Strict stationarity testing and estimation of explosive ARCH models

Francq, Christian and Zakoian, Jean-Michel

April 2010

Strict stationarity testing and estimation of explosive ARCH models

By Christian Francq*

University Lille 3, EQUIPPE-GREMARS

and

By Jean-Michel Zakoïan*

CREST and University Lille 3

This paper studies the asymptotic properties of the quasi-maximum likelihood estimator of ARCH(1) models without strict stationarity constraints, and considers applications to testing problems. The estimator is unrestricted, in the sense that the value of the intercept, which cannot be consistently estimated in the explosive case, is not fixed. A specific behavior of the estimator of the ARCH coefficient is obtained at the boundary of the stationarity region, but this estimator remains consistent and asymptotically normal in every situation. The asymptotic variance is different in the stationary and non stationary situations, but is consistently estimated, with the same estimator, in both cases. Tests of strict stationarity and non stationarity are proposed. Their behaviors are studied under the null assumption and under local alternatives. The tests developed for the ARCH(1) model are able to detect non-stationarity in more general GARCH models. A numerical illustration based on stock indices is proposed.

1. Introduction. Testing for strict stationarity is an important issue in the context of financial time series. A standard assumption is that the prices are non stationary while the returns (or log-returns) are stationary. Numerous statistical tools, such as the unit root tests, have been introduced for testing the non-stationarity of prices. For the log-returns, the most widely used models are arguably the GARCH introduced by Engle (1982) and Bollerslev (1986). No statistical tools are available for testing strict stationarity in the GARCH framework. This is the main aim of this paper to develop such tools. The problem is non standard because, contrary to stationarity in linear time

*The authors are very grateful to Professor Y. Davydov and to Professor L. Horváth for stimulating and instructive discussions on topics related to this paper.

AMS 2000 subject classifications: Primary 62M10; secondary 62F12, 62F05.

Keywords and phrases: ARCH model, Inconsistency of estimators, Local power of tests, Nonstationarity, Quasi Maximum Likelihood Estimation

series models, which solely depends on the lag polynomials, the strict stationarity condition for GARCH models has a non explicit form, involving the distribution of the underlying independent and identically distributed (iid) sequence.

The asymptotic properties of the quasi-maximum likelihood estimator (QMLE) for classical GARCH models have been extensively studied; see Berkes, Horváth and Kokoszka (2003), Francq and Zakoïan (2004) and the references therein. For valid statistical inference based on those results, strict stationarity must hold. Thus, from the point of view of the validity of the asymptotic results for the QMLE, strict stationarity testing in GARCH models is also an important issue. Surprisingly, this issue has not been addressed in the literature, to the best of our knowledge.

1.1. *Modes of divergence in the non stationary case.* The complexity of the statistical problem arises from the specificities of the probabilistic framework, even for the simplest GARCH model. To fix ideas, consider the ARCH(1) model, given by

$$\left\{\begin{array}{l}\epsilon_t = \sqrt{h_t}\eta_t, \quad t = 1, 2, \ldots \\ h_t = \omega_0 + \alpha_0\epsilon_{t-1}^2\end{array}\right. \tag{1.1}$$

with an initial value ϵ_0, where $\omega_0 > 0$, $\alpha_0 \geq 0$, and (η_t) is a sequence of independent and identically distributed (iid) variables such that $E\eta_1 = 0$ and $E\eta_1^2 = 1$. The necessary and sufficient condition for the existence of a strictly stationary solution to (1.1) is (by Nelson, 1990)

$$\gamma_0 < 0 \qquad (i.e.\ \alpha_0 < \exp\left\{-E\log\eta_1^2\right\}), \tag{1.2}$$

where $\gamma_0 = E\log(\alpha_0\eta_1^2)$. More precisely, if (1.2) holds we have

$$h_t - \sigma_t^2 \to 0 \text{ almost surely as } t \to \infty, \tag{1.3}$$

where

$$\sigma_t^2 = \lim_{n\to\infty} \uparrow \sigma_{t,n}^2, \qquad \sigma_{t,n}^2 = \omega_0\left(1 + \sum_{k=1}^{n-1}\alpha_0^k\eta_{t-1}^2\ldots\eta_{t-k}^2\right). \tag{1.4}$$

Let us now turn to the nonstationary case, for which it is necessary to consider separately $\gamma_0 > 0$ and $\gamma_0 = 0$. Under the assumption

$$\gamma_0 > 0 \qquad (i.e.\ \alpha_0 > \exp\left\{-E\log\eta_1^2\right\}), \tag{1.5}$$

$h_t \to \infty$ almost surely as $t \to \infty$, as shown by Nelson (1990). In this case, the increasing sequence $\sigma^2_{t,n}$ goes to infinity almost surely as $t \to \infty$, by the Cauchy root test. The case $\gamma_0 = 0$ is much more intricate. By the Chung-Fuchs theorem, it can be seen that $\sigma^2_{t,n}$ goes to infinity almost surely as $t \to \infty$. However, (1.3) may not hold when $\gamma_0 = 0$. Actually, Klüppelberg, Lindner and Maller (2004) (see also Goldie and Maller (2000)) showed that

$$\text{when } \gamma_0 = 0, \quad h_t \to \infty \text{ in probability} \tag{1.6}$$

instead of almost surely in the case $\gamma_0 > 0$ [1]. The astonishing difficulties encountered in the case $\gamma_0 = 0$ are related to the fact that the sequence $h_t = \sigma^2_{t,t} + \alpha_0^t \eta^2_{t-1} \dots \eta_1^2 \epsilon_0^2$ does not increase with t.

1.2. *The statistical problem.* Denote by $\theta = (\omega, \alpha)'$ the ARCH(1) parameter and define the QMLE as any measurable solution of

$$\hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n)' = \arg\min_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^{n} \ell_t(\theta), \quad \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta), \tag{1.7}$$

where Θ is a compact subset of $(0, \infty)^2$, and $\sigma_t^2(\theta) = \omega + \alpha \epsilon^2_{t-1}$ for $t = 1, \dots, n$ (with an initial value for ϵ_0^2). The rescaled residuals are defined by $\hat{\eta}_t = \eta_t(\hat{\theta}_n)$ where $\eta_t(\theta) = \epsilon_t / \sigma_t(\theta)$ for $t = 1, \dots, n$.

To construct a test of the strict stationarity assumption, we will establish the asymptotic distribution, under (1.2), of the statistic

$$\hat{\gamma}_n = \log \hat{\alpha}_n + \frac{1}{n} \sum_{t=1}^{n} \log \hat{\eta}_t^2.$$

This will be accomplished by deriving the joint distribution of $(\hat{\alpha}_n, \frac{1}{n} \sum_{t=1}^{n} \log \hat{\eta}_t^2)$, under the assumptions used to prove the asymptotic normality of the QMLE $\hat{\theta}_n$.

To study the asymptotic power of the test, it is necessary to analyze the asymptotic behavior of the QMLE when $\gamma_0 \geq 0$. Jensen and Rahbek (2004a, 2004b) were the first to establish an asymptotic theory for estimators of non-stationary GARCH. [2] They considered a constrained QMLE of α_0 (in the sense that the value of the intercept is fixed) which is consistent in the non stationary case, but is inconsistent in the stationary case. We will establish the strong consistency and asymptotic normality of the (unconstrained)

[1] Klüppelberg, Lindner and Maller (2004) noted that the arguments given by Nelson (1990) for the a.s. convergence are in failure when $\gamma_0 = 0$.

[2] See Linton, Pan and Wang (2009) for extensions in the case of non iid errors.

QMLE of α_0, the only component which matters for our testing problem, when $\gamma_0 > 0$. It turns out that, contrary to the strict stationarity case, the asymptotic distribution of $\hat{\alpha}_n$ is extremely simple, and is given by

$$\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\alpha_0^2\right\}, \qquad \text{as } n \to \infty$$

where $\xrightarrow{d}$ stands for the convergence in distribution. When $\gamma_0 = 0$, the QMLE of α_0 will be shown to be weakly consistent with the same asymptotic normal distribution as in the case $\gamma_0 > 0$. The asymptotic variances of $\hat{\alpha}_n$ when $\gamma_0 \geq 0$ and when $\gamma_0 < 0$ do not coincide, but we propose an estimator which is consistent in both situations. This is in accordance with similar results for autoregressive models with random coefficients derived by Aue and Horváth (2009).

The rest of the paper is organized as follows. Section 2 is devoted to the asymptotic properties of the QMLE. In Section 3, we first consider the problem of testing the value of α without any stationarity restriction. Then, we consider strict stationarity testing. The asymptotic distributions of two tests are studied when the null assumption is either the stationarity or the non stationarity. Section 4 is devoted to a power study. We start by establishing the Local Asymptotic Normality (LAN) of the ARCH(1) model without stationarity. The Fisher information matrix is degenerate in the case $\gamma_0 \geq 0$. The LAN property is used to derive the local asymptotic power of the proposed tests. Optimality issues are discussed. Necessary and sufficient conditions on the noise density are derived for the tests to be uniformly locally asymptotically most powerful. We also consider testing stationarity in more general GARCH-type models. Numerical illustrations are provided in Section 6. In particular, the stationarity of eleven major stock returns is analyzed. Proofs and complementary results, in particular the inconsistency of the constrained estimators in the stationary case, are collected in Section 7.

2. Asymptotic properties of the QMLE.

In this paper we consider the standard QMLE, which is the commonly used estimator for GARCH models.

2.1. *Consistency and asymptotic normality of $\hat{\alpha}_n$.*

The following result completes those already established in the stationary case, which we recall for convenience. The asymptotic distribution in the case $\gamma_0 = 0$ will be treated separately.

