\right)_{k=1\ldots4}$, $\left[\hat{\Gamma}_{ij}\right]_{i,j=1\ldots4}$, $\hat{A}$ and $\hat{B}$ are respectively the empirical versions of the elements of $S_n(\nu)$, $\Gamma(\nu)$, A and B when replacing ν_0 by its estimator $\hat{\nu}_n$ with $A(\nu_0) = \Gamma_{11}(\nu_0)\Gamma_{22}(\nu_0) - \Gamma_{12}^2(\nu_0)$ and $B(\nu_0) = \Gamma_{33}(\nu_0)\Gamma_{44}(\nu_0) - \Gamma_{34}^2(\nu_0)$. According to Theorem 2, we have:

$$\hat{D}_n \xrightarrow{\mathcal{D}} \begin{cases} \chi_2^2, & \text{under } K_0; \\ \chi_2^2(\lambda^2), & \text{under } K_1^n, \end{cases}$$

where $\lambda^2 = h_1^2 \frac{A(\nu_0)}{\Gamma_{22}(\nu_0)} + h_3^2 \frac{B(\nu_0)}{\Gamma_{44}(\nu_0)}$. At a nominal level $\alpha = 0.05$, the asymptotic power of the test based on $\hat{D}_n$ which rejects the null hypothesis if $\left\{\hat{D}_n \geq \chi_{2,1-\alpha}^2\right\}$ is given by :

$$1 - \Upsilon_2\left(\chi_{2,1-0.05}^2 - \lambda^2\right) = \exp\left(\frac{\lambda^2 - 5.99}{2}\right) \text{ for } \lambda^2 \leq 5.99 \text{ and } = 1 \text{ elsewhere.}$$

4. Proofs

Proof of Proposition 1. Making use of Vitali's lemma, it suffices then to show that

$$E\left[\frac{1}{t}\left\{\phi_i(\nu_0 + t\mathbf{h}; \nu_0) - 1\right\}\right]^2 \to E\left[\mathbf{h}^\top \dot{\phi}_i(\nu_0)\right]^2 \text{ as } t \to 0.$$

Moreover, using Fatou's lemma we can see that the above statement holds true whenever

$$\limsup_{t\to 0} E\left[\frac{1}{t}\left\{\phi_i(\nu_0 + t\mathbf{h}; \nu_0) - 1\right\}\right]^2 \leq E\left[\mathbf{h}^\top \dot{\phi}_i(\nu_0)\right]^2.$$

Recall that $g_\nu(X_i|y) = \frac{1}{\sigma(y,\rho_0)} f\left(\frac{X_i - m(y,\theta_0)}{\sigma(y,\rho_0)}\right)$. Conditioning by $\mathbf{X}_{i-1}$ one may write

$$\begin{aligned} E\left[\left\{\phi_i(\nu_0 + t\mathbf{h};\nu_0) - 1\right\}^2 \middle| \mathbf{X}_{i-1}\right] &= E\left[\frac{\left\{g^{\frac{1}{2}}_{\nu_0+t\mathbf{h}}(X_i|\mathbf{X}_{i-1}) - g^{\frac{1}{2}}_{\nu_0}(X_i|\mathbf{X}_{i-1})\right\}^2}{g_{\nu_0}(X_i|\mathbf{X}_{i-1})}\middle| \mathbf{X}_{i-1}\right] \\ &= \int\left\{\sqrt{g_{\nu_0+t\mathbf{h}}(X_i|y)} - \sqrt{g_{\nu_0}(X_i|y)}\right\}^2 dy. \end{aligned} \tag{19}$$

Moreover, observe that

$$g^{1/2}_{\nu_0+t\mathbf{h}}(X_i|y) - g^{1/2}_{\nu_0}(X_i|y) = \int_0^1 \frac{d}{ds} g^{1/2}_{\nu_0+ts\mathbf{h}}(X_i|y)ds = \int_0^1 \frac{\frac{d}{ds} g_{\nu+ts\mathbf{h}}(X_i|y)}{2g^{1/2}_{\nu+ts\mathbf{h}}(X_i|y)}ds.$$

