

A Nonparametric Goodness-of-fit Test for a Class of Parametric Autoregressive Models

Joseph Ngatchou Wandji

► To cite this version:

Joseph Ngatchou Wandji. A Nonparametric Goodness-of-fit Test for a Class of Parametric Autoregressive Models. RR-3065, INRIA. 1996. inria-00073627

HAL Id: inria-00073627

<https://hal.inria.fr/inria-00073627>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***A Nonparametric Goodness-of-fit Test for a
Class of Parametric Autoregressive Models***

Joseph NGATCHOU WANDJI

N° 3065

décembre 1996

_____ THÈME 4 _____



*R*apport
de recherche

A Nonparametric Goodness-of-fit Test for a Class of Parametric Autoregressive Models

Joseph NGATCHOU WANDJI

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet IS2

Rapport de recherche n° 3065 — décembre 1996 — 22 pages

Abstract: We derive a nonparametric test for discriminating between generalized autoregressive models. This test is based on a suitably normalized sum of residuals. The null distribution and the power of the test under both fixed and a sequence of local alternatives are studied under mild stationarity and α -mixing conditions. This procedure can be applied to testing linear models against nonlinear models or certain nonlinear models against others. Numerical simulations show that the proposed test is powerful against most of the alternatives considered.

Key-words: Autoregressive models, contiguity, goodness-of-fit tests, mixing, nonlinear models, nonparametric methods.

(Résumé : tsvp)

Un test d'adéquation pour une classe paramétrique de processus autorégressifs

Résumé : Nous proposons un test non paramétrique pour discriminer entre modèles autorégressifs. Ce test est basé sur une somme normalisée de résidus. La loi de la statistique de test sous l'hypothèse nulle, la puissance et la puissance locale sont étudiées sous des conditions faibles de stationnarité et de mélangeance forte. Cette procédure peut être utilisée pour tester des modèles linéaires contre des modèles non linéaires ou certains modèles non linéaires contre d'autres. Des simulations numériques montrent un bon comportement de la puissance sous la plupart des alternatives considérées.

Mots-clé : Contiguïté, mélangeance, méthodes non paramétriques, modèles autorégressifs, modèles non linéaires, tests d'adéquation.

1. Introduction – In time series analysis, it is often hard to select an appropriate class of models. This paper is concerned with testing the form of the second term (“mixed term”) in the right-hand side of the following equation

$$X_{i+1} = \sum_{k=0}^p a_k X_{i-k} + V(X_i, \dots, X_{i-q}, \varepsilon_i, \dots, \varepsilon_{i-Q}) + \sigma(X_i, \dots, X_{i-L}) \varepsilon_{i+1}, \quad (1)$$

where $\{\varepsilon_i : i \in \mathbb{Z}\}$ is a sequence of independent and identically distributed random variables (iid rv’s) such that $E(\varepsilon_1) = 0$, $E(\varepsilon_1^2) = 1$, ε_{i+1} is independent of the σ -field $\sigma(X_j : j \leq i)$, for each $i \in \mathbb{Z}$, and $p, q, Q, L \geq 0$ are given integers. The function V has a given parametric function. The parameters $a_i, i = 0, \dots, p$, the distribution of ε_1 and the function $\sigma(\cdot)$ are not necessarily known.

Tests based on likelihood functions, such as Lagrange multiplier (LM) tests, as well as rank-based tests have been applied to testing ARMA models against bilinear, exponential or threshold models (see, e.g., Saikkonen and Luukonen, 1988; Luukonen, Saikkonen and Terasvirta, 1988; Benghabrit and Hallin, 1992, 1996a and b).

Other tests, essentially of nonparametric flavor, have been derived for testing linear models against unspecified alternatives. But in the case of rejection of the linear hypothesis, it is difficult to choose the correct type of nonlinearity. Some of these tests are based on the bispectrum (Subba Rao and Gabr, 1980 and Hinnich 1982). Others are based on normalized sums of quadratic deviations between a parametric and a nonparametric estimate of the conditional mean or variance function (Poggi and Portier, 1995 and Hjellvik and Tjøstheim, 1995). The tests proposed by Keenan (1985) and Tsay (1986) are based on auxiliary regressions.

A number of papers on tests based on weighted empirical processes of residuals have been proposed for testing the conditional mean or variance function (e.g., An and Cheng, 1991; Diebolt, 1990 and Diebolt and Laïb, 1995), or for testing the mixed term in an autoregressive equation (Diebolt and Ngatchou Wandji, 1996a,b). In all these papers the parameters of the null model are supposed to be known. In An and Cheng (1991), the case where the parameter of the null model is not known is also treated. The Kolmogorov–Smirnov type statistic that they propose in this case is based on subsample sums. This modification of the test statistic allows the authors to obtain its limiting distribution under the null hypothesis, but it involves a significant loss of information and power.

McKeague and Zhang (1994) have introduced a procedure for testing whether the conditional mean or variance function has a specific parametric form. Their test is based on a comparison between a nonparametric and a parametric estimator of the cumulated conditional mean or variance function. But this test, like that of An and Cheng (1991), cannot be applied to testing the existence of certain bilinear or EXPARMA effects, for instance.

Under stationarity and α -mixing conditions, Diebolt and Ngatchou Wandji (1996b) have derived a Kolmogorov–Smirnov type test for testing the presence of a term of the form $V(X_i, \varepsilon_i)$ in the stochastic difference equation

$$X_{i+1} = T(X_i) + V(X_i, \varepsilon_i) + \sigma(X_i) \varepsilon_{i+1}, \quad (2)$$

where $\{\varepsilon_i : i \geq 1\}$ is a sequence of standardized iid rv's such that ε_{i+1} is independent of $\sigma(X_j : j \leq i)$ for all i , the function $T(\cdot)$ is assumed known and can be nonlinear, the function $\sigma(\cdot)$ is not necessarily known and the function $V(x, y)$ has a given parametric form. This test is an extension of a test of linearity derived by the same authors in an earlier paper (Diebolt and Ngatchou Wandji, 1996a), where the particular bilinear case $T(x) = ax$, $V(x, y) = bxy$, $\sigma(x) = \sigma \neq 0$ is considered.

The purpose of the present paper is to extend the test of Diebolt and Ngatchou Wandji (1996b) to the case where the function $T(\cdot)$ is unknown. If we replace $T(\cdot)$ in the test statistic by a consistent estimator, the limiting distribution of the test statistic under the null hypothesis is not analytically tractable in general. In order to overcome such a difficulty, we propose a simplified and more flexible version of the above test in a more general setting than Diebolt and Ngatchou Wandji (1996a).

We consider equation (1) and assume that there exists a $\gamma > 0$ such that:

(\mathcal{A}_1) The sequence $\{(X_i, \dots, X_{i-\max(L,q)}, \varepsilon_i, \dots, \varepsilon_{i-Q})\}$ is stationary, ergodic and α -mixing (see, e.g., Doukhan, 1994, for the definition of α -mixing) with α_k satisfying

$$\sum_{k \geq 0} (k+1) \alpha_k^{\gamma/(2+\gamma)} < \infty.$$

(\mathcal{A}_2) The moment $E(|X_1|^{2+\gamma}) < \infty$.

We wish to test the null hypothesis $\mathcal{H}_0(V = V_0)$ against the alternative $\mathcal{H}_1(V = V_0 + V_*)$, where the functions V_0 and V_* have given parametric forms. We also assume that

(\mathcal{A}_3) Under \mathcal{H}_0 , $E(X_1) = 0$.

The test statistic that we propose is $T_n = \left(n^{-1/2} \sum_{i=1}^n (X_{i+1} - \sum_{k=0}^p \hat{a}_k X_{i-k}) \right)^2 / \hat{\rho}_n^2$, where \hat{a}_k , $k = 0, \dots, p$ and $\hat{\rho}_n^2$ are respectively consistent estimators under \mathcal{H}_0 of a_k , $k = 0, \dots, p$ and $\rho_{\mathcal{H}_0}^2 = E\left(V_0^2(X_i, \dots, X_{i-Q}, \varepsilon_i, \dots, \varepsilon_{i-Q}) + \sigma^2(X_i, \dots, X_{i-L})\right)$.

The assumptions (\mathcal{A}_1) and (\mathcal{A}_2) are satisfied under certain conditions by a large number of models contained in Examples 1–3. The assumption (\mathcal{A}_3) insures the convergence of T_n under the null hypothesis (Theorem 1). It holds when $E(V_0(X_i, \dots, X_{i-Q}, \varepsilon_i, \dots, \varepsilon_{i-Q})) = 0$ and $\sum_{i=0}^p a_i \neq 1$. If (\mathcal{A}_3) does not hold, then T_n either diverges or converges to a rv whose distribution is untractable in general. The consistency of the test requires that under \mathcal{H}_1 , $E(X_1) \neq 0$. Nonetheless when this condition fails, an explicit expression of the asymptotic power can be obtained (Theorem 2), but in this case the consistency of the test is not ensured. See Examples 1–4 in Section 2.

The paper is organized as follows. Section 2 is devoted to the study of the asymptotic null distribution of our test. In Section 3, we study the power of the test under fixed alternatives. We show that under mild conditions, this power converges to one. In Section

4, we study the asymptotic power of the test against a sequence of local alternatives of the form $\mathcal{H}_1^n(V = V_0 + n^{-1/2}\delta)$, where n is the sample size. We show that the sequences $\{\mathcal{H}_1^n : n \geq 1\}$ and $\{\mathcal{H}_0^n = \mathcal{H}_0 : n \geq 1\}$ are contiguous. From this result, we derive an explicit expression for the asymptotic local power. In Section **5**, under first-order diagonal bilinear alternatives, this local power is compared to that of the Lagrange multiplier test (Luukonen and Saikkonen 1988 and Benghabrit and Hallin 1996a), and to the signed-rank and rank tests of Benghabrit and Hallin (1996b). In Section **6**, we present Monte Carlo simulation experiments. The last section we concludes our work.