THEOREM 2.1. *For the ARCH(1) model (1.1), let the QMLE defined in (1.7).*

i) When $\gamma_0 < 0$ *and* $P(\eta_t^2 = 1) < 1$

$$\hat{\alpha}_n \to \alpha_0 \quad \text{and} \quad \hat{\omega}_n \to \omega_0 \qquad a.s. \text{ as } n \to \infty.$$

ii) When $\gamma_0 > 0$,

$$\hat{\alpha}_n \to \alpha_0 \qquad a.s. \text{ as } n \to \infty.$$

iii) When $\gamma_0 = 0$,

$$\hat{\alpha}_n \to \alpha_0 \qquad \text{in probability as } n \to \infty.$$

iv) When $\gamma_0 < 0$, $\kappa_\eta = E\eta_1^4 \in (1, \infty)$ *and* $\theta_0 = (\omega_0, \alpha_0)'$ *belongs to the interior* $\overset{\circ}{\Theta}$ *of* Θ,

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)J^{-1}\right\}, \qquad \text{as } n \to \infty, \tag{2.1}$$

and

$$J = E\begin{bmatrix} \frac{1}{(\omega_0+\alpha_0\epsilon_1^2)^2} & \frac{\epsilon_1^2}{(\omega_0+\alpha_0\epsilon_1^2)^2} \\ \frac{\epsilon_1^2}{(\omega_0+\alpha_0\epsilon_1^2)^2} & \frac{\epsilon_1^4}{(\omega_0+\alpha_0\epsilon_1^2)^2} \end{bmatrix}. \tag{2.2}$$

v) When $\gamma_0 > 0$, $\kappa_\eta \in (1, \infty)$ *and* $\theta_0 \in \overset{\circ}{\Theta}$,

$$\sqrt{n}\left(\hat{\alpha}_n - \alpha_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\alpha_0^2\right\}, \qquad \text{as } n \to \infty. \tag{2.3}$$

To obtain the asymptotic distribution of $\hat{\alpha}_n$ in the case $\gamma_0 = 0$, we need an additional assumption on the distribution of η_t^2. Let $Z_t = \alpha_0\eta_t^2$. Note that$\gamma_0 = E\log Z_t = 0$ entails $E(1 + Z_{t-1} + Z_{t-1}Z_{t-2} + \cdots + Z_{t-1}\dots Z_1) \geq t$ by Jensen's inequality. We introduce the assumption

A: when t tends to infinity,

$$E\left(\frac{1}{1 + Z_1 + Z_1Z_2 + \cdots + Z_1\dots Z_{t-1}}\right) = o\left(\frac{1}{\sqrt{t}}\right).$$

Note that **A** is obviously satisfied when $Z_t = 1$ *a.s.*, since the expectation is then equal to $1/t$.

Theorem 2.2. *Suppose that* $\gamma_0 = 0$, $\theta_0 \in \overset{\circ}{\Theta}$, $\kappa_\eta \in (1, \infty)$ *and* **A** *is satisfied. Then the QMLE* $\hat{\alpha}_n$ *is asymptotically normal and its asymptotic distribution is given by (2.3).*

2.2. *Estimator of the asymptotic variance of $\hat{\alpha}_n$ with or without stationarity.* In view of (2.1)-(2.2), when $\gamma_0 < 0$ the asymptotic distribution of the QMLE $\hat{\alpha}_n$ of θ_0 is given by

$$\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\xi\right\}, \qquad \text{as } n \to \infty, \tag{2.4}$$

with

$$\xi = \frac{\mu(0,2)}{\mu(0,2)\mu(2,2) - \mu^2(1,2)}, \qquad \mu(p,q) = E\frac{\epsilon_1^{2p}}{(\omega_0 + \alpha_0\epsilon_1^2)^q}. \tag{2.5}$$

It is obvious to show that the empirical estimator of ξ

$$\hat{\xi}_n = \frac{\hat{\mu}_n(0,2)}{\hat{\mu}_n(0,2)\hat{\mu}_n(2,2) - \hat{\mu}_n^2(1,2)}, \qquad \text{where} \qquad \hat{\mu}_n(p,q) = \frac{1}{n}\sum_{t=1}^{n} \frac{\epsilon_t^{2p}}{(\hat{\omega}_n + \hat{\alpha}_n\epsilon_t^2)^q},$$

is strongly consistent in the stationary case $\gamma_0 < 0$. The following result shows that this estimator is also a consistent estimator of the asymptotic variance of $\hat{\alpha}_n$ in the nonstationary case $\gamma_0 \geq 0$.

THEOREM 2.3. *Assume $\theta_0 \in \Theta$, $\kappa_\eta \in (1, \infty)$ and let $\hat{\kappa}_\eta = n^{-1}\sum_{t=1}^{n} \hat{\eta}_t^4$, where $\hat{\eta}_t = \epsilon_t/\sigma_t(\hat{\theta}_n)$.*

i) When $\gamma_0 < 0$, we have $\hat{\kappa}_\eta \to \kappa_\eta$ and $\hat{\xi}_n \to \xi$ a.s as $n \to \infty$.
ii) When $\gamma_0 > 0$, we have $\hat{\kappa}_\eta \to \kappa_\eta$ and $\hat{\xi}_n \to \alpha_0^2$ a.s.
*iii) When $\gamma_0 = 0$, we have $\hat{\kappa}_\eta \to \kappa_\eta$ and, if **A** is satisfied, $\hat{\xi}_n \to \alpha_0^2$ in probability.*

In any case, $(\hat{\kappa}_\eta - 1)\hat{\xi}_n$ is a consistent estimator of the asymptotic variance of the QMLE of α_0.

The consequence of Theorem 2.3, from a practical point of view, is extremely important. It means that we can get confidence intervals, or tests for α_0 without assuming stationarity/nonstationarity.

3. Testing. Before considering strict stationarity testing, we start with tests on the parameter α.

3.1. *Testing the ARCH coefficient.* Consider a testing problem of the form

$$H_0 : \alpha_0 \leq \alpha^* \qquad \text{against} \qquad H_1 : \alpha_0 > \alpha^*, \tag{3.1}$$

where α^* is a given positive number. A value of particular interest is $\alpha^* = 1$, because $E\epsilon_t^2 < \infty$ if and only if $\alpha_0 < 1$. Note however that we do not impose any constraint on α^* so that some values of $\alpha_0 \leq \alpha^*$ may correspond to nonstationary ARCH models. A direct consequence of Theorems 2.1-2.2-2.3 is the following result, in which Φ denotes the $\mathcal{N}(0,1)$ cumulative distribution function. Let $\underline{\alpha} \in (0,1)$.

Corollary 3.1. *Assume that $\theta_0 \in \overset{\circ}{\Theta}$ and the assumptions of Theorem 2.3 hold. For the testing problem (3.1), the test defined by the critical region*

$$\mathrm{C}^{\alpha^*} = \left\{ T_n^{\alpha^*} := \frac{\sqrt{n}(\hat{\alpha}_n - \alpha^*)}{\sqrt{(\hat{\kappa}_\eta - 1)\hat{\xi}_n}} > \Phi^{-1}(1-\underline{\alpha}) \right\} \tag{3.2}$$

has the asymptotic significance level $\underline{\alpha}$ and is consistent.

Hence, assumption H_0 can be tested without knowing if the observations are generated by a stationary or an explosive ARCH. However, it is of interest to test if a given series is stationary or not. This cannot be done by testing an assumption of the form (3.1).

3.2. *Strict stationarity testing.* Consider the strict stationarity testing problems

$$H_0 : \gamma_0 < 0 \qquad \text{against} \qquad H_1 : \gamma_0 \geq 0, \tag{3.3}$$

and

$$H_0 : \gamma_0 \geq 0 \qquad \text{against} \qquad H_1 : \gamma_0 < 0, \tag{3.4}$$

where $\gamma_0 = E \log\left(\alpha_0 \eta_1^2\right)$. These hypotheses are not of the form (3.1) because γ_0 not only depends on α_0, but also on the unknown moment $\zeta := E \log \eta_t^2 \in \mathbb{R} \cup \{-\infty\}$. Let $\hat{\zeta}_n = n^{-1} \sum_{t=1}^n \log \hat{\eta}_t^2$. The following result gives the asymptotic joint distribution of $\hat{\zeta}_n$ and $\hat{\alpha}_n$, and the asymptotic distribution of a consistent estimator of γ_0, under either the stationarity or the nonstationarity conditions.

Theorem 3.1. *Assume that $E\eta_1^4 \in (1,\infty)$, $E\left|\log \eta_1^2\right|^2 < \infty$ and $\theta_0 \in \overset{\circ}{\Theta}$.*

i) If the stationarity condition $\gamma_0 < 0$ holds, as $n \to \infty$,

$$\sqrt{n} \begin{pmatrix} \hat{\zeta}_n - \zeta \\ \hat{\theta}_n - \theta_0 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{ 0, \Sigma := \begin{pmatrix} \sigma_u^2 + \sigma_v^2 + 2\sigma_{uv} & -(\sigma_v^2 + \sigma_{uv})\theta_0' \\ -(\sigma_v^2 + \sigma_{uv})\theta_0 & \sigma_v^2 J^{-1} \end{pmatrix} \right\}, \tag{3.5}$$

where $v_t = 1 - \eta_t^2$, $\sigma_v^2 = Ev_t^2$, $u_t = \log \eta_t^2 - \zeta$, $\sigma_u^2 = Eu_t^2$, $\sigma_{uv} = Cov(u_t, v_t)$ *and* J *is given by (2.2). Moreover*

$$\hat{\gamma}_n := \hat{\zeta}_n + \log \hat{\alpha}_n \to \gamma_0 \quad \textit{in probability as } n \to \infty \tag{3.6}$$

and

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}\left(0, \sigma_u^2 + \sigma_v^2 \left\{\frac{\xi}{\alpha_0^2} - 1\right\}\right) \quad \textit{as } n \to \infty. \tag{3.7}$$

ii) If $\gamma_0 > 0$, *or if* $\gamma_0 = 0$ *and* **A** *holds, then*

$$\sqrt{n}\begin{pmatrix} \hat{\zeta}_n - \zeta \\ \hat{\alpha}_n - \alpha_0 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{0, \begin{pmatrix} \sigma_u^2 + \sigma_v^2 + 2\sigma_{uv} & -(\sigma_v^2 + \sigma_{uv})\alpha_0 \\ -(\sigma_v^2 + \sigma_{uv})\alpha_0 & \sigma_v^2\alpha_0^2 \end{pmatrix}\right\}. \tag{3.8}$$

Moreover (3.6) holds and we have

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}\left(0, \sigma_u^2\right) \quad \textit{as } n \to \infty. \tag{3.9}$$

It is interesting to note that the asymptotic distribution of $\hat{\zeta}_n$ is the same in the cases $\gamma_0 < 0$ and $\gamma_0 \geq 0$ and that this distribution is independent of θ_0.

Let $\hat{\sigma}_u^2 = n^{-1}\sum_{t=1}^n \left(\log \hat{\eta}_t^2\right)^2 - \hat{\zeta}_n^2$.

COROLLARY 3.2. *Let the assumptions of Theorem 3.1 hold. For the testing problem (3.3), the test defined by the critical region*

$$\mathrm{C}^{\mathrm{ST}} = \left\{T_n := \sqrt{n}\frac{\hat{\gamma}_n}{\hat{\sigma}_u} > \Phi^{-1}(1 - \underline{\alpha})\right\} \tag{3.10}$$

has its asymptotic significance level bounded by $\underline{\alpha}$, *has the asymptotic probability of rejection* $\underline{\alpha}$ *under* $\gamma_0 = 0$, *and is consistent for all* $\gamma_0 > 0$.

For the testing problem (3.4), the test defined by the critical region

$$\mathrm{C}^{\mathrm{NS}} = \left\{T_n < \Phi^{-1}(\underline{\alpha})\right\} \tag{3.11}$$

has its asymptotic significance level bounded by $\underline{\alpha}$, *has the asymptotic probability of rejection* $\underline{\alpha}$ *under* $\gamma_0 = 0$, *and is consistent for all* $\gamma_0 < 0$.