Simple calculations show that

$$\begin{aligned} \frac{d}{ds} g_{\nu+ts\mathbf{h}}(X_i|y) &= -t\mathbf{h}^\top \nabla g_{\nu+ts\mathbf{h}}(X_i|y) \\ &= \frac{-t(\mathbf{h_1}^\top, \mathbf{h_2}^\top)}{\sigma^2(y,\rho+st\mathbf{h}_2)} \times \left\{\nabla m(y,\theta+st\mathbf{h}_1) f'\left(\frac{X_i - m(y,\theta+st\mathbf{h}_1)}{\sigma(y,\rho+st\mathbf{h}_2)}\right),\right. \\ &\left.\nabla\sigma(y,\rho+st\mathbf{h}_2)\left[f\left(\frac{X_i - m(y,\theta+st\mathbf{h}_1}{\sigma(y,\rho+st\mathbf{h}_2)}\right) + \frac{X_i - m(y,\theta+st\mathbf{h}_1}{\sigma(y,\rho+st\mathbf{h}_2)} f'\left(\frac{X_i - m(y,\theta+st\mathbf{h}_1)}{\sigma(y,\rho+st\mathbf{h}_2)}\right)\right]\right\}. \end{aligned}$$

It results from Hölder's inequality that

$$\begin{aligned} E\left[\left\{\phi_i(\nu_0 + t\mathbf{h};\nu_0) - 1\right\}^2 \middle| \mathbf{X}_{i-1}\right] &= t^2 \int\left\{\int_0^1 \frac{\mathbf{h}^\top \nabla g_{\nu+ts\mathbf{h}}(X_i|y)}{2g^{1/2}_{\nu+ts\mathbf{h}}(X_i|y)} ds\right\}^2 dy \\ &\leq t^2 \int\int_0^1 \left[\frac{\mathbf{h}^\top \nabla g_{\nu+ts\mathbf{h}}(X_i|y)}{2g_{\nu+ts\mathbf{h}}(X_i|y)}\right]^2 g_{\nu+ts\mathbf{h}}(X_i|y) ds dy \end{aligned}$$

Conditions (C1) and (C2) and Lebesgue dominated convergence theorem allow us to write

$$\begin{aligned} \limsup_{t\to 0} E\left[\left\{\frac{1}{t}\phi_i(\nu_0 + t\mathbf{h};\nu_0) - 1\right\}^2 \middle| \mathbf{X}_{i-1}\right] &\leq \int\left[\frac{\mathbf{h}^\top \nabla g_\nu(X_i|y)}{2g_\nu(X_i|y)}\right]^2 g_\nu(X_i|y) dy \\ &= E\left[\mathbf{h}^\top \dot{\phi}_i(\nu_0) \,|\mathbf{X}_{i-1}\right]^2. \end{aligned}$$

Taking the expectation of the two sides of the above inequality, we get

$$\limsup_{t\to 0} E\left[\frac{1}{t}\left\{\phi_i(\nu_0 + t\mathbf{h};\nu_0) - 1\right\}\right]^2 \leq E\left[\mathbf{h}^\top \dot{\phi}_i(\nu_0)\right]^2,$$

which completes the proof. □

Proof of Theorem 2

(i) From Corollary 1, we have under H_1^n that $S_n(\nu_0) \xrightarrow{\mathcal{D}} N(\Gamma(\nu_0)h, \Gamma(\nu_0))$. So that

$$\Gamma^{-\frac{1}{2}}(\nu_0) S_n(\nu_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\Gamma^{\frac{1}{2}}(\nu_0)\mathbf{h}, Id). \tag{20}$$

Cochran's Theorem leads to

$$S_n^\top(\nu_0)\Gamma^{-\frac{1}{2}}(\nu_0)\left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{-\frac{1}{2}}(\nu) S_n(\nu_0) = \zeta^S_{n,\Omega}(\nu_0) \xrightarrow{\mathcal{D}} \chi^2_{K-r}(\lambda^2), \tag{21}$$

with $\lambda^2 = h^\top \Gamma^{\frac{1}{2}}(\nu_0)\left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{\frac{1}{2}}(\nu_0) h = h^\top[\Gamma(\nu_0) - \Gamma(\nu_0)\Omega(\Omega^\top\Gamma(\nu_0)\Omega)^{-1}\Omega^\top\Gamma(\nu_0)]h$, which gives the desired result under H_1^n. In addition we have, under H_0, $\lambda = 0$ since $\left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{\frac{1}{2}}(\nu_0)h = 0$.