2. Asymptotic null distribution – For simplicity, all the results of this paper are obtained with $p = L$. For $p \neq L$ similar results can be stated and proved.

Let $\mathbf{x} = (x_0, \dots, x_p)$ and $\mathbf{y} = (y_0, \dots, y_p)$. Denote by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=0}^p x_j y_j$ the standard Euclidean inner product of \mathbb{R}^{p+1} and by $\hat{\theta}_n = (\hat{a}_0, \dots, \hat{a}_p)$ a strongly or weakly consistent estimator of $\theta = (a_0, \dots, a_p)$ under \mathcal{H}_0 . Let $\mathbf{X}_i = (X_i, \dots, X_{i-p})$, $\mathbf{Z}_i = (X_i, \dots, X_{i-q}, \varepsilon_i, \dots, \varepsilon_{i-Q})$ and

$$S_n = n^{-1/2} \sum_{i=1}^n \left(X_{i+1} - \langle \hat{\theta}_n, \mathbf{X}_i \rangle \right) = n^{-1/2} \sum_{i=1}^n \left(X_{i+1} - \sum_{j=0}^p \hat{a}_j X_{i-j} \right). \quad (3)$$

Before defining our test, we make the following additional technical assumptions and give examples.

(A₄) Under \mathcal{H}_0 , $E \left\{ \left[V_0(Z_i) + \sigma(X_i)\varepsilon_{i+1} \right] \left[V_0(Z_j) + \sigma(X_j)\varepsilon_{j+1} \right] \right\} = 0$ for all $i \neq j$.

(A₅) Under \mathcal{H}_0 , $E(\sigma^2(X_1)) \neq 0$.

Let $q \geq Q$. Possible null nonlinear models for our test have the form

$$\begin{aligned} X_{i+1} &= \sum_{j=0}^p a_j X_{i-j} + \sum_{j=1}^Q \beta_j(X_{i-j}, \dots, X_{i-(q-Q)-j}, \varepsilon_{i-j}, \dots, \varepsilon_{i-(q-Q)-j}) \varepsilon_{i-j+1} \\ &\quad + \sigma(X_i, \dots, X_{i-q}) \varepsilon_{i+1}, \end{aligned} \quad (4)$$

with $\sum_{j=0}^p a_j \neq 1$ and the β_j real-valued functions. For such models, $E(X_1) = 0$.

Example 1 – If we take $V(Z_i) = 0$ and $\sigma(X_i, \dots, X_{i-L}) = \sigma$ in (1), it then reduces to a linear $AR(p)$ model. Let $P(z) = 1 - a_0 z - a_1 z^2 \dots - a_p z^{p+1}$. If the roots of P have modulus larger than one, then under some additional technical assumptions (e.g., Mokkadem, 1990), (A₁)–(A₃) hold. Thus our test can be used for testing a linear $AR(p)$ model against various nonlinear models.

Example 2 – Let $q \geq Q$. If we take

$$V(Z_i) = \sum_{j=0}^q b_j \varepsilon_{i-j} + \sum_{t=0}^q \sum_{\ell=-1}^{q-Q} b_{t\ell} X_{i-\ell-t} \varepsilon_{i-t} \text{ and } \sigma(X_i) = 1$$

in equation (1), it then reduces to a bilinear model. Under some technical assumptions, it can be proved, using Theorem 4.2 of Pham (1985) and Theorem 4.3 of Mokkadem (1990), that for any integers t and ℓ the sequence $\left\{ (X_i, \dots, X_{i-t}, \varepsilon_i, \dots, \varepsilon_{i-\ell}) \right\}_{i \in \mathcal{Z}}$ is geometrically ergodic and α -mixing. Thus, the assumption (\mathcal{A}_1) holds. Also, by Mokkadem (1990) the assumption (\mathcal{A}_2) holds whenever $|\varepsilon_1|$ has a moment of order $2(2 + \gamma)$. If we take, e.g., $a = a_0$, $a_j = 0$ for $j = 1, \dots, p$, $|a| < 1$ and $V_0(\mathbf{Z}_i) = bX_{i-1}\varepsilon_i$ with $a^2 + b^2 < 1$, we have $E\{V_0(\mathbf{Z}_i)\} = 0$ for all a and b , and (\mathcal{A}_3) holds. Note that $|a| < 1$ and $a^2 + b^2 < 1$ are sufficient conditions for the existence of a stationary solution in this case (e.g., Guégan and Ngatchou Wandji, 1996). The test we propose can be used for testing, e.g., $V = V_0$ against $V = V_0 + V_*$, with V_0 as above and $V_*(\mathbf{Z}_i) = \tau X_i \varepsilon_i$ or any other bilinear term such that $E(X_1) \neq 0$ under \mathcal{H}_1 .

Example 3 – If the left-hand side of (1) is replaced by $h(X_i, \dots, X_{i-p}, \varepsilon_i; \theta) + \varepsilon_{i+1}$, where $h : \mathbb{R}^{p+1} \times \Theta \rightarrow \mathbb{R}$ and $\theta \in \Theta \subset \mathbb{R}^d$, $d \geq 1$, we obtain a general model which contains many common models, such as EXPARMA and threshold models. Under some technical assumptions, Theorems 2.2 and 2.3 of Hili (1995) imply that the sequence $\{(X_i, \dots, X_{i-p}, \varepsilon_i)\}$ is geometrically ergodic and α -mixing. Moreover, if there exists $s > 0$ such that $E(|\varepsilon_1|^s) < \infty$, then $E(|X_1|^{s'}) < \infty$ for all $s' \leq s$ in the stationary regime. It then follows that under some additional assumptions the requirements (\mathcal{A}_1) and (\mathcal{A}_2) are satisfied. For instance, if

$$h(X_i, \dots, X_{i-p}, \varepsilon_i; \theta) = aX_i + \tau I_{(-\infty, r]}(X_{i-1})\varepsilon_i + bX_i, \varepsilon_i,$$

where r is a known threshold then taking $V_0(\mathbf{Z}_i) = \tau I_{(-\infty, r]}(X_{i-1})\varepsilon_i$, we have $E\{V_0(\mathbf{Z}_i)\} = 0$, and (\mathcal{A}_3) is satisfied when $|a| \neq 1$. Therefore, our test can be applied to testing, e.g., $V = V_0$ against $V = V_0 + V_*$, with V_0 as above and $V_*(\mathbf{Z}_i) = bX_i \varepsilon_i$ or $bX_i \exp(-\lambda^2 X_i)$ or any other nonlinear term such that $E(X_1) \neq 0$ under \mathcal{H}_1 .

Example 4 – If we take $V(\mathbf{Z}_i) = \tau \varepsilon_{i-1} \varepsilon_i + b \varepsilon_i^2$ and $\sigma(X_i) = \sigma$ in (1), it then reduces to an extension of a nonlinear moving average model (Robinson, 1977). When the conditions (\mathcal{A}_1) – (\mathcal{A}_5) are in force, our procedure can be used to testing e.g., $V = V_0$ against $V = V_0 + V_*$, where $V_0(\mathbf{Z}_i) = \tau \varepsilon_{i-1} \varepsilon_i$ and $V_*(\mathbf{Z}_i) = b \varepsilon_i^2$.

Note that most of the above nonlinear models cannot be tested using McKeague and Zhang's (1994) or An and Chen's (1991) methods. See additional examples in Section 6.

Consider model (2). When $T(x) = ax$ and $\sigma(x) = \sigma$, the unknown parameter a can be estimated under the null hypothesis by the least squares estimator \hat{a}_n . If we substitute \hat{a}_n for a in the test statistic M_n defined in Diebolt and Ngatchou (1996a), we then obtain $M_n = n^{-1/2} \sup_{t \in J} \left| \sum_{i=1}^n (X_{i+1} - \hat{a}_n X_i) I_{[c_1, t]}(X_i) \right|$, with J is a compact interval $[c_1, c_2]$ and $I_K(x) = 1$ if $x \in K$, 0 if $x \in K^c$. Under the null hypothesis and under stationarity and

α -mixing, it can be proved that M_n converges in distribution to

$$\sup_{t \in J} |(W \circ \Lambda)(t) + \xi \Theta(t)|,$$

where $\Lambda(\cdot)$ and $\Theta(\cdot)$ are real-valued functions, $W(s)$ is a standard Wiener process and ξ is a centered Gaussian rv with variance $\sigma^2/(1-a^2)$, which is not independent of W . But the distribution of such a supremum is not analytically known in general and the available asymptotic expressions for the tail as $u \rightarrow \infty$ of $P\{\sup_{t \in J} |(W \circ \Lambda)(t) + \xi \Theta(t)| > u\}$ are not sufficiently precise for moderate values of u to allow us to tabulate this distribution.

If we replace $I_{[c_1, t]}(X_i)$ by 1 in the expression of M_n , it then converges in distribution to $|\mathcal{N}(0, \sigma^2)|$ under the null hypothesis and under stationarity and α -mixing. Let $\rho_{\mathcal{H}_0}^2 = E\{V_0^2(Z_1) + \sigma^2(X_1)\}$. Our test statistic is defined by

$$T_n = (S_n / \hat{\rho}_n)^2, \quad (5)$$

where $\hat{\rho}_n$ a consistent estimator of $\rho_{\mathcal{H}_0}$ under \mathcal{H}_0 .