4. Asymptotic local powers. The section investigates the asymptotic behavior under local alternatives of the tests (3.2) on α_0 and of the strict stationarity tests (3.10) and (3.11). We first establish the LAN of the ARCH model without imposing any stationarity constraint. This LAN property will be used to derive the asymptotic properties of our tests, but the result is of independent interest (see van der Vaart (1998) for a general reference on LAN and its applications, and see Drost and Klaassen (1997), Drost, Klaassen and Werker (1997) and Ling and McAleer (2003) for applications to GARCH and other stationary processes).

4.1. *LAN without stationarity constraint.* Assume that η_t has a density f with third-order derivatives, that

$$\lim_{|y|\to\infty} y^2 f'(y) = 0, \tag{4.1}$$

and that for some positive constants K and δ

$$|y|\left|\frac{f'}{f}(y)\right| + y^2\left|\left(\frac{f'}{f}\right)'(y)\right| + y^2\left|\left(\frac{f'}{f}\right)''(y)\right| \leq K\left(1+|y|^{\delta}\right), \tag{4.2}$$

$$E\,|\eta_1|^{2\delta} < \infty. \tag{4.3}$$

These regularity conditions are satisfied for numerous distributions, in particular for the gaussian distribution with $\delta = 2$, and entail the existence of the Fisher information for scale

$$\iota_f = \int \left\{1 + yf'(y)/f(y)\right\}^2 f(y)dy < \infty.$$

Given the initial value ϵ_0, the density of the observations $(\epsilon_1, \dots, \epsilon_n)$ satisfying (1.1) is given by $L_{n,f}(\theta_0) = \prod_{t=1}^n \sigma_t^{-1}(\theta_0) f\left\{\sigma_t^{-1}(\theta_0)\epsilon_t\right\}$. Around $\theta_0 \in \overset{\circ}{\Theta}$, let a sequence of local parameters of the form $\theta_n = \theta_0 + \tau_n/\sqrt{n}$, where (τ_n) is a bounded sequence of $\mathbb{R}^2$. Without loss of generality, assume that n is sufficiently large so that $\theta_n \in \Theta$. Under the strict stationarity condition $\gamma_0 < 0$, Drost and Klaassen (1997) showed that the log-likelihood ratio $\Lambda_{n,f}(\theta_n, \theta_0) = \log L_{n,f}(\theta_n)/L_{n,f}(\theta_0)$ satisfies the LAN property

$$\Lambda_{n,f}(\theta_n, \theta_0) = \tau_n' S_{n,f}(\theta_0) - \frac{1}{2}\tau_n'\mathfrak{I}_f\tau_n + o_{P_{\theta_0}}(1), \quad S_{n,f}(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \mathfrak{I}_f\right\} \tag{4.4}$$

under P_{θ_0} as $n \to \infty$. The following proposition shows that (4.4) holds regardless of γ_0.

Proposition 4.1. *When $\theta_0 \in \overset{\circ}{\Theta}$, under (4.1)-(4.3) we have the LAN property (4.4). When $\gamma_0 \geq 0$, the Fisher information is the degenerate matrix*

$$\mathfrak{I}_f = \frac{\iota_f}{4}\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-2} \end{pmatrix} \tag{4.5}$$

4.2. *Local asymptotic powers of the tests.* Because the information matrix (4.5) is singular, when $\gamma_0 \geq 0$ the LAN property does not entail the convolution theorem of Hájek (see Theorem 2.2 of Drost and Klaassen (1997)). The LAN property, with the help of Le Cam's third lemma, allows however to easily compute local asymptotic powers of tests. In view of Corollary 3.2,

$$\lim_{n\to\infty} P_{\theta_0}\left(\mathrm{C}^{\mathrm{ST}}\right) = \lim_{n\to\infty} P_{\theta_0}\left(\mathrm{C}^{\mathrm{NS}}\right) = \underline{\alpha},$$

when $\theta_0 = (\omega_0, \alpha_0)'$ is such that $\alpha_0 = \exp(-E\log\eta_t^2)$. Denote by $P_{n,\tau}$, where $\tau = (\tau_1, \tau_2)'$, the distribution of the observations $(\epsilon_1, \dots, \epsilon_n)$ when the parameter is of the form $\left(\omega_0 + \tau_1/\sqrt{n}, \exp(-E\log\eta_t^2) + \tau_2/\sqrt{n}\right)'$. We should use the notation $(\epsilon_{1,n}, \dots, \epsilon_{n,n})$ instead of $(\epsilon_1, \dots, \epsilon_n)$ because the parameter varies with n, but we will avoid this heavy notation. Local alternatives for the C^{ST}-test (resp. the C^{NS}-test) are obtained for $\tau_2 > 0$ (resp. $\tau_2 < 0$).

Proposition 4.2. *Under the assumptions of Theorem 3.1 and Proposition 4.1, the local asymptotic powers of the strict stationarity tests (3.10) and (3.11) are given by*

$$\lim_{n\to\infty} P_{n,\tau}\left(\mathrm{C}^{\mathrm{ST}}\right) = \Phi\left\{\frac{\tau_2}{\alpha_0\sigma_u} - \Phi^{-1}(1-\underline{\alpha})\right\} \tag{4.6}$$

and

$$\lim_{n\to\infty} P_{n,\tau}\left(\mathrm{C}^{\mathrm{NS}}\right) = \Phi\left\{\Phi^{-1}(\underline{\alpha}) - \frac{\tau_2}{\alpha_0\sigma_u}\right\}.$$

We now compute the local asymptotic power of the test defined by (3.2). We thus consider a sequence of local parameters of the form $\theta_n^{\alpha^*} = (\omega_0, \alpha^*)' + \tau/\sqrt{n}$ where $\tau = (\tau_1, \tau_2)'$ with $\tau_2 > 0$. We denote by $P_{n,\tau}^{\alpha^*}$ the distribution of the observations under the assumption that the ARCH(1) parameter is $\theta_n^{\alpha^*}$.

Proposition 4.3. *Let the assumptions of Proposition 4.1 and Theorem 2.3 be satisfied. For testing (3.1), the test defined by the rejection region (3.2) has the local asymptotic power*

$$\lim_{n\to\infty} P_{n,\tau}^{\alpha^*}\left(\mathrm{C}^{\alpha^*}\right) = \Phi\left\{\frac{\tau_2}{\sqrt{(\kappa_\eta - 1)\xi_0}} - \Phi^{-1}(1-\underline{\alpha})\right\}, \tag{4.7}$$

where $\xi_0 = \alpha^{*2}$ when $E\log\alpha^*\eta_1^2 \geq 0$ and $\xi_0 = \xi$ defined by (2.5) when $E\log\alpha^*\eta_1^2 < 0$.

4.3. *Optimality issues.* Let T_2 be a subset of $\mathbb{R}$ containing 0. When $\gamma_0 \geq 0$, the relations (4.4)-(4.5) imply that the limiting distribution of $\Lambda_{n,f}(\theta_0 + \tau/\sqrt{n}, \theta_0)$ is that of the log-likelihood ratio in the statistical model $\mathcal{N}(\tau_2, 4\alpha_0^2/\iota_f)$ of parameter τ_2. In other words, the so-called local experiments $\{L_{n,f}(\theta_0 + (0,\tau_2)'/\sqrt{n}), \tau_2 \in T_2\}$ converge to the gaussian experiment $\{\mathcal{N}(\tau_2, 4\alpha_0^2/\iota_f), \tau_2 \in T_2\}$ (see van der Vaart (1998) for details about the notion of statistical experiments). Testing $\alpha_0 \leq \exp(-E \log \eta_1^2)$ against $\alpha_0 > \exp(-E \log \eta_1^2)$ corresponds to testing $\tau_2 \leq 0$ against $\tau_2 > 0$ in the limiting experiment. The uniformly most powerful test based on $X \sim \mathcal{N}(\tau_2, 4\alpha_0^2/\iota_f)$ is the Neyman-Pearson test of rejection region $C = \left\{X/\sqrt{4\alpha_0^2/\iota_f} > \Phi^{-1}(1-\underline{\alpha})\right\}$. This optimal test has the power

$$P_{\tau_2}(C) = \Phi\left(\frac{\tau_2}{\sqrt{4\alpha_0^2/\iota_f}} - \Phi^{-1}(1-\underline{\alpha})\right). \tag{4.8}$$

A test of (3.10) whose level and power jointly converge to $\underline{\alpha}$ and to the bound in (4.8), respectively, will be called asymptotically optimal. A similar result holds for the dual testing problem (3.11). For the testing problem (3.1), the optimal local asymptotic power is

$$\Phi\left(\frac{\tau_2}{\sqrt{4\xi_0/\iota_f}} - \Phi^{-1}(1-\underline{\alpha})\right), \tag{4.9}$$

where ξ_0 is defined in Proposition 4.3.

Proposition 4.4. *Under the assumptions of Proposition 4.2, the strict stationarity tests (3.10) (and/or (3.11)) is asymptotically optimal if and only if*

$$f(y) = \frac{1}{2\sqrt{|\delta|\pi}e^{-\delta/4}} e^{\frac{(\log|y|)^2}{\delta}} y^{-2}, \quad \delta < 0. \tag{4.10}$$

The test (3.2) is optimal for the testing problem (3.1) if and only if

$$f(y) = \frac{a^a}{\Gamma(a)} e^{-ay^2} |y|^{2a-1}, \quad a > 0, \quad \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt. \tag{4.11}$$

Figure 1 displays the densities (4.11) and (4.10) for different values of a and δ. Note that the gaussian density is obtained in (4.11) for $a = 1/2$. The result was expected because the C^{α^*}-test is based on the QMLE of α_0, and the QMLE is efficient in the gaussian case. Note however that the

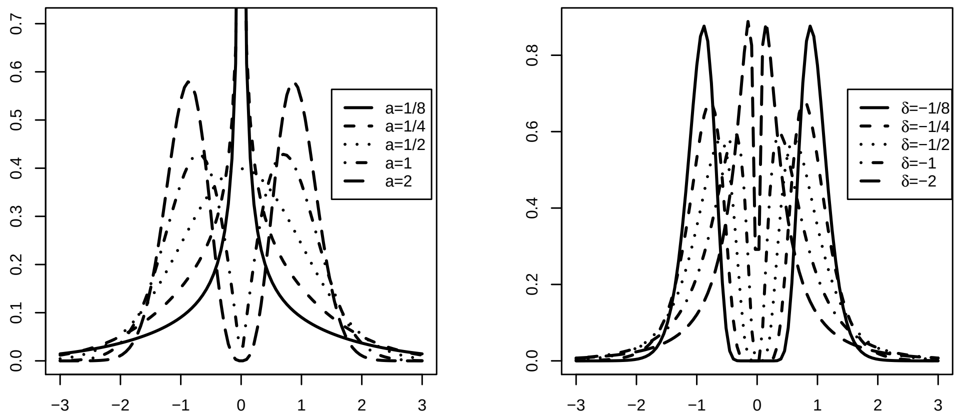

FIG 1. *Densities (4.11) of η_t for which the test (3.2) on α_0 is asymptotically optimal (left panel) and densities (4.10) for which the strict stationarity tests (3.10) and (3.11) are asymptotically optimal (right panel).*