(ii) The most stringency is a result of the weak convergence of local experiments to Gaussian shifts (see Le Cam [6]).

(iii) Follows from (i) and (ii). □

Proof of Proposition 2

(1) Simple algebra calculus allow to write the symmetric matrix $\Gamma_n(\nu)$ as

$$\Gamma_n(\nu) = -\frac{1}{n}\sum_{i=d}^{n}\begin{bmatrix} -\hat{I}_0\frac{\nabla m(\mathbf{X}_{i-1},\theta)\nabla m^\top(\mathbf{X}_{i-1},\theta)}{\sigma^2(\mathbf{X}_{i-1},\rho)} & \hat{I}_1\frac{\nabla m(\mathbf{X}_{i-1},\theta)\nabla\sigma^\top(\mathbf{X}_{i-1},\rho)}{\sigma^2(\mathbf{X}_{i-1},\rho)}\epsilon_i\ell''(\epsilon_i) \\ \hat{I}_1\frac{\nabla\sigma(\mathbf{X}_{i-1},\rho)\nabla m^\top(\mathbf{X}_{i-1},\theta)}{\sigma^2(\mathbf{X}_{i-1},\rho)}\epsilon_i\ell''(\epsilon_i) & (\hat{I}_2-1)\frac{\nabla\sigma(\mathbf{X}_{i-1},\rho)\nabla\sigma^\top(\mathbf{X}_{i-1},\rho)}{\sigma^2(\mathbf{X}_{i-1},\rho)}(\epsilon_i^2\ell''(\epsilon_i)-1)\end{bmatrix}.$$

Statement (1) follows then from an application of the ergodic theorem combined with the independence between X_{i-1} and ϵ_i and Conditions (C1) and (C3).

(2) We have to show that $\Gamma_n(\nu_n)-\Gamma_n(\nu_0)\xrightarrow{P}0$ uniformly in any $2q\times 1$ vector bounded h as $n\to\infty$. This result will be proved if we can prove that

$$\sup_{\mathbf{h},||\mathbf{h}||_{2q}<M}||\Gamma_n(\nu_n)-\Gamma_n(\nu_0)||_{\mathcal{M}} = o_P(1). \tag{22}$$

In order to check condition (22), let us write the $2q\times 2q$ matrix $\Gamma_n(\nu)$ as $\begin{pmatrix}\Gamma^{11}(\nu) & \Gamma^{12}(\nu)\\ \Gamma^{21}(\nu) & \Gamma^{22}(\nu)\end{pmatrix}$. Moreover, for $1\le s\le t\le q$ and $1\le k,m\le 2$, denote by $\Gamma^{km}_{st}(\nu)$ the (s,t)th real-valued element of the $q\times q$ matrix $\Gamma^{km}(\nu)$. We have then to check condition (22) for each $\Gamma^{km}(\nu)$. To this end, making use of Taylor expansion of the function $\Gamma^{km}_{st}(\nu)$ around ν_0 we may write

$$\Gamma^{km}_{st}(\nu_n)-\Gamma^{km}_{st}(\nu_0) = (\theta_n-\theta_0)^\top\frac{\partial}{\partial\theta}\Gamma^{km}_{st}(\nu_n^*)+(\rho_n-\rho_0)^\top\frac{\partial}{\partial\rho}\Gamma^{km}_{st}(\nu_n^*),\quad \nu_n^*\text{ is a point between }\nu_n\text{ and }\nu_0,$$

where $\theta_n,\theta_0,\rho_n,\rho_0$ are such that

$$\theta_n=\theta_0+\frac{h_1}{\sqrt{n}}\quad\text{and}\quad\rho_n=\rho_0+\frac{h_2}{\sqrt{n}},$$

with h_1 and h_2 are $q\times 1$ vectors. Thus, we have for $1\le k,m\le 2$:

$$\Gamma^{km}(\nu_n)-\Gamma^{km}(\nu_0) = \left(\frac{h_1^\top}{\sqrt{n}}\frac{\partial}{\partial\theta}\Gamma^{km}_{st}(\nu^*)+\frac{h_2^\top}{\sqrt{n}}\frac{\partial}{\partial\rho}\Gamma^{km}_{st}(\nu^*)\right)_{1\le s,t\le q}. \tag{23}$$