Theorem 1 *Under \mathcal{H}_0 , if the assumptions (\mathcal{A}_1) - (\mathcal{A}_5) hold, then S_n converges in distribution to a Gaussian $\mathcal{N}(0, \rho_{\mathcal{H}_0}^2)$ random variable.*

Proof of Theorem 1 – We first show that $E\{V_0^2(Z_1) + \sigma^2(X_1)\} < \infty$. Under \mathcal{H}_0 , $V_0(Z_i) + \sigma(X_i)\varepsilon_{i+1} = X_{i+1} - \sum_{j=0}^p a_j X_{i-j}$. Raising each member of this equation to the square and taking their expectation, it follows from the C_r -inequality (see, e.g., Shorack and Wellner, 1986) that

$$E\{V_0^2(Z_1) + \sigma^2(X_1)\} \leq C_2 \left(1 + \sum_{i=0}^p |a_i|^2\right) E(|X_1|^2) < \infty.$$

Now define $\alpha(t) = \alpha_{[t]}$, where $[t]$ denotes the integer part of t , and denote by Q_ξ the tail quantile function of a rv ξ . It can be proved (see Proposition 3 in Appendix 1) that if $\{\xi_i : i \in \mathbb{Z}\}$ is a stationary and strongly mixing sequence of centered real-valued rv's such that $\sum_{k \geq 0} (k+1) \alpha_k^{(r-2)/r} < \infty$ and $E(|\xi_1|^r) < \infty$ for some $r > 2$, then

$$\int_0^1 \alpha(u) Q_{\xi_0}^2(u) du < \infty.$$

The above moment condition implies that the sequence $\{\xi_i : i \in \mathbb{Z}\}$ satisfies the assumptions of Theorem 1 of Rio (1995) (see Appendix 1). Now, the sum S_n defined by (2) can be written in the form $S_n = A_n - B_n$, where

$$A_n = n^{-1/2} \sum_{i=1}^n \left(X_{i+1} - \langle \theta, X_i \rangle \right) \quad \text{and} \quad B_n = \left\langle \hat{\theta}_n - \theta, n^{-1/2} \sum_{i=1}^n X_i \right\rangle. \quad (6)$$

Any j^{th} -coordinate $(n^{-1/2} \sum_{i=1}^n X_i)_j = n^{-1/2} \sum_{i=1}^n X_{i-j}$ of $n^{-1/2} \sum_{i=1}^n X_i$ is a sum of α -mixing rv's. Under the assumptions (\mathcal{A}_1) - (\mathcal{A}_3) , Theorem 1 of Rio (1995) implies that $n^{-1/2} \sum_{i=1}^n X_{i-j}$ converges in distribution to $\mathcal{N}(0, \rho_B^2)$, where ρ_B^2 is the sum of the absolutely convergent series $\sum_{i \in \mathbf{Z}} \text{cov}(X_0, X_i)$. Since, any j^{th} -coordinate $(\hat{a}_j - a_j)$ of $\hat{\theta}_n - \theta$ converges almost surely (a.s.) or in probability (P) to 0, it results that B_n also converges a.s. or P to 0. Under \mathcal{H}_0 ,

$$A_n = n^{-1/2} \sum_{i=1}^n \left\{ V_0(Z_i) + \sigma(X_i) \varepsilon_{i+1} \right\}. \quad (7)$$

Since $V_0(Z_i) + \sigma(X_i) \varepsilon_{i+1} = X_{i+1} - \langle \theta, X_i \rangle$ under \mathcal{H}_0 , we have $E(A_n) = 0$ by (\mathcal{A}_3) . By the C_r -inequality we have

$$E \left\{ \left| V_0(Z_i) + \sigma(X_i) \varepsilon_{i+1} \right|^{2+\gamma} \right\} \leq C_{2+\gamma} \left(1 + \sum_{i=0}^p |a_i|^{2+\gamma} \right) E(|X_1|^{2+\gamma}) < \infty, \quad (8)$$

in view of (\mathcal{A}_2) . It then follows from (\mathcal{A}_1) that A_n converges in distribution to $\mathcal{N}(0, \rho_A^2)$, where ρ_A^2 is the sum of the absolutely convergent series

$$\sum_{i \in \mathbf{Z}} \text{cov} \left(V_0(Z_0) + \sigma(X_0) \varepsilon_1, V_0(Z_i) + \sigma(X_i) \varepsilon_{i+1} \right). \quad (9)$$

So, S_n converges in distribution to $\mathcal{N}(0, \rho_{\mathcal{H}_0}^2)$, with $\rho_{\mathcal{H}_0}^2 = \rho_A^2$. From Theorem 1 of Rio (1995), $n^{-1} \text{var}(A_n)$ converges to ρ_A^2 . Since under (\mathcal{A}_4) the rv's $\xi_i = V_0(Z_i) + \sigma(X_i) \varepsilon_{i+1}$ are centered and ε_{i+1} is independent of $\sigma(X_j : j \leq i)$, we have $n^{-1} \text{var}(A_n) = n^{-1} E(A_n^2) = E \left\{ V_0^2(Z_1) + \sigma^2(X_1) \right\} = \rho_A^2 = \rho_{\mathcal{H}_0}^2$. \diamond

Remark 1 – When the assumptions (\mathcal{A}_4) - (\mathcal{A}_5) are not satisfied, the above theorem still holds with $\rho_{\mathcal{H}_0}^2$ defined by (9). However, in this case $\rho_{\mathcal{H}_0}^2$ may be equal to 0, or may be difficult to estimate.

Corollary 1 *Under \mathcal{H}_0 , if the assumptions (\mathcal{A}_1) - (\mathcal{A}_5) hold, then T_n converges in distribution to a $\chi^2(1)$ random variable.*

The following simple proposition gives sufficient conditions under which estimates of $\rho_{\mathcal{H}_0}^2$ can be obtained.

Proposition 1 *Let $U_n = n^{-1} \sum_{i=1}^n \left\{ X_{i+1} - \langle \hat{\theta}_n, X_i \rangle \right\}^2$. Under \mathcal{H}_0 , if $\hat{\theta}_n \xrightarrow{a.s.} \theta$ and the assumption (\mathcal{A}_1) holds, then U_n converges a.s. to $\rho_{\mathcal{H}_0}^2 = E \left\{ V_0^2(Z_1) + \sigma^2(X_1) \right\}$.*

Remark 2 – Let $V_n = n^{-1} \sum_{i=1}^n \left(X_{i+1}^2 - \langle \hat{\theta}_n, X_i \rangle^2 \right)$. Under the assumptions of Proposition 1, if under \mathcal{H}_0 the technical condition $E \left\{ \langle \theta, X_1 \rangle V_0(Z_1) \right\} = 0$ holds, it can be proved that

V_n converges a.s. to $\rho_{\mathcal{H}_0}^2$. This assumption holds, e.g. for the models described by (4).

3. Power of the test against fixed alternatives – This section is devoted to the study of the power of our test against fixed alternatives $\mathcal{H}_1(V = V_0 + V_*)$. Two cases arise: $E(X_1) = 0$ and $E(X_1) \neq 0$. The first situation can happen for, for instance, when $V_0(Z_1)$ and $V_*(Z_1)$ are subdiagonal bilinear terms in a bilinear equation. For example, put $\langle \theta, X_i \rangle = aX_i$ with $|a| < 1$, $V(Z_i) = \tau X_{i-2}\varepsilon_i + bX_{i-1}\varepsilon_i$ and $\sigma(X_i) = \sigma$ in (1). If we take $V_0(Z_i) = \tau X_{i-2}\varepsilon_i$ and $V_*(Z_i) = bX_{i-1}\varepsilon_i$, we then have $E\{V_0(Z_1) + V_*(Z_1)\} = 0$ which implies that $E(X_1) = 0$. The other situation can happen, for instance, when at least one of $V_0(Z_1)$ and $V_*(Z_1)$ is a diagonal bilinear term in a bilinear equation. For example, put $\langle \theta, X_i \rangle = aX_i$ with $|a| < 1$, $V(Z_i) = bX_i\varepsilon_i$ with b such that $a^2 + b^2 < 1$, and $\sigma(X_i) = \sigma$ in (1). If we take $V_0(Z_i) = 0$ and $V_*(Z_i) = bX_i\varepsilon_i$, one can prove that $E\{V_0(Z_1) + V_*(Z_1)\} \neq 0$, i.e. $E(X_1) \neq 0$. More generally, if $E(V_0(Z_1) + V_*(Z_1)) = 0$ and $\sum_{i=0}^p a_i \neq 1$, then $E(X_1) = 0$. Also, if $E(V_0(Z_1) + V_*(Z_1)) \neq 0$ and $\sum_{i=0}^p a_i \neq 1$, then $E(X_1) \neq 0$. We have the following theorem:

Theorem 2 Under \mathcal{H}_1 , assume that (\mathcal{A}_1) and (\mathcal{A}_2) hold and that $\hat{\theta}_n \xrightarrow{a.s.} \theta_{\mathcal{H}_1}$ for some $\theta_{\mathcal{H}_1} \in \mathbb{R}^{p+1}$ and $\hat{\rho}_n^2 \xrightarrow{a.s.} \rho_{\mathcal{H}_1}^2$ for some positive $\rho_{\mathcal{H}_1}^2$ as $n \rightarrow \infty$.