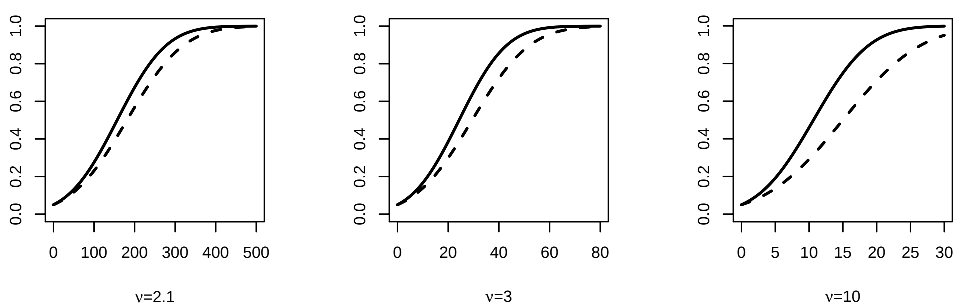

FIG 2. *Optimal asymptotic power (in full line) and local asymptotic power of the strict stationarity test (3.10) (in dotted line) when η_t follows a standardized Student distribution with ν degrees of freedom. .*

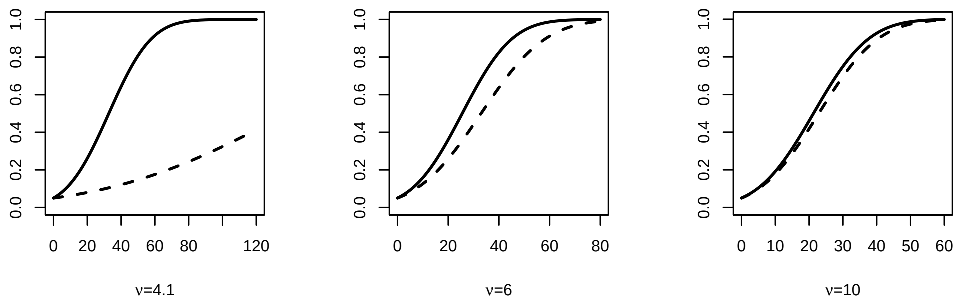

FIG 3. *Optimal asymptotic power (in full line) and local asymptotic power of the test (3.1) (in dotted line) for testing $\alpha_0 < \alpha^*$ where $\alpha^* = 2\exp(-\zeta)$ when η_t follows a standardized Student distribution with ν degrees of freedom.*

C^{α^*}-test is also asymptotically optimal when η_t follows some non gaussian distributions. The strict stationarity tests are not optimal in the gaussian case. For densities which do not belong to the class (4.10), there is a price to pay for the estimation of ζ and/or for using an estimator of α_0 which is asymptotically less accurate than the MLE. This point is illustrated by Figures 2–3, in which the local asymptotic powers of the different tests (in dotted lines) are compared to the optimal asymptotic powers given by (4.8) and (4.9). In these two figures, the noise η_t is assumed to satisfy a Student distribution with $\nu > 2$ degrees of freedom, standardized in such a way that $E\eta_t^2 = 1$. Figure 3 considers tests of the null assumption $\alpha_0 \leq \alpha^*$, where $\alpha^* = 2\exp(-E\log\eta_t^2)$ is such that $\gamma_0 = 0$ for this particular distribution. It can be seen, in Figure 3, that the local asymptotic power is far from the optimal power when ν is small, but the discrepancy decreases as ν increases. By contrast, the discrepancy increases with ν in Figure 2. This is not surprising since the normal distribution belongs to the class defined by (4.11), but not to that defined by (4.10).

5. Testing non stationarity in non linear GARCH. In this section we study the behaviour of the stationarity tests of Section 3.2 when the data are generated by the following GARCH-type model:

$$\left\{ \begin{array}{lcl} \epsilon_t & = & \sqrt{h_t}\eta_t, \quad t = 1, 2, \dots \\ h_t & = & \omega(\eta_{t-1}) + a(\eta_{t-1})h_{t-1} \end{array} \right. \tag{5.1}$$

with an initial value h_0, under the same assumptions on (η_t) as in Model (1.1). In this model, $\omega : \mathbb{R} \to [\underline{\omega}, +\infty)$, for some $\underline{\omega} > 0$, and $a : \mathbb{R} \to \mathbb{R}^+$. This model belongs to the so-called class of augmented GARCH models (see Hörmann, 2008) and encompasses many classes of GARCH(1,1) models introduced in the literature: for instance, with constant $\omega(\cdot)$, the standard GARCH(1,1) when $a(x) = \alpha_0 x^2 + \beta_0$; the GJR model when $a(x) = \alpha_1(\max\{x,0\})^2 + \alpha_2(\min\{x,0\})^2 + \beta_0$. It can be shown that, if $E\log^+ a(\eta_t) < \infty$,

$$\Gamma := E\log a(\eta_t) < 0 \tag{5.2}$$

is a necessary and sufficient condition for the strict stationarity of this model (see e.g. Francq and Zakoïan, 2006a). Our aim is to test strict stationarity, without estimating the non parametric Model (5.1). We shall see that, surprisingly, the tests developed for the standard ARCH(1) model still work in this framework. Recall that the tests are founded on the statistics

$$\hat{\gamma}_n = \log\hat{\alpha}_n + \frac{1}{n}\sum_{t=1}^n \log\hat{\eta}_t^2, \quad \hat{\sigma}_u^2 = \frac{1}{n}\sum_{t=1}^n \left(\log\hat{\eta}_t^2\right)^2 - \left(\frac{1}{n}\sum_{t=1}^n \log\hat{\eta}_t^2\right)^2$$

where $\hat{\alpha}_n$ denotes the QMLE of the ARCH coefficient in an ARCH(1) model and the squared rescaled residuals are given by

$$\hat{\eta}_t^2 = \frac{\epsilon_t^2}{\hat{\omega}_n + \hat{\alpha}_n\epsilon_{t-1}^2}, \quad t = 1,\dots,n.$$

Proposition 5.1. *Let $\epsilon_1,\dots,\epsilon_n$ denote observations from Model (5.1). Assume $0 < E|\log\eta_1^2|^2 < \infty$, $E|\log a(\eta_1)|^2 < \infty$, $E\left\{a(\eta_1)/\eta_1^2\right\} < \infty$, and $E|\omega(\eta_1)|^s < \infty$ for some $s > 0$.*
If $\Gamma > 0$ then

$$\hat{\gamma}_n \to \Gamma, \quad \textit{and} \quad \hat{\sigma}_u^2 \to Var\log\left\{\eta_1^2\frac{a(\eta_0)}{\eta_0^2}\right\} > 0, \quad a.s.$$

If $\Gamma < 0$ then, under regularity conditions implying the strong consistency of $\hat{\theta}_n$ to the unique pseudo-true value

$$(\omega^*,\alpha^*)' = \arg\min_{\theta\in\Theta} E\left\{\frac{\epsilon_t^2}{\omega + \alpha\epsilon_{t-1}^2} + \log\left(\omega + \alpha\epsilon_{t-1}^2\right)\right\}$$

and if $Var\log\epsilon_t^2 < \infty$, we have, for some Γ^*,

$$\hat{\gamma}_n \to \Gamma^* < 0, \quad \textit{and} \quad \hat{\sigma}_u^2 \to Var\log\left\{\frac{\epsilon_t^2}{\omega^* + \alpha^*\epsilon_{t-1}^2}\right\} > 0, \quad a.s.$$

Thus, the (non)stationarity tests developed in the ARCH(1) case lead, asymptotically, to the right decision, even if the ARCH(1) model is misspecified (at least for the augmented GARCH(1,1), except in the limit case where $\Gamma = 0$). More precisely, we have the following result.

Corollary 5.1. *Let the assumptions of Proposition 5.1 hold.*
If $\Gamma > 0$ *then*

$$P(\mathrm{C}^{\mathrm{NS}}) \to 0 \quad \textit{and} \quad P(\mathrm{C}^{\mathrm{ST}}) \to 1$$

where C^{ST} *and* C^{ST} *are defined in Corollary 3.2.*
If $\Gamma < 0$ *then*

$$P(\mathrm{C}^{\mathrm{ST}}) \to 0 \quad \textit{and} \quad P(\mathrm{C}^{\mathrm{NS}}) \to 1.$$

6. Numerical illustrations. Before illustrating our asymptotic results for the tests, we study the behaviour of the QMLE in finite samples.

6.1. *Inconsistency of* $\hat{\omega}_n$ *in the non stationary case.* The asymptotic behavior of the score leads us to think that the QMLE of ω_0 is inconsistent without the strict stationarity condition. A detailed discussion is provided in Section 7.3. Figure 4 presents some numerical evidence on the performance of the QMLE in finite samples through a simulation study. In all experiments, we use the sample size $n = 200$ and $n = 4,000$ with 100 replications. The data of the top panel are generated from the second-order stationary ARCH(1) model (1.1) with the true parameter $\theta_0 = (1, 0.95)'$. The data of the middle panel are generated from the strictly stationary ARCH(1) model with $\theta_0 = (1, 1.5)'$ and infinite variance. In those two panels the results are very similar, confirming that the second-order stationarity condition is not necessary for the use of the QMLE. The bottom panel, obtained for the explosive ARCH(1) model with $\theta_0 = (1, 4)'$, confirms the asymptotic results for the QMLE of α_0. It also illustrates the impossibility of estimating the parameter ω_0 with a reasonable accuracy under the nonstationarity condition (1.5). The results concerning ω_0 even worsen when the sample size increases.

6.2. *Finite sample properties of the tests.*

6.2.1. *On simulated data.* To assess the performance of the tests developed in Section 3, we simulated $N = 1,\ 000$ independent trajectories of size $n = 100$, $n = 500$ and $n = 1,\ 000$ of an ARCH(1) model. We used different values of α_0 and a double Gamma distribution for η_t, with shape parameter $k = 3$ and scale parameter $s = 1/2\sqrt{3}$. The density of that distribution is $f(\eta) = \eta^{k-1}/\{2(k-1)!\}e^{-|\eta|/s}$, where k and s are such that $E\eta_0^2 = 1$ and such that the assumptions of Proposition 5.1 are satisfied with a standard GARCH(1,1) volatility.