Therefore, to check condition (22), it suffices then to verify that each component of (23) goes to 0 in probability as $n\to\infty$. Consider the first term $\frac{h_1^\top}{\sqrt{n}}\frac{\partial}{\partial\theta}\Gamma^{11}_{st}(\nu^*)+\frac{h_2^\top}{\sqrt{n}}\frac{\partial}{\partial\rho}\Gamma^{11}_{st}(\nu^*)$ in which

$$\Gamma^{11}_{st}(\nu) = -\frac{1}{n}\sum_{i=d}^{n}\left[\frac{\partial m(\mathbf{X}_{i-1},\theta)}{\partial\theta_s}\times\frac{\partial m(\mathbf{X}_{i-1},\theta)}{\partial\theta_t}\right]\times\frac{\ell''(\epsilon_i)}{\sigma^2(\mathbf{X}_{i-1},\rho)}$$

Using (C4) and (C6), one can easily see that

$$\begin{aligned}\left|\frac{h_1^\top}{\sqrt{n}}\frac{\partial}{\partial\theta}\Gamma^{11}_{st}(\nu^*)+\frac{h_2^\top}{\sqrt{n}}\frac{\partial}{\partial\rho}\Gamma^{11}_{st}(\nu^*)\right| &\le \frac{M}{\sqrt{n}}\left\{||\frac{\partial}{\partial\theta}\Gamma^{11}_{st}(\nu^*)||_q+||\frac{\partial}{\partial\rho}\Gamma^{11}_{st}(\nu^*)||_q\right\}\\ &\le \frac{M}{n^{3/2}}\sum_{i=d}^{n}\frac{\gamma^2(\mathbf{X}_{i-1},\theta)}{L(\mathbf{X}_{i-1})}[|\ell''(\epsilon^*)|+M_0(\mathbf{X}_{i-1})|\ell'''(\epsilon_i^*)|]\\ &\quad+\frac{M}{n^{3/2}}\sum_{i=d}^{n}\frac{\gamma(\mathbf{X}_{i-1},\theta)}{L(\mathbf{X}_{i-1})}\{|\ell'(\epsilon_i^*)|+M_0(\mathbf{X}_{i-1})|\ell''(\epsilon_i^*)|\}\\ &\quad+\frac{M}{n^{3/2}}\sum_{i=d}^{n}\frac{\gamma'^2(\mathbf{X}_{i-1},\theta)}{L^2(\mathbf{X}_{i-1})}\{M_1(\mathbf{X}_{i-1})|\ell''(\epsilon_i^*)|+M_0(\mathbf{X}_{i-1})|\epsilon_i^*\ell'''(\epsilon_i^*)|\}\\ &\quad+\frac{M}{n^{3/2}}\sum_{i=d}^{n}\frac{\gamma'(\mathbf{X}_{i-1},\theta)}{L(\mathbf{X}_{i-1})}\{M_1(\mathbf{X}_{i-1})|\ell''(\epsilon_i^*)|+M_0(\mathbf{X}_{i-1})|\epsilon_i^*\ell'''(\epsilon_i^*)|\}.\end{aligned} \tag{24}$$