(i) If $E(X_1) \neq 0$, then at the asymptotic level of significance α

$$P\{T_n > \chi_\alpha\} \longrightarrow 1 \text{ as } n \rightarrow \infty.$$

(ii) If $E(X_1) = 0$, then at the asymptotic level of significance α

$$P\{T_n > \chi_\alpha\} \longrightarrow 1 - \Phi(\sqrt{\chi_\alpha} \rho / \rho_{\mathcal{H}_1}) + \Phi(-\sqrt{\chi_\alpha} \rho / \rho_{\mathcal{H}_1}) \text{ as } n \rightarrow \infty,$$

where ρ^2 is the sum of the absolutely convergent series

$$\sum_{i \in \mathcal{Z}} \text{cov}\left\{\left(X_1 - \langle \theta_{\mathcal{H}_1}, X_0 \rangle\right), \left(X_{i+1} - \langle \theta_{\mathcal{H}_1}, X_i \rangle\right)\right\},$$

Φ is the standard normal repartition function and χ_α is the $(1 - \alpha)$ -quantile of the $\chi^2(1)$ distribution.

We now give an example for which the assumptions $\hat{\theta}_n \xrightarrow{a.s.} \theta_{\mathcal{H}_1}$ and $\hat{\rho}_n^2 \xrightarrow{a.s.} \rho_{\mathcal{H}_1}^2$ of Theorem 2 hold. Consider the second order subdiagonal bilinear model, i.e. the model (1) where $\langle \theta, X_j \rangle = aX_j$ with $|a| < 1$, $V_0(Z_j) = \tau X_{j-1}\varepsilon_j$, $V_*(Z_j) = bX_j\varepsilon_j$ and $\sigma(X_j) = \sigma$. Consider the moment estimator of a under \mathcal{H}_0 , $\hat{a}_n = \sum_{j=1}^n X_{j+1}X_j / \sum_{j=1}^n X_j^2$. Under \mathcal{H}_1 and under ergodicity and moment conditions, it can be proved that $\hat{a}_n \xrightarrow{a.s.} a + \varepsilon = E(X_1X_0)/E(X_0^2) \leq 1$, with $\varepsilon = E(X_1^2\varepsilon_1 + X_1X_0\varepsilon_1)/E(X_1^2)$. It follows that $\hat{\rho}_n^2 \xrightarrow{a.s.} \{1 - (a + \varepsilon)^2\}E(X_1^2)$ as $n \rightarrow \infty$.

One can prove that ϵ is finite when $E(|X_1|^4)$ is finite.

Remark 3 – (i) Theorem 2 still holds if $\hat{\theta}_n \xrightarrow{P} \theta_{\mathcal{H}_1}$ for some $\theta_{\mathcal{H}_1} \in \mathbb{R}^{p+1}$ and $\hat{\rho}_n^2 \xrightarrow{P} \rho_{\mathcal{H}_1}^2$ for some positive $\rho_{\mathcal{H}_1}^2$ as $n \rightarrow \infty$.

(ii) If we only assume that under \mathcal{H}_1 , the $\hat{\theta}_n$'s remains in a compact subset of \mathbb{R}^{p+1} , then we can prove that if $E(X_1) \neq 0$ under \mathcal{H}_1 , then $P\{T_{\phi(n)} > \chi_\alpha\} \rightarrow 1$ as $n \rightarrow \infty$ for certain subsequences $\{T_{\phi(n)}\}$ of $\{T_n\}$.

Proof of Theorem 2 – For (i), denote $v = E\left(X_1 - \langle \theta_{\mathcal{H}_1}, X_0 \rangle\right) \in \mathbb{R}$ and $K = E\left(X_0\right) \in \mathbb{R}^{p+1}$.

Under \mathcal{H}_1 , we can write $S_n = C_n + D_n + n^{1/2} \left(v + \langle \theta_{\mathcal{H}_1}, K \rangle\right)$, where

$$C_n = n^{-1/2} \sum_{i=1}^n \left(X_{i+1} - \langle \theta_{\mathcal{H}_1}, X_i \rangle - v\right), \quad (10)$$

$$D_n = \left\langle \omega_n, n^{-1/2} \sum_{i=1}^n \left(X_i - K\right) \right\rangle, \quad (11)$$

with $\omega_n \in \mathbb{R}^{p+1} \rightarrow 0$ a.s. as $n \rightarrow \infty$. Under the assumptions (\mathcal{A}_1) – (\mathcal{A}_2) , it can be shown as for B_n that the sequence D_n converges a.s. to 0 as $n \rightarrow \infty$. Using the same arguments as for A_n , it can be proved that the sequence C_n converges in distribution to $\mathcal{N}(0, \rho_C^2)$, where ρ_C^2 is the sum of the absolutely convergent series

$$\sum_{i \in \mathcal{I}} \text{cov} \left\{ \left(X_1 - \langle \theta_{\mathcal{H}_1}, X_0 \rangle - v\right), \left(X_{i+1} - \langle \theta_{\mathcal{H}_1}, X_i \rangle - v\right) \right\}.$$

It then follows that $S_n/\hat{\rho}_n$ is asymptotic to $\mathcal{N}\left(0, \rho_C^2/\rho_{\mathcal{H}_1}^2\right) + n^{1/2}\rho_{\mathcal{H}_1}^{-1} \left(v + \langle \theta_{\mathcal{H}_1}, K \rangle\right)$. Since $v + \langle \theta_{\mathcal{H}_1}, K \rangle = E(X_1) \neq 0$, we have

$$P\{T_n > \chi_\alpha\} \sim P\left\{\left[\mathcal{N}\left(0, \rho_C^2/\rho_{\mathcal{H}_1}^2\right) + n^{1/2}\rho_{\mathcal{H}_1}^{-1}E(X_1)\right]^2 > \chi_\alpha\right\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

For (ii), we first remark that under \mathcal{H}_1 we have $S_n = E_n + F_n$, where

$$E_n = n^{-1/2} \sum_{i=1}^n \left(X_{i+1} - \langle \theta_{\mathcal{H}_1}, X_i \rangle\right) \quad \text{and} \quad F_n = \left\langle \omega_n, n^{-1/2} \sum_{i=1}^n X_i \right\rangle, \quad (12)$$

with $\omega_n \in \mathbb{R}^{p+1} \rightarrow 0$ a.s. as $n \rightarrow \infty$. In the same manner as for B_n in the proof of Theorem 1, it can be shown that F_n converges a.s. to 0 as $n \rightarrow \infty$. The sequence E_n is a normalized sum of centered α -mixing rv's satisfying (\mathcal{A}_1) . One can show, using Theorem 1 of Rio (1995), that E_n converges in distribution to $\mathcal{N}(0, \rho_E^2)$, where $\rho_E^2 = \rho^2$ is the sum of the absolutely convergent series

$$\sum_{i \in \mathcal{I}} \text{cov} \left\{ \left(X_1 - \langle \theta_{\mathcal{H}_1}, X_0 \rangle\right), \left(X_{i+1} - \langle \theta_{\mathcal{H}_1}, X_i \rangle\right) \right\}.$$

Since we have assumed that under \mathcal{H}_1 , $\hat{\rho}_n^2$ converges a.s. to $\rho_{\mathcal{H}_1}^2$, it results that $S_n/\hat{\rho}_n = E_n/\hat{\rho}_n + F_n/\hat{\rho}_n$ converges in distribution to $\mathcal{N}\left(0, \rho^2/\rho_{\mathcal{H}_1}^2\right)$, and

$$P\{T_n > \chi_\alpha\} \longrightarrow 1 + \Phi(-\sqrt{\chi_\alpha} \rho/\rho_{\mathcal{H}_1}) - \Phi(\sqrt{\chi_\alpha} \rho/\rho_{\mathcal{H}_1}) \text{ as } n \rightarrow \infty. \diamond$$

Remark 4 – Assertion (ii) of Theorem 2 implies that one can obtain, under some technical assumptions, expressions for the theoretical asymptotic power of the test even in the case $E(X_1) = 0$. In this case, the power can be close to one if $\rho/\rho_{\mathcal{H}_1}$ is small enough. However, $\rho_{\mathcal{H}_1}$ and ρ may be very difficult to calculate.

4. Power of the test against local alternatives – In this section, we study the power of the test against local alternatives of the form $\mathcal{H}_1^n\left(V = V_0 + n^{-1/2}\delta\right)$, where $E[\delta(Z_1)] \neq 0$. We restrict ourselves to models which are invertible and for which ε_t can be written in terms of $X_t, X_{t-1}, \dots, X_0, \dots, X_{-m}, \varepsilon_0, \dots, \varepsilon_{-m}$, where $m = \max(p, q, Q)$. This concerns many time series models, e.g. linear and some nonlinear AR models such as EXPAR and TAR models (e.g., Hili, 1995), some bilinear models (e.g., Pham, 1981 and Guégan and Ngatchou Wandji, 1996). We also assume that the ε_i 's are iid with density f satisfying the following assumptions:

$$(\mathcal{B}_1) \quad f(x) > 0 \text{ for all } x \in \mathbb{R}, \int x f(x) dx = E(\varepsilon_1) = 0 \text{ and } \int x^2 f(x) dx = \text{var}(\varepsilon_1) = 1.$$

$$(\mathcal{B}_2) \quad f \text{ is differentiable with derivative } f'; I(f) = \int \phi_f^2(x) f(x) dx < \infty \text{ and } \int |x \phi_f(x)| f(x) dx < \infty, \text{ where } \phi_f = f'/f.$$

$$(\mathcal{B}_3) \quad \phi_f \text{ is differentiable with derivative } \phi'_f, \text{ and } \phi'_f \text{ is } C_\phi\text{-Lipschitzian of order one for some constant } C_\phi > 0.$$

$$(\mathcal{B}_4) \quad \lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0 \text{ and } \lim_{x \rightarrow +\infty} f'(x) = \lim_{x \rightarrow -\infty} f'(x) = 0.$$

$$(\mathcal{B}_5) \quad \int |\phi_f(x)|^{2+\gamma} f(x) dx < \infty \text{ where } \gamma > 0 \text{ is as in the Introduction.}$$

Remark 5 – Condition (\mathcal{B}_4) implies that $\int \phi_f(x) f(x) dx = 0$ and $\int \phi_f^2(x) f(x) dx = - \int \phi'_f(x) f(x) dx$.