The results concerning the test (3.2) on α_0 are presented in Tables 1-2. With the density f, we have $\gamma_0 = 0$ for $\alpha_0 = 1.895$. It has to be noted that

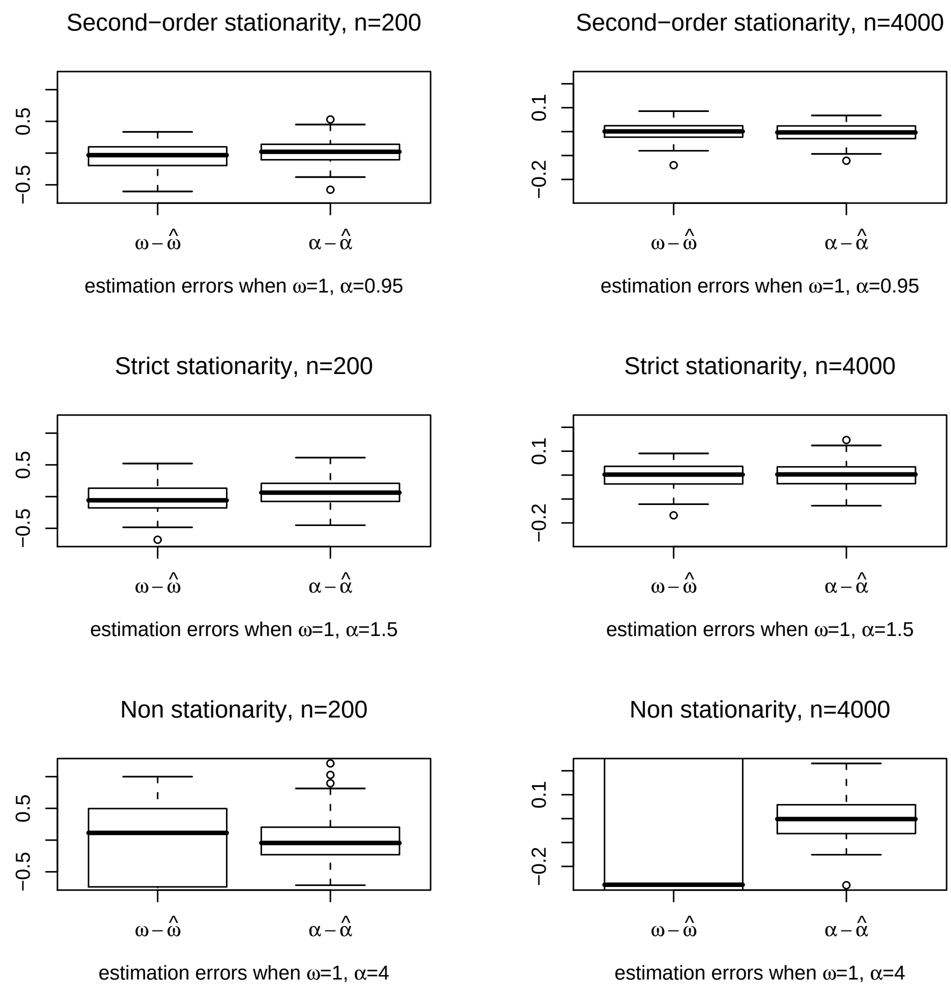

FIG 4. *Boxplots of estimation errors for the QMLE of the parameters* ω_0 *and* α_0 *of an ARCH(1), with* $\eta_t \sim \mathcal{N}(0,1)$.

Table 1

Relative frequency of rejection (in %) for the test (3.2) of the null hypothesis $H_0 : \alpha_0 \leq 1$ against $H_1 : \alpha_0 > 1$ at the nominal level $\underline{\alpha} = 5\%$ when the errors follow a double Gamma distribution.

α_0	0.8	0.9	0.95	1	1.1	1.2	1.3
$n = 100$	0.0	0.4	1.2	2.3	7.7	17.8	32.9
$n = 500$	0.0	0.2	0.8	2.4	22.8	66.9	90.8
$n = 1,000$	0.0	0.0	0.7	4.9	38.9	88.6	99.4

Table 2

As Table 1, but for testing the null hypothesis $H_0 : \alpha_0 \leq 3$ against $H_1 : \alpha_0 > 3$.

α_0	2.8	2.9	2.95	3	3.1	3.2	3.3
$n = 100$	0.4	1.3	1.4	2.5	4.0	8.9	12.5
$n = 500$	0.1	0.5	1.9	3.5	10.0	28.5	47.9
$n = 1,000$	0.0	0.4	1.0	5.2	19.7	46.6	78.6

the test behaves similarly when the value tested corresponds to a stationary solution (Table 1) or to a non stationary process (Table 2).

We now illustrate the behavior of the strict stationarity tests (3.10) and (3.11), through simulations of ARCH(1) models with values of α_0 corresponding to $\gamma_0 < 0$ ($\alpha_0 \in \{1.6, 1.7, 1.8\}$), $\gamma_0 = 0$ ($\alpha_0 = 1.895$) and $\gamma_0 > 0$ ($\alpha_0 \in \{2, 2.1, 2.2\}$). Tables 3-4 show that, as expected, the frequency of rejection of the C^{ST}-test increases with γ_0 while, obviously, that of the C^{NS}-test decreases. The rejection frequencies of the two tests approach the nominal level when $\gamma_0 = 0$ and n increases.

Now consider testing strict stationarity in a GARCH(1,1) model using the tests developed for the ARCH(1). Tables 5-6 confirm the theoretical result of Section 5. More precisely, for n sufficiently large, the tests give the right conclusion when $\Gamma < 0$ and $\Gamma > 0$. Note that when $\Gamma = 0$ the rejection frequencies are far from the nominal 5% level corresponding to an ARCH(1). This is not surprising since, except in the ARCH(1) case, the

Table 3

Relative frequency of rejection of the stationarity hypothesis $H_0 : \gamma_0 < 0$ of the test (3.10) at the nominal level $\underline{\alpha} = 5\%$ in the ARCH(1) case. The parameter $\alpha_0 = 1.895$ corresponds to $\gamma_0 = 0$.

α_0	1.6	1.7	1.8	1.895	2	2.1	2.2
$n = 100$	0.2	1.4	2.9	6.4	11.8	21.3	33.8
$n = 500$	0.0	0.0	0.6	4.7	25.0	57.4	83.4
$n = 1,000$	0.0	0.0	0.3	5.9	38.2	82.2	98.3

TABLE 4
As Table 3, but for testing the nonstationarity hypothesis $H_0 : \gamma_0 \geq 0$ with the test (3.11).

α_0	1.6	1.7	1.8	1.895	2	2.1	2.2
$n = 100$	42.1	28.4	17.1	11.6	5.5	2.1	0.8
$n = 500$	90.0	61.4	23.8	6.7	0.9	0.0	0.0
$n = 1,000$	99.3	86.9	39.8	5.5	0.3	0.0	0.0

TABLE 5
As Table 3, but for standard GARCH(1,1) models with $\beta_0 = 0.8$. The parameter $\alpha_0 = 0.226$ corresponds to $\Gamma = 0$.

α_0	0.1	0.15	0.2	0.226	0.35	0.4	0.5
$n = 100$	0.0	0.0	0.0	0.0	0.0	0.0	0.6
$n = 500$	0.0	0.0	0.0	0.0	1.6	49.9	96.2
$n = 1,000$	0.0	0.0	0.0	0.0	55.1	96.0	99.8

asymptotic relative frequencies of rejection are unknown under $\Gamma = 0$.

6.2.2. *On real data.* The strict stationarity tests were then applied to the daily returns of 11 major stock market indices. We considered the CAC, DAX, DJA, DJI, DJT, DJU, FTSE, Nasdaq,[3] Nikkei, SMI and SP500, from January 2, 1990, to January 22, 2009, except for the indices for which such historical data do not exist. Table 7 displays the test statistics T_n computed on each series. Note that, as $n \to \infty$,

$$T_n = \sqrt{n}\frac{\hat{\gamma}_n - \gamma_0}{\hat{\sigma}_u} + \sqrt{n}\frac{\gamma_0}{\hat{\sigma}_u} \to -\infty$$

in probability when $\gamma_0 < 0$, and $T_n \to +\infty$ in probability when $\gamma_0 > 0$. Because the values of T_n given in Table 7 are very small, a nonstationary augmented GARCH(1,1) model is not plausible, for any of these series.

[3] Since the Nasdaq index level was halved on January 3, 1994, one outlier has been eliminated for this series.

TABLE 6
As Table 4, but for testing the nonstationarity hypothesis $H_0 : \Gamma \geq 0$ with the test (3.11).

α_0	0.1	0.15	0.2	0.226	0.35	0.4	0.5
$n = 100$	100.0	100.0	98.5	95.3	26.6	10.7	2.9
$n = 500$	100.0	100.0	100.0	93.9	1.1	0.2	0.1
$n = 1,000$	100.0	100.0	100.0	89.1	0.5	0.1	0.0

Table 7

Test statistic T_n of the strict stationarity tests (3.10) and (3.11). The test statistic is the realization of a random variable which is asymptotically $\mathcal{N}(0,1)$ distributed when $\gamma_0 = 0$, tends to $-\infty$ under the strict stationarity hypothesis $\gamma_0 < 0$, tends to $+\infty$ when $\gamma_0 > 0$.

CAC	DAX	DJA	DJI	DJT	DJU	FTSE	Nasdaq	Nikkei	SMI	SP500
-86.1	-79	-80.8	-78.7	-87.2	-69.3	-75.3	-81.4	-86.7	-71.5	-80.8

7. Proofs and complementary results.

7.1. *Asymptotic behaviors of* (h_t). As noted in the introduction, when $\gamma_0 \neq 0$ the asymptotic behavior of the sequences (h_t) (defined by (1.1)) and (σ_t^2) (defined by (1.4)) is the same and is easily obtained by the Cauchy rule. When $\gamma_0 = 0$ the asymptotic behavior of σ_t^2 can be obtained by the Chung-Fuchs theorem. The behavior of h_t is different in this case and is described in the result below.