To obtain the desired result, we have to show that each term in (24) is $o_P(1)$. The other terms in (23) can be handled similarly. The above terms contain summation of quantities of the form ${\epsilon_i^*}^k\ell^{(m)}(\epsilon_i^*)$. To be concise, we evaluate only one of these terms which is of the form

$$\frac{1}{n^{3/2}}\sum_{i=d}^{n}\psi(\mathbf{X}_{i-1}){\epsilon_i^*}^k\ell^{(m)}(\epsilon_i^*)$$

for a given measurable function ψ with $k = 0, 1, 2$ and $m = 1, 2, 3$ where $\ell^{(m)}$ stands for the derivative of order m of the function ℓ. We have

$$\epsilon_i^{*k}\ell^{(m)}(\epsilon_i^*) = (\epsilon_i^{*k} - \epsilon_i^k)\left(\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right) + \epsilon_i^k\left(\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right) + (\epsilon_i^{*k} - \epsilon_i^k)\ell^{(m)}(\epsilon_i) + \epsilon_i^k\ell^{(m)}(\epsilon_i).$$

Observe that

$$\begin{aligned}\frac{1}{n^{3/2}}\left|\sum_{i=d}^n \psi(\mathbf{X}_{i-1})\epsilon_i^{*k}\ell^{(m)}(\epsilon_i^*)\right| \le\ & \max_{d\le i\le n}\frac{\left|(\epsilon_i^{*k} - \epsilon_i^k)\left(\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right)\right|}{\sqrt{n}}\frac{1}{n}\sum_{i=d}^n|\psi(\mathbf{X}_{i-1})| \\ &+ \max_{d\le i\le n}\frac{\left|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right|}{\sqrt{n}}\frac{1}{n}\sum_{i=d}^n|\epsilon_i\psi(\mathbf{X}_{i-1})| \\ &+ \max_{d\le i\le n}\frac{\left|\epsilon_i^{*k} - \epsilon_i^k\right|}{\sqrt{n}}\frac{1}{n}\sum_{i=d}^n|\ell^{(m)}(\epsilon_i)\psi(\mathbf{X}_{i-1})| \\ &+ \frac{1}{\sqrt{n}}\frac{1}{n}\sum_{i=d}^n|\epsilon_i^k\ell^{(m)}(\epsilon_i)\psi(\mathbf{X}_{i-1})|.\end{aligned} \tag{25}$$

By the ergodic theorem, the independence between the X_{i-1}'s and ϵ_i's and condition (C6), each sum in (25) is asymptotically bounded a.s.

To obtain the desired result, it suffices then to check that each max term in (25) is $o_P(1)$, that is

$$\max_{d\le i\le n}\left|\epsilon_i^{*k} - \epsilon_i^k\right| = o_P(1) \quad \text{and} \quad \max_{d\le i\le n}\frac{\left|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right|}{\sqrt{n}} = o_P(1). \tag{26}$$

Making use of Taylor expansion of ϵ_i^{*k} (as a function of the parameter $\nu^* = (\theta^*, \rho^*)$) around ν_0, we may write, under H_0

$$\epsilon_i^{*k} - \epsilon_i^k = k(\theta_n^* - \theta_0)^\top\epsilon_i^{**k-1}\frac{\nabla m(\mathbf{X}_{i-1}, \theta_n^{**})}{\sigma(\mathbf{X}_{i-1}, \rho_n^{**})} - k(\rho_n^* - \rho_0)^\top\epsilon_i^{**k}\frac{\nabla\sigma(\mathbf{X}_{i-1}, \rho_n^{**})}{\sigma(\mathbf{X}_{i-1}, \rho_n^{**})},$$

where θ_n^{**} and ρ_n^{**} are intermediate points between θ_0 and θ_n^*, ρ_0 and ρ_n^*, respectively. Condition (C5) and the continuity of the functions $m(.,.)$ and $\sigma(.,.)$ with respect to θ and ρ yield

$$\max_{d\le i\le n}|\epsilon_i^{*k} - \epsilon_i^k| = O_P(1)\max_{d\le i\le n}\frac{|\epsilon_i^{k-1}M_0(\mathbf{X}_{i-1})|}{\sqrt{n}} + O_P(1)\max_{d\le i\le n}\frac{|M_1(\mathbf{X}_{i-1})\epsilon_i^k|}{\sqrt{n}}.$$