(\mathcal{B}_1) – (\mathcal{B}_5) are satisfied, e.g., by normal densities. Such assumptions have been made in Guégan and Ngatchou Wandji (1996) and Benghabrit and Hallin (1992 and 1996a,b).

In the remainder of this paper, we need the following additional technical assumptions:

$$(\mathcal{C}_1) \quad \text{There exists } \zeta > 0 \text{ such that } |\sigma(x_0, \dots, x_p)| \geq \zeta \text{ for all } (x_0, \dots, x_p) \in \mathbb{R}^{p+1}.$$

$$(\mathcal{C}_2) \quad \text{Under } \mathcal{H}_0, E(|\delta(Z_1)|^3) < \infty.$$

(C₃) Under \mathcal{H}_0 , $E\left\{Y_s \phi_f(\varepsilon_{s+1}) V_0(Z_t)\right\} = E\left\{Y_s^2 \phi'_f(\varepsilon_{s+1}) \left[V_0(Z_t) + \sigma(X_t) \varepsilon_{t+1}\right]\right\} = 0$ for all $s \neq t$, where $Y_t = \delta(Z_t)/\sigma(X_t)$.

Remark 6 – Conditions (C₂) and (C₃) hold, for example, if $V_0(Z_t) = 0$, $\sigma(X_t) = \sigma$ and $\delta(Z_t)$ is a function of the ε_t 's only or of the $X_{t-j} \varepsilon_{t-j}$'s, $j \geq 0$ only. They also hold with $V_0(Z_t) = b X_{t-2} \varepsilon_{t-1}$, $\delta(Z_t) = \tau \varepsilon_t$ and $\sigma(X_t) = \sigma$, for all b, τ . Many other examples can be found through Examples 1–4.

To calculate the likelihood ratio, we follow the approach of Benghabrit and Hallin (1992, 1996a,b). Denote by $X^{(n)} = (X_1, \dots, X_n)$ a finite realization of the stationary solution of (1) under \mathcal{H}_1^n . Denote by $\tilde{X}_0 = (X_0, \dots, X_{-m}, \varepsilon_0, \dots, \varepsilon_{-m})$, where $m = \max(p, q, Q)$, the observable starting values, and assume that under \mathcal{H}_1^n , the distribution of \tilde{X}_0 admits a nowhere vanishing density $g_{\delta/\sqrt{n}}$. We also assume that the function $g_\delta(\cdot)$ is continuous at $\delta = 0$. The likelihood functions for $(\tilde{X}_0, X^{(n)})$ under \mathcal{H}_0 and \mathcal{H}_1^n are respectively

$$L_n(\theta; V_0, 0, \sigma) = g_0(\tilde{X}_0) \prod_{t=1}^n f(W_{t+1}) \quad \text{and} \quad L_n(\theta; V_0, \delta, \sigma) = g_{\delta/\sqrt{n}}(\tilde{X}_0) \prod_{t=1}^n f(W_{t+1} - n^{-1/2} Y_t),$$

where $W_{t+1} = \left(X_{t+1} - \langle \theta, X_t \rangle - V_0(Z_t)\right)/\sigma(X_t)$ and $Y_t = \delta(Z_t)/\sigma(X_t)$.

To simplify the expression of the likelihood functions and the loglikelihood ratio, we will consider likelihoods associated with the starting values \tilde{X}_0 . The influence of these starting values on the asymptotic results is null. The loglikelihood ratio Λ_n decomposes into :

$$\Lambda_n = \Lambda_n(\theta; V_0, \delta, \sigma) = \sum_{t=1}^n \left[\log f\left(W_{t+1} - n^{-1/2} Y_t\right) - \log f\left(W_{t+1}\right) \right] + o_P(1). \quad (13)$$

Theorem 3 Under (A₁)–(A₅), (B₁)–(B₅) and (C₁)–(C₃), the sequences of hypotheses $\{\mathcal{H}_1^n : n \geq 1\}$ and $\{\mathcal{H}_0^n = \mathcal{H}_0 : n \geq 1\}$ are contiguous.

Proof of Theorem 3 – By Proposition 7 of Le Cam (1986), p. 90 (see Appendix 2), it suffices to show that under the requirements (B₁)–(B₅), (C₁)–(C₃) and under \mathcal{H}_0 , Λ_n converges in distribution to a Gaussian rv with mean $-\lambda^2/2$ and variance $\lambda^2 > 0$. By a Taylor expansion of $\log f$ around W_{t+1} , we have :

$$\Lambda_n = n^{-1/2} \sum_{t=1}^n Y_t \phi_f(W_{t+1}) - (2n)^{-1} \left[\sum_{t=1}^n Y_t^2 \phi'_f(W_{t+1} + \mu_t^n n^{-1/2} Y_t) \right] + o_P(1), \quad (14)$$

for some sequence $\mu_t^n \in (0, 1)$. In view of (B₃), we have

$$\left| n^{-1} \sum_{t=1}^n Y_t^2 \phi'_f(W_{t+1} + \mu_t^n n^{-1/2} Y_t) - n^{-1} \sum_{t=1}^n Y_t^2 \phi'_f(W_{t+1}) \right|$$

$$\begin{aligned}
&\leq n^{-1} \sum_{t=1}^n \left| Y_t^2 \left[\phi'_f(W_{t+1} + \mu_t^n n^{-1/2} Y_t) - \phi'_f(W_{t+1}) \right] \right| \\
&\leq n^{-3/2} C_\phi \sum_{t=1}^n |Y_t|^3.
\end{aligned} \tag{15}$$

By ergodicity, under (\mathcal{B}_2) and (\mathcal{C}_2) the right-hand side of (15) converges a.s. to 0. Therefore, we have :

$$\Lambda_n = n^{-1/2} \sum_{t=1}^n Y_t \phi_f(W_{t+1}) - (2n)^{-1} \left[\sum_{t=1}^n Y_t^2 \phi'_f(W_{t+1}) \right] + o_P(1). \tag{16}$$

Recall that under \mathcal{H}_0 , the $W_{i+1} = \varepsilon_{i+1}(\theta; V_0, 0, \sigma)$'s are iid independent of $\sigma(X_j : j \leq i)$, with density f satisfying the requirements (\mathcal{B}_1) – (\mathcal{B}_5) . By ergodicity, the second term in the right-hand side of (16) converges a.s. to

$$(1/2) E \left\{ Y_1^2 \phi'_f(W_{t+1}) \right\} = (1/2) E \left(Y_1^2 \right) I(f).$$

By Theorem 1 of Rio (1995) under (\mathcal{B}_5) and (\mathcal{C}_2) , the first term in the right-hand side of (16) converges in distribution to $\mathcal{N}(0, \lambda^2)$ with

$$\lambda^2 = E \left(Y_1^2 \right) \int \phi_f^2(x) f(x) dx = -E \left(Y_1^2 \right) I(f).$$

Finally, $\Lambda_n \xrightarrow{d} \mathcal{N}(-\lambda^2/2, \lambda^2)$. \diamond

Theorem 4 Assume that (\mathcal{A}_1) – (\mathcal{A}_5) , (\mathcal{B}_1) – (\mathcal{B}_5) and (\mathcal{C}_1) – (\mathcal{C}_3) are satisfied. At the asymptotic level of significance α , the asymptotic power of the test against \mathcal{H}_1^n is

$$1 - \Phi(\sqrt{\chi_\alpha} - \pi) + \Phi(-\sqrt{\chi_\alpha} - \pi),$$

where $\pi = \rho_{\mathcal{H}_0}^{-1} E(\delta(Z_1)) \int x \phi_f(x) f(x) dx$, $\rho_{\mathcal{H}_0}$ is defined in Theorem 1 and χ_α is the $(1-\alpha)$ -quantile of the $\chi^2(1)$ distribution.

Remark 7 – Under the assumption that $\lim_{x \rightarrow -\infty} x f(x) = \lim_{x \rightarrow +\infty} x f(x) = 0$, it can be proved that $\pi = -\rho_{\mathcal{H}_0}^{-1} E(\delta(Z_1))$. This assumption is satisfied by the common densities.