Proposition 7.1. *For the ARCH(1) model (1.1), the following properties hold:*

i) When $\gamma_0 > 0$, $\quad h_t \to \infty$ a.s. at an exponential rate:

$$\text{for any } \rho \in (e^{-\gamma_0}, 1), \quad \rho^t h_t \to \infty \quad \text{and} \quad \rho^t \epsilon_t^2 \to \infty \quad \text{a.s. as } t \to \infty.$$

ii) (Klüppelberg, Lindner and Maller (2004))
When $\gamma_0 = 0$, $\quad h_t \to \infty \quad$ and $\quad \epsilon_t^2 \to \infty$ in probability.

iii) Let ψ be a decreasing bijection from $(0,\infty)$ to $(0,\infty)$ such that $E\psi(\epsilon_1^2) < \infty$. When $\gamma_0 = 0$,

$$\frac{1}{n}\sum_{t=1}^{n} \psi(\epsilon_t^2) \to 0 \quad \text{and} \quad \frac{1}{n}\sum_{t=1}^{n} \psi(h_t) \to 0 \quad \text{in } L^1 \text{ as } n \to \infty. \tag{7.1}$$

Proof. To prove *i)* we note that

$$h_t = \omega_0 + \alpha_0 \eta_{t-1}^2 h_{t-1} = \omega_0 \left\{ 1 + \sum_{i=1}^{t-1} \alpha_0^i \eta_{t-1}^2 \dots \eta_{t-i}^2 \right\} + \alpha_0^t \eta_{t-1}^2 \dots \eta_1^2 \epsilon_0^2$$

$$\geq \omega_0 \prod_{i=1}^{t-1} \alpha_0 \eta_i^2, \tag{7.2}$$

Thus, for any constant $\rho \in (e^{-\gamma_0}, 1)$, we have

$$\begin{aligned}\liminf_{t\to\infty} \frac{1}{t}\log \rho^t h_t &\geq \lim_{t\to\infty}\frac{1}{t}\left\{\log\rho\omega_0 + \sum_{i=1}^{t-1}\log\rho\alpha_0\eta_i^2\right\}\\ &= E\log\rho\alpha_0\eta_1^2 = \log\rho + \gamma_0 > 0.\end{aligned}$$

It follows that $\log\rho^t h_t$, and hence $\rho^t h_t$, tends to $+\infty$ a.s as $n\to\infty$. For any real-valued function f, let $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$, so that $f(x) = f^+(x) - f^-(x)$. Since $E\log^+\eta_1^2 \leq E\eta_1^2 = 1$, we have $E|\log\eta_1^2| = \infty$ if and only if $E\log\eta_1^2 = -\infty$. Thus $\gamma_0 > 0$ implies $E|\log\eta_1^2| < \infty$, which entails that $\log\eta_t^2/t \to 0$ a.s. as $t\to\infty$. Therefore, $\liminf_{t\to\infty} t^{-1}\log\rho^t\eta_t^2 h_t \geq E\log\rho\alpha_0\eta_1^2 > 0$, and $\rho^t\epsilon_t^2 = \rho^t\eta_t^2 h_t \to +\infty$ a.s. by already given arguments.

The proof of *ii)* follows from Klüppelberg et al. (2004). Their condition $E\left|\log\lambda\varepsilon_1^2\right| < \infty$ becomes in our notations $E\left|\log\alpha_0\eta_1^2\right| < \infty$, and this condition is satisfied because $E(\log\alpha_0\eta_1^2)^+ - E(\log\alpha_0\eta_1^2)^- = \gamma_0 = 0$ and $E(\log\alpha_0\eta_1^2)^+ \leq \alpha_0$.

To prove *iii)* note that $\gamma_0 \geq 0$ implies $\alpha_0 \geq \exp\{-\log E\eta_1^2\} = 1$, by Jensen's inequality. Thus $h_t > \epsilon_{t-1}^2$ and $\psi(h_t) < \psi(\epsilon_{t-1}^2)$. Therefore, the second convergence in (7.1) will follow from first convergence. It suffices to consider the case $\epsilon_0 = 0$ and to show that $E\psi(\epsilon_t^2) \to 0$ as $t\to\infty$. Note that, even if ϵ_t^2 does not increase with probability one, ϵ_{t+1}^2 is stochastically greater than ϵ_t^2 because

$$\begin{aligned}\epsilon_{t+1}^2 &= (\omega_0 + \omega_0\alpha_0\eta_t^2 + \cdots + \omega_0\alpha_0^{t-1}\eta_t^2\cdots\eta_2^2 + \omega_0\alpha_0^t\eta_t^2\cdots\eta_1^2)\eta_{t+1}^2\\ &\geq (\omega_0 + \omega_0\alpha_0\eta_t^2 + \cdots + \omega_0\alpha_0^t\eta_t^2\cdots\eta_2^2)\eta_{t+1}^2\\ &\overset{d}{=} \epsilon_t^2\end{aligned}$$

where $\overset{d}{=}$ stands for equality in distribution. The dominated convergence theorem and *i)-ii)* then entail

$$E\psi(\epsilon_t^2) = \int_0^\infty P\left\{\epsilon_t^2 < \psi^{-1}(u)\right\}du \to \int_0^\infty \lim_{t\to\infty}\downarrow P\left\{\epsilon_t^2 < \psi^{-1}(u)\right\}du = 0,$$

which completes the proof. □

7.2. *Asymptotic normality of the unrestricted QMLE of α_0.*

LEMMA 7.1. *When $\gamma_0 > 0$, we have*

$$\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left| \frac{\partial}{\partial\omega} \ell_t(\theta) \right| < \infty \quad a.s., \tag{7.3}$$

$$\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left\| \frac{\partial^2}{\partial\omega\partial\theta} \ell_t(\theta) \right\| < \infty \quad a.s., \tag{7.4}$$

$$\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial\alpha^2} \ell_t(\omega, \alpha_0) - \frac{1}{\alpha_0^2} \right| = o(1) \quad a.s., \tag{7.5}$$

$$\frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta} \left| \frac{\partial^3}{\partial\alpha^3} \ell_t(\theta) \right| = O(1) \quad a.s., \tag{7.6}$$

When $\gamma_0 = 0$,

$$\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial\alpha^2} \ell_t(\omega, \alpha_0) - \frac{1}{\alpha_0^2} \right| = o(1) \quad \text{in probability}, \tag{7.7}$$

$$\frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta} \left| \frac{\partial^3}{\partial\alpha^3} \ell_t(\theta) \right| = O(1) \quad \text{in probability}. \tag{7.8}$$

Proof. Using Proposition 7.1, there exist a real random variable K and a constant $\rho \in (e^{-\gamma_0}, 1)$, independent of θ and t, such that

$$\left| \frac{\partial}{\partial\omega} \ell_t(\theta) \right| = \left| \frac{-(\omega_0 + \alpha_0\epsilon_{t-1}^2)\eta_t^2}{(\omega + \alpha\epsilon_{t-1}^2)^2} + \frac{1}{\omega + \alpha\epsilon_{t-1}^2} \right| \le K\rho^t(\eta_t^2 + 1). \tag{7.9}$$

Since $\sum_{t=1}^{\infty} K\rho^t(\eta_t^2 + 1)$ has a finite expectation, it is almost surely finite. Thus (7.3) is proved, and (7.4) can be obtained by the same arguments. We have

$$\begin{aligned}
\frac{\partial^2 \ell_t(\omega, \alpha_0)}{\partial\alpha^2} - \frac{1}{\alpha_0^2} &= \left\{ 2\frac{(\omega_0 + \alpha_0\epsilon_{t-1}^2)\eta_t^2}{\omega + \alpha_0\epsilon_{t-1}^2} - 1 \right\} \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \\
&= \left(2\eta_t^2 - 1\right) \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} + r_{1,t} \\
&= 2\left(\eta_t^2 - 1\right) \frac{1}{\alpha_0^2} + r_{1,t} + r_{2,t}
\end{aligned}$$

where

$$\sup_{\theta\in\Theta} |r_{1,t}| = \sup_{\theta\in\Theta} \left| \frac{2(\omega_0 - \omega)\eta_t^2}{(\omega + \alpha_0\epsilon_{t-1}^2)} \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} \right| = o(1) \quad \text{a.s.}$$

and

$$\begin{aligned}\sup_{\theta\in\Theta}|r_{2,t}| &= \sup_{\theta\in\Theta}\left|(2\eta_t^2-1)\left\{\frac{\epsilon_{t-1}^4}{(\omega+\alpha_0\epsilon_{t-1}^2)^2}-\frac{1}{\alpha_0^2}\right\}\right| \\ &= \sup_{\theta\in\Theta}\left|(2\eta_t^2-1)\left\{\frac{\omega^2+2\alpha_0\epsilon_{t-1}^2}{\alpha_0^2(\omega+\alpha_0\epsilon_{t-1}^2)^2}\right\}\right| = o(1) \quad \text{a.s.}\end{aligned}$$

as $t\to\infty$. Therefore (7.5) is established. To prove (7.6), it suffices to remark that

$$\begin{aligned}\left|\frac{\partial^3}{\partial\alpha^3}\ell_t(\theta)\right| &= \left|\left\{2-6\frac{(\omega_0+\alpha_0\epsilon_{t-1}^2)\eta_t^2}{\omega+\alpha\epsilon_{t-1}^2}\right\}\left(\frac{\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2}\right)^3\right| \\ &\leq \left\{2+6\left(\frac{\omega_0}{\omega}+\frac{\alpha_0}{\alpha}\right)\eta_t^2\right\}\frac{1}{\alpha^3}.\end{aligned}$$

We obtain (7.7) and (7.8) similarly in view of Proposition 7.1 *iii)*. □

Proof of Theorem 2.1. *i)* and *iv)* have already been proven (see Berkes, Horváth and Kokoszka (2003) and Francq and Zakoïan (2004)).

To prove *ii)* note that $(\hat\omega_n,\hat\alpha_n)=\arg\min_{\theta\in\Theta}Q_n(\theta)$, where

$$Q_n(\theta)=\frac{1}{n}\sum_{t=1}^n\{\ell_t(\theta)-\ell_t(\theta_0)\}.$$

We have

$$\begin{aligned}Q_n(\theta) &= \frac{1}{n}\sum_{t=1}^n\eta_t^2\left\{\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}-1\right\}+\log\frac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)} \\ &= \frac{1}{n}\sum_{t=1}^n\eta_t^2\frac{(\omega_0-\omega)+(\alpha_0-\alpha)\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2}+\log\frac{\omega+\alpha\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}.\end{aligned}$$

For any $\theta\in\Theta$, we have $\alpha\neq 0$. Letting

$$O_n(\alpha)=\frac{1}{n}\sum_{t=1}^n\eta_t^2\frac{(\alpha_0-\alpha)}{\alpha}+\log\frac{\alpha}{\alpha_0}$$

and

$$d_t=\frac{\alpha(\omega_0-\omega)-\omega(\alpha_0-\alpha)}{\alpha(\omega+\alpha\epsilon_{t-1}^2)},$$

we have, by Proposition 7.1,

$$Q_n(\theta)-O_n(\alpha) = \frac{1}{n}\sum_{t=1}^n\eta_t^2d_{t-1}+\frac{1}{n}\sum_{t=1}^n\log\frac{(\omega+\alpha\epsilon_{t-1}^2)\alpha_0}{(\omega_0+\alpha_0\epsilon_{t-1}^2)\alpha}\to 0 \quad \text{a.s.}$$

Moreover this convergence is uniform on the compact set Θ:

$$\lim_{n\to\infty} \sup_{\theta\in\Theta} |Q_n(\theta) - O_n(\alpha)| = 0 \quad \text{a.s.} \tag{7.10}$$

Let α_0^- and α_0^+ denote two constants such that $0 < \alpha_0^- < \alpha_0 < \alpha_0^+$. Introducing $\hat{\sigma}_\eta^2 = n^{-1} \sum_{t=1}^n \eta_t^2$, the solution of

$$\alpha_n^* = \arg\min_\alpha O_n(\alpha)$$

is $\alpha_n^* = \alpha_0 \hat{\sigma}_\eta^2$. This solution belongs to the interval (α_0^-, α_0^+) for sufficiently large n. Thus

$$\alpha_n^{**} = \arg \min_{\alpha\notin(\alpha_0^-,\alpha_0^+)} O_n(\alpha) \in \{\alpha_0^-, \alpha_0^+\} \tag{7.11}$$

and

$$\lim_{n\to\infty} O_n(\alpha_n^{**}) = \min\left\{\lim_{n\to\infty} O_n(\alpha_0^-), \lim_{n\to\infty} O_n(\alpha_0^+)\right\} > 0. \tag{7.12}$$

This result and (7.10) show that almost surely

$$\lim_{n\to\infty} \min_{\theta\in\Theta,\, \alpha\notin(\alpha_0^-,\alpha_0^+)} Q_n(\theta) > 0.$$

Since $\min_\theta Q_n(\theta) \le Q_n(\theta_0) = 0$, it follows that

$$\lim_{n\to\infty} \arg\min_{\theta\in\Theta} Q_n(\theta) \in (0, \infty) \times (\alpha_0^-, \alpha_0^+).$$

Because the interval (α_0^-, α_0^+) containing α_0 can be chosen arbitrarily small, *ii)* is proven.