Now using Markov's inequality, it follows from Conditions C1-(1) and C3-(6) that

$$\max_{d\le i\le n}|\epsilon_i^{*k} - \epsilon_i^k| = o_P(1). \tag{27}$$

To deal now with the second term in (26) we have for any $\eta > 0$ and some $\gamma > 0$ that

$$\begin{aligned}\mathbb{P}\left\{\max_{1\le i\le n}\frac{\left|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right|}{\sqrt{n}} \ge \eta\right\} =\ & \mathbb{P}\left\{\max_{1\le i\le n}\left|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)\right|^{\gamma+1} \ge (\sqrt{n}\eta)^{\gamma+1}\right\} \\ \le\ & \mathbb{P}\left\{\max_{1\le i\le n}|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)|^{\gamma+1}1_{\{|\epsilon_i|\le L\}} \ge (\sqrt{n}\eta)^{(\gamma+1)/2}\right\} \\ +\ & \mathbb{P}\left\{\max_{d\le i\le n}|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)|^{\delta}1_{\{|\epsilon_i|>L\}} \ge (\sqrt{n}\eta)^{(\gamma+1)/2}\right\},\end{aligned} \tag{28}$$

where L is a large positive constant. Since $\ell^{(m)}$ is continuous, it is then uniformly continuous on the compact set $[-L, L]$. This fact combined with the statement (27) and Conditions C1-(1) and C3-(6) implies that the first term in the right hand side of (28) is $o_P(1)$.

For the second term in the right hand side of (28), observe that

$$\left\{\max_{d\le i\le n}|\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)|^{\gamma+1}1_{\{|\epsilon_i|>L\}} \ne 0\right\} \subset \{\exists i_0, d \le i_0 \le n; |\epsilon_{i_0}| > L\}.$$

By choosing $L = L_n = n$, we obtain by stationarity, Markov's inequality and Condition C3-(6) that

$$\mathbb{P}\left\{\max_{d\le i\le n} |\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)|^{\gamma+1} 1_{\{|\epsilon_i|>L\}} \neq 0\right\} \le nP(|\epsilon_0| > L_n) = nL_n^{-1-\gamma} E|\epsilon_d|^{1+\gamma} = O(n^{-\gamma}),$$

il follows then by Borel-Contelli's Lemma that

$$\limsup_{n\to\infty} \max_{d\le i\le n} |\ell^{(m)}(\epsilon_i^*) - \ell^{(m)}(\epsilon_i)| 1_{\{|\epsilon_i|>L\}} = 0 \quad \text{a.s..}$$

The first term in (26) may be handled similarly. This achieves the proof of part 2) of the proposition.

(3) This statement is obtained as a consequence of the uniform convergence established in part 2) as well as the result in part 1). □

Proof of Theorem 3

By a Taylor expansion of Λ_n and S_n around ν_0, we get

$$\Lambda_n = \mathbf{h}^\top S_n(\nu_0) - \frac{1}{2}\mathbf{h}^T W_n(\nu_n^*)\mathbf{h} \quad \text{and} \quad S_n(\nu_n) = S_n(\nu_0) - \frac{1}{2} W_n(\nu_n^*)\mathbf{h},$$

where ν_n^* is an intermediate point between ν_0 and $\nu_n = \nu_0 + h/\sqrt{n}$. The statements i) and ii) follow then from the second part of Proposition 2, whereas iii) is a direct consequence of ii). □

Proof of Corollary 2

We have from Theorem 3 that

$$S_n(\hat{\nu}_n) = S_n(\nu_0) - \Gamma(\nu_0)\sqrt{n}(\hat{\nu}_n - \nu_0) + o_P(1), \text{ under } K_0,$$

then $\Gamma^{-\frac{1}{2}}(\nu_0)S_n(\hat{\nu}_n) - \Gamma^{-\frac{1}{2}}(\nu_0)S_n(\nu_0) = \Gamma^{\frac{1}{2}}(\nu_0)\sqrt{n}(\hat{\nu}_n - \nu_0) + o_P(1)$.