Proof of Theorem 4 – By Le Cam's third lemma (see, e.g., Hájek and Šidáček, 1967), it suffices to show that under \mathcal{H}_1^n , $S_n / \hat{\rho}_n \xrightarrow{d} \mathcal{N}(\pi, 1)$. We first prove that under \mathcal{H}_0 , $(\Lambda_n, S_n / \hat{\rho}_n)$ converges in distribution to a bidimensional Gaussian rv. We have

$$\left| \frac{S_n}{\hat{\rho}_n} - \frac{S_n}{\rho_{\mathcal{H}_0}} \right| = \left| S_n \right| \left| \frac{\hat{\rho}_n - \rho_{\mathcal{H}_0}}{\rho_{\mathcal{H}_0} \hat{\rho}_n} \right| \xrightarrow{P} 0.$$

Hence, $S_n/\hat{\rho}_n = S_n/\rho_{\mathcal{H}_0} + o_P(1)$ and we can write $(\Lambda_n, S_n/\hat{\rho}_n) = (\tilde{\Lambda}_n, \tilde{S}_n) + (o_P(1), o_P(1))$, where $\tilde{\Lambda}_n$ and \tilde{S}_n are respectively Λ_n and $S_n/\hat{\rho}_n$ minus the $o_P(1)$'s terms. For $\omega_1, \omega_2 \in \mathbb{R}$ we have

$$\omega_1 \tilde{\Lambda}_n + \omega_2 \tilde{S}_n = n^{-1/2} \sum_{t=1}^n \left[\omega_1 Y_t \phi_f(W_{t+1}) + \omega_2 \left\{ X_{t+1} - \langle \theta, X_t \rangle \right\} \rho_{\mathcal{H}_0}^{-1} \right] \quad (17)$$

$$+ n^{-1} \sum_{t=1}^n \omega_2 Y_t^2 \phi'_f(W_{t+1}). \quad (18)$$

By ergodicity, the term in (18) converges a.s. to $\omega_2 E(Y_t^2) I(f)$. For the study of the convergence of the right-hand side of (17), we remark that under (\mathcal{A}_3) and (\mathcal{B}_4) ,

$$E \left[\omega_1 Y_1 \phi_f(W_1) + \omega_2 \left\{ X_2 - \langle \theta, X_1 \rangle \right\} \rho_{\mathcal{H}_0}^{-1} \right] = 0.$$

Under the assumptions (\mathcal{A}_2) , (\mathcal{B}_5) and (\mathcal{C}_1) – (\mathcal{C}_3) , Theorem 1 of Rio (1995) implies that the right-hand side of (17) converges to $\mathcal{N}(0, \nu^2)$ with $\nu^2 = \sum_{t \in \mathbb{Z}} \text{cov}(\psi_1, \psi_t)$, where $\psi_t = \omega_1 Y_t \phi_f(W_{t+1}) + \omega_2 \left\{ X_{t+1} - \langle \theta, X_t \rangle \right\} \rho_{\mathcal{H}_0}^{-1}$. Therefore, $(\tilde{\Lambda}_n, \tilde{S}_n)$ converges to a bidimensional Gaussian rv. To characterize the limiting covariance matrix, we only need to calculate $\lim_{n \rightarrow \infty} \text{cov}(\tilde{\Lambda}_n, \tilde{S}_n)$. Under \mathcal{H}_0 , (\mathcal{A}_1) – (\mathcal{A}_5) , (\mathcal{B}_1) – (\mathcal{B}_5) and (\mathcal{C}_1) – (\mathcal{C}_3) , we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{cov}(\tilde{\Lambda}_n, \tilde{S}_n) &= \lim_{n \rightarrow \infty} \rho_{\mathcal{H}_0}^{-1} n^{-1} \sum_{t=1}^n \sum_{s=1}^n E \left\{ Y_t \phi_f(W_{t+1}) \left[V_0(Z_s) + \sigma(X_s) W_{s+1} \right] \right\} \\ &\quad + \lim_{n \rightarrow \infty} \rho_{\mathcal{H}_0}^{-1} n^{-3/2} \sum_{t=1}^n \sum_{s=1}^n E \left\{ Y_t^2 \phi'_f(W_{t+1}) \left[V_0(Z_s) + \sigma(X_s) W_{s+1} \right] \right\} \\ &= \rho_{\mathcal{H}_0}^{-1} E \{ \delta(Z_1) \} \int x \phi_f(x) f(x) dx = \pi. \end{aligned}$$

In conclusion, $(\Lambda_n, S_n/\hat{\rho}_n) \xrightarrow{d} \mathcal{N}(\mu, \Sigma)$, where $\mu = (-\lambda^2/2, 0)$ and $\Sigma = \begin{pmatrix} \lambda^2 & \pi \\ \pi & 1 \end{pmatrix}$. By Le Cam's third Lemma, $S_n/\hat{\rho}_n \xrightarrow{d} \mathcal{N}(\pi, 1)$ under \mathcal{H}_1^n . \diamond

5. Comparison with other tests – We now compare the asymptotic local power obtained above with those of the rank-based tests derived in Benghabrit and Hallin (1996b) and with that of the Lagrange multiplier (LM) test (Benghabrit and Hallin 1996a). This comparison is made for the first-order bilinear process

$$X_{i+1} = aX_i + n^{-1/2} \beta X_i \varepsilon_i + \varepsilon_{i+1}. \quad (19)$$

Here, $V_0(X_i, \varepsilon_i) = 0$, $\delta(X_i, \varepsilon_i) = \beta X_i \varepsilon_i$, $\sigma(X_i) = 1$ and the ε_i 's are standard Gaussian rv's. It can be checked that the requirements (\mathcal{B}_1) – (\mathcal{B}_5) and (\mathcal{C}_1) – (\mathcal{C}_3) hold, and that $\rho_{\mathcal{H}_0} = 1$

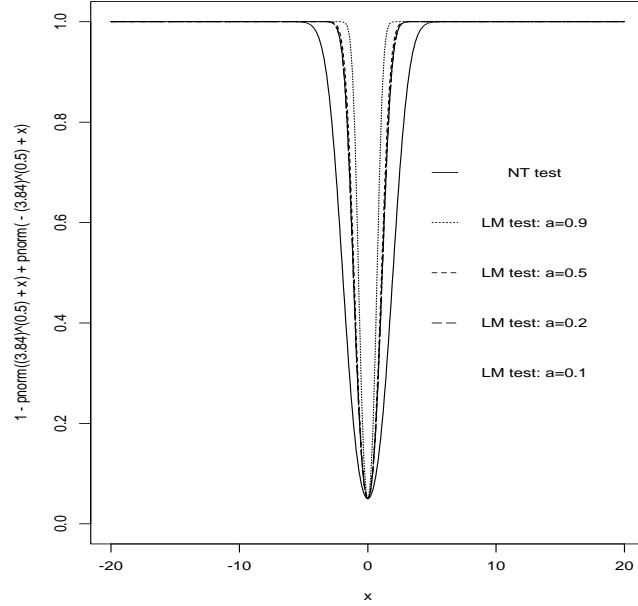


Figure 1: Asymptotic local power of the NT and LM tests against a first-order diagonal bilinear model, with Gaussian noise, for $\alpha = 5\%$ and $a=0.9; 0.5; 0.2; 0.1$.

and $\int x\phi_f(x)f(x)dx = -1$. We need to calculate $E(\delta(X_i, \varepsilon_i)) = \beta E(X_i\varepsilon_i)$ under \mathcal{H}_0 . To this end, it suffices to calculate $E(X_i\varepsilon_i)$ under \mathcal{H}_0 . If we put $\beta = 0$ in (19), we are under \mathcal{H}_0 . Then, multiplying each member of this equation by ε_{i+1} and taking the expectations, we obtain $E(X_i\varepsilon_i) = 1$, and therefore $\pi = -\beta$. The asymptotic local power of our test at the asymptotic level of significance α is then

$$P_{NT}(a, \alpha, \beta) = 1 - \Phi(\sqrt{\chi_\alpha} + \beta) + \Phi(-\sqrt{\chi_\alpha} + \beta).$$

From Benghabrit and Hallin (1996a) it can be seen that the asymptotic local power of the LM test for (19) at the asymptotic level of significance α is

$$P_{LM}(a, \alpha, \beta) = 1 - \Phi\{\sqrt{\chi_\alpha} + \beta[3 + a^2(1 - a^2)^{-1}]^{1/2}\} + \Phi\{-\sqrt{\chi_\alpha} + \beta[3 + a^2(1 - a^2)^{-1}]^{1/2}\}.$$

This power is equal to those of the rank-based tests derived in Benghabrit and Hallin (1996b).

We have plotted the theoretical local power of our test and that of the LM test for $\alpha = 0.05$ and $a = 0.1; 0.2; 0.5; 0.9$ (see Fig. 1). For $|a| \leq 0.2$ the curves of the LM test are the same. For each value of a and for $|\beta| > 2\sqrt{\chi_\alpha}/\{(3 - 2a^2)^{1/2}(1 - a^2)^{-1/2} - 1\}$ and $\beta = 0$, the local powers are nearly equal. For $\beta \in (-2\sqrt{\chi_\alpha}/\{(3 - 2a^2)^{1/2}(1 - a^2)^{-1/2} - 1\}, 0) \cup (0, 2\sqrt{\chi_\alpha}/\{(3 - 2a^2)^{1/2}(1 - a^2)^{-1/2} - 1\})$, the power of the LM test is larger than that of our test. For $a \leq 0.5$, the gap between the two plots is not too large. For $a = 0.9$, this gap becomes very large. We also remark that both powers grow quickly to 1 as β moves away from 0.

		$n = 300$				$n = 600$			
a		-0.5		0.2		-0.5		0.2	
b		NT	LM	NT	LM	NT	LM	NT	LM
-0.70		1.00	0.97	1.00	0.96	1.00	0.98	1.00	0.93
-0.50		1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
-0.30		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.10		0.40	0.85	0.40	0.86	0.70	0.99	0.68	1.00
-0.05		0.14	0.34	0.13	0.34	0.22	0.62	0.22	0.56
-0.00		0.047	0.051	0.044	0.051	0.053	0.053	0.051	0.043
0.05		0.14	0.36	0.13	0.34	0.23	0.56	0.23	0.57
0.10		0.41	0.86	0.38	0.83	0.67	0.99	0.66	0.98
0.30		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.50		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.70		1.00	0.97	1.00	0.97	1.00	1.00	1.00	0.99

Table 1 : Experimental powers of the NT and LM tests, for testing a linear AR(1) model against a first-order diagonal bilinear model with Gaussian noise at level of significance 5%, for $a = -0.5, 0.2$, for $n = 300, 600$ and with 10,000 replications.