Turning to the proof of *iii)* we first note that Proposition 7.1 entails $n^{-1} \sum_{t=1}^n E \sup_{\theta\in\Theta} |d_{t-1}| \to 0$, which implies $\frac{1}{n} \sum_{t=1}^n \sup_{\theta\in\Theta} \eta_t^2 d_{t-1} \to 0$ in L^1. Using the elementary inequality $|\log(x/y)| \le |x-y|/x + |x-y|/y$ for $x, y > 0$, Proposition 7.1 also entails

$$n^{-1} \sum_{t=1}^n \log \frac{(\omega + \alpha\epsilon_{t-1}^2)\alpha_0}{(\omega_0 + \alpha_0\epsilon_{t-1}^2)\alpha} \to 0$$

in L^1 uniformly in θ. It follows that (7.10) can be replaced by

$$\lim_{n\to\infty} E \sup_{\theta\in\Theta} |Q_n(\theta) - O_n(\alpha)| = 0. \tag{7.13}$$

We now note that the asymptotic behavoir of $O_n(\alpha)$ is not affected by the assumption on γ_0. Thus (7.11) and (7.12) still hold. For all $\varepsilon > 0$, we clearly have

$$P\left\{\min_{\theta\in\Theta,\,\alpha\notin(\alpha_0^-,\alpha_0^+)} Q_n(\theta)\leq\varepsilon\right\} \leq P\left\{\min_{\theta\in\Theta,\,\alpha\notin(\alpha_0^-,\alpha_0^+)} O_n(\theta)\leq 2\varepsilon\right\} + P\left\{\sup_{\theta\in\Theta}|Q_n(\theta)-O_n(\alpha)|>\varepsilon\right\}.$$

By (7.11)-(7.12), the first term of the right-hand side of the last inequality tends to zero when $\varepsilon < \lim_{n\to\infty} O_n(\alpha_n^{**})$. The second term tends to zero by (7.13) and the Markov inequality. Because $Q_n(\hat\theta_n) \leq Q_n(\theta_0) = 0$, we conclude that $P\left\{\hat\alpha_n \in (\alpha_0^-, \alpha_0^+)\right\} \to 1$, which shows the weak consistency in *iii)*.

It remains to prove the asymptotic normality of $\hat\alpha_n$ when $\gamma_0 > 0$. Notice that we cannot use the fact that the derivative of the criterion cancels at $\hat\theta_n = (\hat\omega_n, \hat\alpha_n)$ since we have no consistency result for $\hat\omega_n$. Thus the minimum could lie on the boundary of Θ, even asymptotically. However, the partial derivative with respect to α is asymptotically equal to zero at the minimum since $\hat\alpha_n \to \alpha_0$ and (ω_0, α_0) belongs to the interior of Θ. Hence, an expansion of the criterion derivative gives

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\omega}\ell_t(\hat\theta_n) \\ 0 \end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\theta}\ell_t(\theta_0) + J_n\sqrt{n}(\hat\theta_n-\theta_0) \tag{7.14}$$

where J_n is a 2×2 matrix whose elements have the form

$$J_n(i,j) = \frac{1}{n}\sum_{t=1}^n \frac{\partial^2}{\partial\theta_j\partial\theta_j}\ell_t(\theta_i^*)$$

where $\theta_i^* = (\omega_i^*, \alpha_i^*)$ is between $\hat\theta_n$ and θ_0. By Proposition 7.1 *i)* and from the central limit theorem we have

$$\begin{aligned} \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\alpha}\ell_t(\theta_0) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n (1-\eta_t^2)\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2} \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^n (1-\eta_t^2)\frac{1}{\alpha_0} + o_P(1) \\ &\xrightarrow{d} \mathcal{N}\left(0, \frac{\kappa_\eta - 1}{\alpha_0^2}\right). \end{aligned} \tag{7.15}$$

By (7.4) in Lemma 7.1 and the compactness of Θ, we have

$$J_n(2,1)\sqrt{n}(\hat{\omega}_n-\omega_0) \leq \sum_{t=1}^{\infty}\sup_{\theta\in\Theta}\left\|\frac{\partial^2}{\partial\omega\partial\theta}\ell_t(\theta)\right\|\frac{1}{\sqrt{n}}(\hat{\omega}_n-\omega_0)\to 0 \quad \text{a.s.} \tag{7.16}$$

An expansion of the function

$$\alpha \mapsto \frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\alpha^2}\ell_t(\omega_2^*,\alpha)$$

gives

$$J_n(2,2)=\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\alpha^2}\ell_t(\omega_2^*,\alpha_0)+\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^3}{\partial\alpha^3}\ell_t(\omega_2^*,\alpha^*)(\alpha_2^*-\alpha_0)$$

where α^* is between α_2^* and α_0. Using (7.5), (7.6) and *ii)* we get

$$J_n(2,2)\to\frac{1}{\alpha_0^2} \quad \text{a.s.} \tag{7.17}$$

The conclusion follows, by considering the second component in (7.14) and from (7.15), (7.16) and (7.17). □

Proof of Theorem 2.2. The proof of the asymptotic normality still relies on the Taylor expansion (7.14). By the Lindeberg central limit theorem for martingale differences (see Billingsley, 1995, p. 476), the asymptotic normality (7.15) of the score vector is obtained by showing that

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(1-\eta_t^2)\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}\right) &= \frac{\kappa_\eta-1}{n}\sum_{t=1}^{n}E\left|\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}\right|^2\\ &\to (\kappa_\eta-1)\alpha_0^{-2},\end{aligned}$$

which is a consequence of Proposition 7.1 *iii)*, and by noting that for all $\varepsilon>0$

$$\begin{aligned}&\frac{\kappa_\eta-1}{n}\sum_{t=1}^{n}E\left|\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}\right|^2 1_{\left\{\left|\frac{1-\eta_t^2}{\sqrt{n}}\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}\right|>\varepsilon\right\}}\\ \leq\ &\frac{\kappa_\eta-1}{\alpha_0^2}P\left(|1-\eta_t^2|>\alpha_0\varepsilon\sqrt{n}\right)\to 0.\end{aligned}$$

To deal with the second term in the right-hand side of (7.14) we cannot use (7.16) because (7.4) requires $\gamma_0>0$. Instead, noting that $\sigma_t^2(\theta_2^*)/\sigma_t^2(\theta_0)$ is bounded, we use

$$\begin{aligned}|J_n(2,1)\sqrt{n}(\hat{\omega}_n-\omega_0)| &\le \frac{K}{\sqrt{n}}\sum_{t=1}^n\left(\frac{2\eta_t^2\sigma_t^2(\theta_0)}{\sigma_t^2(\theta_2^*)}+1\right)\frac{\epsilon_{t-1}^2}{\sigma_t^4(\theta_2^*)}\\ &\le \frac{K}{\sqrt{n}}\sum_{t=1}^n\left(\eta_t^2+1\right)\frac{\epsilon_{t-1}^2}{\sigma_t^4(\theta_2^*)},\end{aligned}$$

where $K>0$ is a generic constant whose value can change along the proof. Hence,

$$E|J_n(2,1)\sqrt{n}(\hat{\omega}_n-\omega_0)| \le \frac{K}{\sqrt{n}}\sum_{t=1}^n E\frac{\epsilon_{t-1}^2}{(\omega_2^*+\alpha_2^*\epsilon_{t-1}^2)^2}\le\frac{K}{\sqrt{n}}\sum_{t=1}^n E\frac{1}{\sigma_{t-1}^2(\theta_0)}.$$

Moreover,

$$\sigma_t^2(\theta_0)=\omega_0(1+Z_{t-1}+Z_{t-1}Z_{t-2}+\cdots+Z_{t-1}\dots Z_1)+Z_{t-1}\dots Z_0\sigma_0^2.$$

By Assumption **A**, it follows that

$$E|J_n(2,1)\sqrt{n}(\hat{\omega}_n-\omega_0)|\to 0 \text{ as } n\to\infty. \tag{7.18}$$

Finally, similarly to (7.17)

$$J_n(2,2)\to\frac{1}{\alpha_0^2}\quad\text{in probability} \tag{7.19}$$

using *iii)* in Theorem 2.1, (7.7) and (7.8). The conclusion follows as in the case $\gamma_0>0$. □

Proof of Theorem 2.3. The convergence results in *i)* can be shown in a standard way, using Taylor expansions of the functions $\hat{\kappa}_\eta=\kappa_\eta(\hat{\theta}_n)$ and $\hat{\mu}_n(p,q)=\mu_n(p,q)(\hat{\theta}_n)$ around θ_0, and the ergodic theorem together with the consistency of $\hat{\theta}_n$.