On the other hand $\Gamma^{\frac{1}{2}}(\nu_0)\sqrt{n}(\hat{\nu}_n - \nu_0) \in M(\Gamma^{\frac{1}{2}}(\nu_0)\Omega)$, under K_0. Thus

$$\left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{-\frac{1}{2}}(\nu_0)S_n(\hat{\nu}_n) = \left[\Gamma^{\frac{1}{2}}(\nu_0)\Omega\right]_\perp \Gamma^{-\frac{1}{2}}(\nu_0)S_n(\nu_0) + o_P(1),$$

which leads by Cochran's Theorem to

$$\left\|\left[Id - P_{\mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\Omega)}\right]\Gamma^{-\frac{1}{2}}(\nu_0)S_n(\nu_0)\right\|_K^2 = \left\|\left[Id - P_{\mathcal{M}(\Gamma^{\frac{1}{2}}(\nu_0)\Omega)}\right]\Gamma^{-\frac{1}{2}}(\nu_0)S_n(\hat{\nu}_n)\right\|_K^2 + o_P(1).$$

The result can be obtained using Proposition 2 when substituting ν_0 by its estimator. The contiguity of the hypotheses allows to get the same conclusion under the local alternatives. □

Proof of Proposition 3

1. To prove the strict stationarity of the model based on functions (16), it suffices to check the conditions (S1)-(S4) of Theorem 3.2.11 in Tanuguchi and Kakizawa [9, page 86]. Conditions (S1) and (S2) are satisfied since the ϵ_i's are iid and by (AP2) the functions g_k and v_k are continuous on R. Moreover, (S3) holds whenever g_k and v_k are Lipschitzian functions and $E|\epsilon_1| < \infty$ which are satisfied by (AP1) and (AP2). The condition (S4) is also satisfied under (AP3).
   A sufficient condition for the geometric ergodicity can be obtained for the above model, if we check that
   $$\limsup_{|x|\to\infty} \frac{E|m(x,\theta) + \sigma(x,\rho)\epsilon_1|}{|x|} < 1$$
   (see Doukhan [2, pages 106-107]), which is fulfilled by (AP3).
2. It's clear that assumption (C2) is satisfied, whereas (C1) holds by taking $M_0(x) = \max_{1\le k\le r} |g_k(x)|/\kappa$ and $M_1(x) = \max_{1\le k\le r} v_k(x)/\kappa$. The assumption (C5) is fulfilled with the conditional least squared estimators. Finally, (C4) and (C6) are clearly satisfied by (AP2). □

Acknowledgment
The authors would like to thank the Editors and the anonymous reviewers for their comments and suggestions that improved the quality of the paper.

References

[1] Chebana, F. and Laïb, N., 2008. Local asymptotic optimal test for nonlinear autoregressive models. Comptes Rendus de l'Académie des Sciences Paris, Serie I, 346, 671-676.
[2] Doukhan, P.M., 1994. Volume 85 of Lecture Notes in Statistics. Springer-Verlag, New York.
[3] Hall, W.J. and Mathiason, D.J., 1990. On large-sample estimation and testing in parametric models. *Internat. Statist. Rev.*, **58**, 77-97.
[4] Hwang, S.Y. and Basawa I.V., 1993. Asymptotic optimal inference for a class of nonlinear time series models. *Stochastic Process. Appl.*, **46**(1), 91-113.
[5] Laïb, N., 2003. Non-parametric testing of conditional variance functions in time series. *Aust. N. Z. J. Stat.*, **45**(4), 461-475.
[6] Le, C.L., 1986. **Asymptotic methods in statistical decision theory**. Springer-Verlag, New York.
[7] McKeague, I.W. and Zhang, M., 1994. Identification of nonlinear time series from first order cumulative characteristics. *Ann. Statist.*, **22**(1), 495-514.
[8] Roussas, G.G., 1972. Contiguity of probability measures: some applications in statistics. *Cambridge University Press, London.*
[9] Taniguchi, M. and Kakizawa, Y., 2000. **Asymptotic theory of statistical inference for time series.** Springer-Verlag, New York.
[10] Tong, H., 1998. Nonlinear time series. The Clarendon Press Oxford University Press, New York, 1990. Cambridge University Press, Cambridge.
[11] Wald, A., 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, **54**, 426-482.