6. Simulation experiments – In this section, we study the power of our test by doing Monte Carlo simulations for the following stochastic difference equations:

- (i) $X_{i+1} = aX_i + bX_i\varepsilon_i + \varepsilon_{i+1}$.
- (ii) $X_{i+1} = aX_i + \tau X_{i-1}\varepsilon_i + bX_i\varepsilon_i + \varepsilon_{i+1}$.
- (iii) $X_{i+1} = aX_i + bX_i \exp(-\lambda X_i^2) + \varepsilon_{i+1}$.
- (iv) $X_{i+1} = aX_i + \tau X_{i-1}\varepsilon_i + b\varepsilon_i \exp(-\lambda X_i^2) + \varepsilon_{i+1}$.

$$(v) X_{i+1} = aX_{i-1} + \tau\varepsilon_i + bX_i\varepsilon_i + \sqrt{c + dX_i^2}\varepsilon_{i+1}.$$

$$(vi) X_{i+1} = aX_{i-1} + \tau X_{i-1}\varepsilon_i + bX_i I(X_i < 0) + \varepsilon_{i+1}.$$

In all these examples, we test the hypothesis $\mathcal{H}_0(b = 0)$ against the alternative $\mathcal{H}_1(b \neq 0)$. The ε_i 's are standard iid Gaussian rv's and the parameters τ, c and d are given. The tests proposed by An and Cheng (1991) cannot be used for testing these hypotheses. The test of McKeague and Zhang (1994) can only be used for case (iii).

	$n = 300$		$n = 600$	
a	-0.5	0.2	-0.5	0.2
b				
	NT	NT	NT	NT
-0.70	0.97	1.00	0.99	1.00
-0.50	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00
-0.10	0.37	0.39	0.63	0.66
-0.05	0.13	0.14	0.22	0.24
-0.00	0.05	0.054	0.052	0.049
0.05	0.13	0.11	0.22	0.19
0.10	0.364	0.31	0.65	0.62
0.30	1.00	1.00	1.00	1.00
0.50	1.00	1.00	1.00	1.00
0.70	1.00	1.00	1.00	1.00

Table 2 : Experimental power of the NT test, for testing the presence of a of first-order diagonal term in a second-order subdiagonal bilinear model with Gaussian noise at level of significance 5%, for $a = -0.5; 0.2$, for $n = 300, 600$ and with 10,000 replications.

For the cases (i)-(ii), $\hat{\theta}_n = \hat{a}$ is the least squares estimate of a . For the case (iv), $\hat{a} = \sum_{i=1}^n X_{i+1}X_{i-1} / \sum_{i=1}^n X_iX_{i-1}$. For the cases (v) and (vi), $\hat{a} = \sum_{i=1}^n X_{i+1}X_{i-1} / \sum_{i=1}^n X_i^2$. For all these cases, $\hat{\rho}_n^2 = n^{-1}(1 - \hat{a}^2) \sum_{i=1}^n X_i^2$. The power is simulated against the alternatives (i)-(vi) as a function of the parameter of interest b at the asymptotic level of significance $\alpha = 5\%$, for $a = -0.5$ and 0.2 , sample sizes $n = 300$ and 600 , and making 10,000 replications.

The experimental power of the NT test against (i) has been compared with the experimental power of the NP test of Diebolt and Ngatchou Wandji (1996a). The results are very similar and are not presented here. Table 1 shows that in the vicinity of the null hypothesis, the experimental power of the LM test is better than that of the NT test. Far from the null hypothesis, the power of the NT test is often better than that of the LM test. Under the alternative (ii), which is a second-order subdiagonal bilinear model, the power of our test is good both in the neighborhood of the null hypothesis and far from it (see Table 2). Under the alternatives (iii) and (iv) with fixed $\tau = 0.1$, these powers are weak. No value of the power is above 10% (see Tables 3 and 4). One could expect that the power would increase with $|b|$, as it does against the alternatives (i) and (ii) with fixed $\tau = 0.1$. But it has an unexpected behavior. However, the level of significance is close to the nominal one. It seems

that very few nonparametric type tests are powerful against exponential models (see Table 5 (vii) of Hjellvik and Tjøstheim, 1995, p. 364).

Under the alternatives (v) with fixed $\tau = 0.01, c = 0.99, d = 0.1$ and (vi) with fixed $\tau = 0.01$, the power of the test is good both in the vicinity of the null hypothesis and far from it (see Tables 5 and 6).

	$n = 300$		$n = 600$	
a	-0.5	0.2	-0.5	0.2
b	NT	NT	NT	NT
-0.70	0.054	0.053	0.057	0.051
-0.50	0.050	0.053	0.053	0.045
-0.30	0.048	0.050	0.050	0.051
-0.10	0.046	0.050	0.047	0.049
-0.05	0.048	0.045	0.050	0.043
-0.00	0.049	0.049	0.053	0.051
0.05	0.047	0.046	0.048	0.041
0.10	0.047	0.044	0.047	0.050
0.30	0.052	0.050	0.050	0.050
0.50	0.047	0.050	0.053	0.052
0.70	0.056	0.050	0.051	0.055

Table 3 : Experimental power of the NT test for testing a linear $AR(1)$ model against a first-order $EXPAR$ model with Gaussian noise at level of significance 5%, for $a = -0.5, 0.2$, for $n = 300, 600$ and with 10,000 replications.

	$n = 300$		$n = 600$	
a	-0.5	0.2	-0.5	0.2
b	NT	NT	NT	NT
-0.70	0.085	0.032	0.010	0.032
-0.50	0.016	0.038	0.018	0.032
-0.30	0.038	0.040	0.028	0.044
-0.10	0.044	0.048	0.040	0.047
-0.05	0.048	0.044	0.045	0.042
-0.00	0.051	0.050	0.051	0.051
0.05	0.052	0.047	0.051	0.048
0.10	0.073	0.047	0.055	0.052
0.30	0.073	0.057	0.069	0.058
0.50	0.079	0.063	0.089	0.062
0.70	0.099	0.067	0.094	0.073

Table 4 : Experimental power of the NT test for testing the presence of a first-order exponential term in a second-order subdiagonal bilinear model with Gaussian noise at level of significance 5%, for $a = -0.5, 0.2$, for $n = 300, 600$ and with 10,000 replications.

	$n = 300$		$n = 600$	
a	-0.5	0.2	-0.5	0.2
b	NT	NT	NT	NT
-0.70	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00
-0.10	0.38	0.38	0.67	0.67
-0.05	0.14	0.14	0.24	0.22
-0.00	0.051	0.048	0.051	0.052
0.05	0.13	0.13	0.22	0.23
0.10	0.41	0.39	0.67	0.66
0.30	1.00	1.00	1.00	1.00
0.50	1.00	1.00	1.00	1.00
0.70	1.00	1.00	1.00	1.00

Table 5 : Experimental power of the NT test for testing the presence of a first-order diagonal term in a second-order heteroscedastic model with Gaussian noise at level of significance 5%, for $a = -0.5, 0.2$, for $n = 300, 600$ and with 10,000 replications.

	$n = 300$		$n = 600$	
a	-0.5	0.2	-0.5	0.2
b	NT	NT	NT	NT
-0.70	1.00	0.97	1.00	1.00
-0.50	0.84	0.78	1.00	1.00
-0.30	0.56	0.38	0.88	0.73
-0.10	0.11	0.085	0.18	0.134
-0.05	0.058	0.047	0.08	0.068
-0.00	0.048	0.046	0.049	0.050
0.05	0.073	0.072	0.10	0.088
0.10	0.13	0.121	0.21	0.20
0.30	0.71	0.68	0.94	0.92
0.50	0.98	0.98	1.00	1.00
0.70	1.00	1.00	1.00	1.00

Table 6 : Experimental power of the NT test for testing the presence of a threshold term in a second-order subdiagonal bilinear model with Gaussian noise at level of significance 5%, for $a = -0.5, 0.2$, for $n = 300, 600$ and with 10,000 replications.

7. Conclusion – We have proposed a new test for discriminating between autoregressive models. This test is certainly not an optimal one. But our simulation experiments show that it can be useful for testing nonlinear time series models to which none of the existing tests can be applied. This test has the advantage of being simple to use. We have proved its convergence under the assumption that the process is not centered at 0. An expression

for its asymptotic local power has been obtained. We have made a detailed comparison of the local power of our test with the corresponding local powers of the tests studied in Benghabrit and Hallin (1996a,b) in the special case where $p = q = Q = 1$, $V_0 = 0$, $\sigma = 1$, $V_*(x, y) = bxy$ and ε is Gaussian. We have conducted Monte Carlo simulations to study the power of our test in several situations. Despite the fact that this test is not sufficiently powerful against EXPARMA models, it appears to be a good tool for testing time series models.

Appendix 1

Define $\alpha(t) = \alpha_{[t]}$ and denote by f^{-1} the càdlàg (right-continuous with a left limit) generalized inverse function of a nonincreasing function f : $f^{-1}(u) = \inf\{t : f(t) \leq u\}$. For any random variable ξ with distribution function F_ξ , denote by Q_ξ the tail quantile function, which is the inverse of the tail function: $t \mapsto P\{\xi > t\} = 1 - F_\xi(t)$.

Theorem 5 (Rio (1995)) *Let $\{\xi_i : i \in \mathbf{Z}\}$ be a stationary and strongly mixing sequence of real centered random variables such that*

$$M_{2,\alpha}^\xi = \int_0^1 \alpha(u)[Q_{\xi_0}(u)]^2 du < \infty. \quad (20)$$

Let $S_n = \sum_{i=1}^n \xi_i$. Then :

- (i) *The series $\sum_{i \in \mathbf{Z}} \text{cov}(\xi_0, \xi_i)$ is absolutely convergent to a nonnegative number ρ^2 and $n^{-1} \text{var}(S_n)$ converges to ρ^2 .*
- (ii) *$n^{-1/2} S_n$ converges in distribution to the (possibly degenerate) normal distribution $\mathcal{N}(0, \rho^2)$.*

The following proposition gives a sufficient condition for (20) to hold.

Proposition 2 *Let $\{\xi_i : i \in \mathbf{Z}\}$ be a stationary and strongly mixing sequence of real-valued centered random variables. If $\sum_{k \geq 0} (k+1)\alpha_k^{(r-2)/r} < \infty$ and $E(|\xi_1|^r) < \infty$ for some $r > 2$, then*

$$M_{2,\alpha}^\xi = \int_0^1 \alpha(u)[Q_{\xi_0}(u)]^2 du < \infty.$$

Proof – Assume that $\sum_{k \geq 0} (k+1)\alpha_k^{(r-2)/r} < \infty$ and $E(|\xi_1|^r) < \infty$ for some $r > 2$. The proof follows from the following equalities and inequalities:

$$\begin{aligned} M_{2,\alpha}^\xi &= \int_0^{\alpha_0} \alpha(u)[Q_{\xi_0}(u)]^2 du \\ &= \sum_{k \geq 0} \int_{\alpha_{k+1}}^{\alpha_k} \alpha(u)[Q_{\xi_0}(u)]^2 du \\ &= \sum_{k \geq 0} \int_0^1 I_{[\alpha_{k+1}, \alpha_k]}(u) \alpha(u)[Q_{\xi_0}(u)]^2 du \\ \text{(Hölder)} &\leq \sum_{k \geq 0} \left\{ \int_0^1 (I_{[\alpha_{k+1}, \alpha_k]}(u) \alpha(u))^{r/(r-2)} du \right\}^{(r-2)/r} \left\{ \int_0^1 [Q_{\xi_0}(u)]^r du \right\}^{2/r} \\ &\leq \left\{ \sum_{k \geq 0} (k+1)\alpha_k^{(r-2)/r} \right\} \left\{ \int_0^1 [Q_{\xi_0}(u)]^r du \right\}^{2/r} \end{aligned}$$

$$\leq \left\{ \sum_{k \geq 0} (k+1) \alpha_k^{(r-2)/r} \right\} [E(|\xi_1|^r)]^{2/r} < \infty. \diamond$$

Appendix 2

Consider a sequence of measurable spaces $(\Omega_n, \mathcal{A}_n)$. Let P_n and Q_n be two sequences of probability measures on $(\Omega_n, \mathcal{A}_n)$.

Definition 1 The sequence $\{Q_n\}$ is contiguous to the sequence $\{P_n\}$ if for each event $A_n \in \mathcal{A}_n$ $\lim_{n \rightarrow \infty} P_n(A_n) = 0 \implies \lim_{n \rightarrow \infty} Q_n(A_n) = 0$.

Suppose that P_n and Q_n are mutually absolutely continuous and let $\Lambda_n = \log \frac{dQ_n}{dP_n}$.

Proposition 3 (Le Cam 1986, p. 90) Assume that under P_n the distribution of Λ_n is asymptotically normal $\mathcal{N}(a_n, \lambda_n^2)$. Then, $\{Q_n\}$ is contiguous to $\{P_n\}$ if and only if

(i) $\limsup_n \lambda_n^2 < \infty$;

(ii) $\lim_{n \rightarrow \infty} a_n + \lambda_n^2/2 = 0$.

If so, $\{Q_n\}$ and $\{P_n\}$ are contiguous and under Q_n , Λ_n is asymptotically normal $\mathcal{N}(a_n + \lambda_n^2/2, \lambda_n^2)$.

Lemma 1 (Le Cam's third Lemma, Hajek and Sidak 1967, p. 208) Let S_n be measurable with respect to \mathcal{A}_n . Assume that the pair (S_n, Λ_n) is under P_n asymptotically jointly normal $\mathcal{N}(\mu, \Sigma)$ with $\mu = (-\lambda_1^2/2, \mu_1)$ and $\Sigma = \begin{pmatrix} \lambda_1^2 & \lambda_{12} \\ \lambda_{12} & \lambda_2^2 \end{pmatrix}$. Then S_n is under Q_n asymptotically normal $\mathcal{N}(\mu_1 + \lambda_{12}, \lambda_1^2)$.

Acknowledgments – I am grateful to Jean Diebolt for fruitful comments and suggestions on the manuscript. I am indebted to Gilles Celeux for providing me a good environment of work in INRIA Rhône-Alpes.

References

- An, H. Z. and Cheng, B. (1991). A Kolmogorov–Smirnov type statistic with application to test for nonlinearity in times series. *International Statist. Review*, **59**, 287–307.
- Benghabrit, Y. and Hallin, M. (1992). Optimal Rank-based tests against first-order super-diagonal bilinear dependence. *J. Statist. Plann. Inference* **32**, 45–61.
- Benghabrit, Y. and Hallin, M. (1996a). Locally asymptotically optimal tests for autoregressive against bilinear serial dependence. *Statist. Sinica* **6**, 147–169.
- Benghabrit, Y. and Hallin, M. (1996b). Rank-based tests for autoregressive against bilinear serial dependence. *Nonparametr. Statist.* **6**, 253–272.
- Diebolt, J. (1990). Testing the function defining a nonlinear autoregressive time series. *Stochastic Process. Appl.* **36**, 85–106.
- Diebolt, J. and Laïb, N. (1995). Nonparametric tests for correlation or autoregression models under mixing conditions. *C. R. Acad. Scie. Paris Ser I*, **320**, 135–139.
- Diebolt, J. and Ngatchou Wandji, J. (1995). Un test non-paramétrique pour un modèle bilinéaire diagonal d'ordre un. *C. R. Acad. Scie. Paris Ser I*, **320**, 1135–1139.
- Diebolt, J. and Ngatchou Wandji, J. (1996a). A nonparametric test for the linearity against bilinearity for autoregressive processes. *Prépublications Mathématiques de l'Université Paris-Nord*, **96–04**.

- Diebolt, J. and Ngatchou Wandji, J. (1996b). A nonparametric test for generalized first-order autoregressive models. *Scand. J. Statist.*, to appear.
- Doukhan, P. (1994). *Mixing: Properties and Examples*. Lecture Notes in Statistics **85**.
- Guégan, D. and Ngatchou Wandji, J. (1996). Power of the Lagrange multiplier test for certain subdiagonal bilinear models. *Statist. Probab. Letters* **29**, 201–212.
- Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press.
- Hili, O. (1995). On the estimation of nonlinear time series models. *Stochastics and Stochastic Rep.* **52**, 207–226.
- Hinnich, M.J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* **3**, 169–176.
- Hjellvik, V. and Tjøstheim, D. (1995). Nonparametric tests of linearity for time series. *Biometrika* **82**, 355–68.
- Keenan, D. (1985). A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika* **72**, 39–44.
- Le Cam, L. (1986). *Asymptotic methods in Statistical decision theory*. Springer-Verlag.
- Luukonen, R., Saikkonen, P. and Terasvirta, T. (1988). Testing linearity in univariate time series models. *Scand. J. Statist.* **15**, 167–175.
- McKeague, I.W. and Zhang, M.J. (1994). Identification of nonlinear time series from first order cumulative characteristics. *Ann. Statist.*, **22**, 495–514.
- Mokkadem, A. (1990). Propriétés de mélange des processus autoregressifs polynomiaux. *Ann. Inst. Henri Poincaré Probab. Statist.*, **26**, 219–260.
- PHAM, T. D. and TRAN, L. T. (1981). On the first order bilinear time series models. *J. Appl. Probab.* **18**, 617–627.
- PHAM, T. D. (1985). Bilinear Markovian representation and bilinear models. *Stochastic Process. Appl.* **20**, 295–306.
- Poggi, J.M. and Portier, B. (1995). Un test de linéarité pour les modèles autorégressifs fonctionnels. *C. R. Acad. Scie. Paris Ser I*, **321**, 113–116.
- Rio, E. (1995). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23**, 1188–1203.
- Robinson, P.M. (1977). The estimation of a nonlinear moving average model. *Stochastic Process. Appl.* **5**, 81–90.
- Saikkonen, P. and Luukonen, R. (1988). Lagrange multiplier tests for testing non linearities in time series models. *Scand. J. Statist.* **15**, 55–68.
- Shorack, G. and Wellner, J. (1986). *Empirical Processes With Applications to Statistics*. Wiley: New York.
- Subba Rao, T. and Gabr, M. M. (1980). A test for linearity of stationary time series. *J. Time Ser. Anal.* **1**, 145–158.
- Tsay, R.S. (1986). Nonlinearity tests for time series. *Biometrika*. 461–466.



Unit e de recherche INRIA Lorraine, Technop ole de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS L ES NANCY
Unit e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit e de recherche INRIA Rh one-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

 diteur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399