Now consider the case *ii)*. For some $\theta^*=(\omega^*,\alpha^*)'$ between $\hat{\theta}_n$ and θ_0 we have

$$\hat{\kappa}_\eta=\frac{1}{n}\sum_{t=1}^n\eta_t^4-\frac{2}{n}\sum_{t=1}^n\frac{\epsilon_t^4}{\sigma_t^4(\theta^*)}\frac{1}{\sigma_t^2(\theta^*)}\frac{\partial\sigma_t^2(\theta^*)}{\partial\theta'}(\theta^*-\theta_0):=\frac{1}{n}\sum_{t=1}^n\eta_t^4+R_n. \tag{7.20}$$

By Proposition 7.1, for some constants $K>0$ and $\rho\in(0,1)$,

$$|R_n|\le\frac{K}{n}\sum_{t=1}^n\eta_t^4\left(\rho^t|\omega^*-\omega_0|+|\alpha^*-\alpha_0|\right)=o_P(1)$$

where the last equality follows from the strong consistency of $\hat{\alpha}_n$ and the fact that $|\omega^* - \omega_0|$ is bounded by compactness of Θ. Hence the first part of *ii)* is proven. Now note that

$$\begin{aligned}\hat{\mu}_n(2,2) &= \frac{1}{\hat{\alpha}_n^2} + \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\epsilon_t^4}{\left(\hat{\omega}_n + \hat{\alpha}_n\epsilon_t^2\right)^2} - \frac{1}{\hat{\alpha}_n^2}\right\} \\ &= \frac{1}{\hat{\alpha}_n^2} - \frac{\hat{\omega}_n^2}{\hat{\alpha}_n^2}\hat{\mu}_n(0,2) - 2\frac{\hat{\omega}_n}{\hat{\alpha}_n}\hat{\mu}_n(1,2).\end{aligned}$$

Similarly we have

$$\hat{\mu}_n(1,2) = -\frac{\hat{\omega}_n}{\hat{\alpha}_n}\hat{\mu}_n(0,2) + \frac{1}{\hat{\alpha}_n}\hat{\mu}_n(0,1).$$

It follows that

$$\hat{\xi}_n = \hat{\alpha}_n^2\left\{1 - \frac{\hat{\mu}_n^2(0,1)}{\hat{\mu}_n(0,2)}\right\}^{-1}. \tag{7.21}$$

In order to show that $\hat{\xi}_n \to \alpha_0^2$, it thus remains to show that $\hat{\mu}_n^2(0,1)/\hat{\mu}_n(0,2) = o(1)$ *a.s.* First note that $\hat{\mu}_n(0,2) \geq n^{-1}\hat{\omega}_n^{-2}$. Since

$$\sigma_t^2(\hat{\theta}_n) = \hat{\omega}_n + \hat{\alpha}_n\eta_{t-1}^2\sigma_{t-1}^2(\theta_0) \geq \omega_0\hat{\alpha}_n\alpha_0^{t-2}\eta_{t-1}^2\eta_{t-2}^2\cdots\eta_1^2,$$

we have

$$n\hat{\mu}_n(0,1) = \sum_{t=1}^{n}\frac{1}{\sigma_{t+1}^2(\hat{\theta}_n)} \leq \frac{\alpha_0}{\omega_0\hat{\alpha}_n}\sum_{t=1}^{\infty}\frac{1}{\alpha_0^t\eta_t^2\eta_{t-1}^2\cdots\eta_1^2}.$$

By the Cauchy root test, the last series is almost surely finite because

$$\limsup_{t\to\infty}\left(\frac{1}{\alpha_0^t\eta_t^2\eta_{t-1}^2\cdots\eta_1^2}\right)^{1/t} = \exp(-\gamma_0) < 1 \quad a.s.$$

We thus have shown that $\hat{\mu}_n^2(0,1) = O(n^{-1})$ *a.s.*, which completes the proof of *ii)*.

Turning to *iii)*, we note that in (7.20), $R_n =: S_n(\theta^* - \theta_0)$ with

$$E|S_n| \leq \frac{K}{n}\sum_{t=1}^{n}E\left(\frac{1}{\alpha^*\epsilon_{t-1}^2}(1,\epsilon_{t-1}^2)\right) \to (0,K)$$

by Proposition 7.1 *iii)*. Since the first component of $\theta^* - \theta_0$ is bounded, by compactness of Θ, and the second component tends to zero in probability, by

Theorem 2.1 *iii)*, the first convergence is established. To complete the proof we note that $\hat{\mu}_n(0,2) \geq Kn^{-1}$ and that, in view of Assumption **A**,

$$E\sqrt{n}|\hat{\mu}_n(0,1)| = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} E\frac{1}{\sigma_{t+1}^2(\hat{\theta}_n)} = o(1).$$

The conclusion follows from (7.21). □

7.3. *Inconsistency of $\hat{\omega}_n$ when $\gamma_0 \geq 0$.* The previous results do not give any insight on the asymptotic behavior of the QMLE of ω_0. Similarly to (7.15) it can be shown that the score vector satisfies

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-2}\end{pmatrix}\right\}. \tag{7.22}$$

The form of the asymptotic variance shows that, for n sufficiently large and almost surely, the variation of the log-likelihood $n^{-1/2}\sum_{t=1}^{n}\log \ell_t(\theta)$ is negligible when θ varies between (ω_0, α_0) and $(\omega_0 + h, \alpha_0)$ for small h.

Note that a score vector with a degenerate asymptotic variance J can arise when a central limit theorem with a non standard rate of convergence applies. This is for instance the case in regressions with trends, or in unit root and cointegration models. In such situations, the rate of convergence of the QMLE is obtained by finding a diagonal matrix Λ_n such that the asymptotic distribution of $\Lambda_n\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0)$ is not degenerated. If, for instance, $\Lambda_n = \mathrm{diag}(n^{-1}, n^{-1/2})$ then the second component of QMLE is expected to converge at the standard rate $\sqrt{n}$, and the first one at the faster rate n. The situation here is completely different. In the proof of Theorem 2.1 it is shown that $\frac{\partial}{\partial\omega}\ell_t(\theta_0) = O_P(\rho^t)$ with $|\rho| < 1$ (see Equation (7.9) below). The equation (7.22) can thus be extended as

$$\Lambda_n\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-2}\end{pmatrix}\right\}, \quad \Lambda_n = \begin{pmatrix} \lambda_n & 0 \\ 0 & n^{-1/2}\end{pmatrix}$$

for any sequence λ_n tending to zero as $n \to \infty$. It means that the log-likelihood is completely flat in the direction where α_0 is fixed and ω_0 varies. Thus there is little hope concerning the existence of any consistent estimator of ω_0. This is in accordance with the numerical illustrations provided in Section 6.

7.4. *A constrained QMLE of α_0.* The asymptotic behaviour of the QMLE $\hat{\alpha}_n$ being independent of ω_0 when $\gamma_0 > 0$, and the QMLE of ω_0 being probably inconsistent in view of the previous remark, it seems natural to avoid estimating ω_0. To this aim a *constrained* QMLE of α_0, in which the first component of θ is fixed to an arbitrary value ω, can be introduced. The estimator

$$\hat{\alpha}_n^c(\omega) = \arg\min_{\alpha\in\Theta_1} \frac{1}{n}\sum_{t=1}^{n} \ell_t(\omega,\alpha), \qquad \Theta_1 \text{ compact } \subset (0,\infty) \tag{7.23}$$

was studied by Jensen and Rahbek (2004a). They proved that, when $\gamma_0 > 0$, (2.3) continues to hold when the QMLE $\hat{\alpha}_n$ is replaced by the constrained QMLE $\hat{\alpha}_n^c(\omega_0)$.[4] In the appendix we prove that:
under the assumptions of Theorem 2.2, in particular $\gamma_0 = 0$,

$$\sqrt{n}\left(\hat{\alpha}_n^c(\omega) - \alpha_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\alpha_0^2\right\}, \quad \text{as } n \to \infty. \tag{7.24}$$

However, the next result shows that the restricted QMLE of α_0 is generally inconsistent in the stationary case.

Proposition 7.2. *Let (ϵ_t) be a stationary solution of the ARCH(1) model with parameters ω_0 and α_0, such that $E\epsilon_t^4 < \infty$. Then, if $\omega \neq \omega_0$*

$$\hat{\alpha}_n^c(\omega) \textit{ does not converge in probability to } \alpha_0. \tag{7.25}$$

On the contrary, Theorems 2.1-2.2 show that

$$\text{the QMLE of } \alpha_0 \text{ is always CAN} \tag{7.26}$$

(under **A** when $\gamma_0 = 0$).

Proof of (7.24). First note that, by the arguments used to prove *iii)* in Theorem 2.1, we have

$$\hat{\alpha}_n^c(\omega) \to \alpha_0 \qquad \text{in probability as } n \to \infty. \tag{7.27}$$

A Taylor expansion of the criterion derivative gives

$$\begin{aligned} 0 &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{\partial}{\partial\alpha}\ell_t(\omega, \hat{\alpha}_n^c(\omega)) \\ &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{\partial}{\partial\theta}\ell_t(\omega, \alpha_0) + \left(\frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2}{\partial\alpha^2}\ell_t(\omega, \alpha^*)\right)\sqrt{n}(\hat{\theta}_n - \theta_0) \qquad (7.28) \end{aligned}$$

[4] In fact, the result was announced under the assumption $\gamma_0 \geq 0$ but their proof is only valid under $\gamma_0 > 0$ because the a.s. convergence of ϵ_t^2 to infinity is used (see their Lemma 1).

where α^* is between $\hat{\alpha}_n^c(\omega)$ and α_0. Another Taylor expansion yields, for α^{**} between $\hat{\alpha}_n^c(\omega)$ and α_0

$$\begin{aligned}\left|\frac{1}{n}\sum_{t=1}^n \frac{\partial^2}{\partial\alpha^2}\ell_t(\omega,\alpha_0) - \frac{1}{n}\sum_{t=1}^n \frac{\partial^2}{\partial\alpha^2}\ell_t(\omega,\alpha^*)\right| &\leq |\alpha^*-\alpha_0|\frac{1}{n}\sum_{t=1}^n\left|\frac{\partial^2}{\partial\alpha^3}\ell_t(\omega,\alpha^{**})\right| \\ &\leq |\alpha^*-\alpha_0|\frac{1}{n}\sum_{t=1}^n\sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\alpha^3}\ell_t(\theta)\right| \\ &= o(1) \quad \text{in probability,}\end{aligned}$$

using (7.27) and (7.8). Therefore, using (7.7), the term in parentheses in (7.28) converges to $1/\alpha_0^2$. To conclude, it remains to prove that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\alpha}\ell_t(\omega,\alpha_0) \quad \xrightarrow{d} \quad \mathcal{N}\left(0, \frac{\kappa_\eta - 1}{\alpha_0^2}\right). \tag{7.29}$$

We have

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\alpha}\ell_t(\omega,\alpha_0) \\ = \quad &\frac{1}{\sqrt{n}}\sum_{t=1}^n (1-\eta_t^2)\frac{\epsilon_{t-1}^2}{\omega+\alpha_0\epsilon_{t-1}^2} + \frac{1}{\sqrt{n}}\sum_{t=1}^n \eta_t^2\frac{(\omega-\omega_0)\epsilon_{t-1}^2}{(\omega+\alpha_0\epsilon_{t-1}^2)^2}.\end{aligned}$$

The last term tends to zero in probability, using **A**, similarly to (7.18). The first term converges in distribution to the normal law of (7.29) by exactly the same arguments as in the proof of Theorem 2.2. □

Proof of Proposition 7.2. The ergodic theorem entails that, almost surely,

$$\begin{aligned}L_n(\alpha) &= \frac{1}{n}\sum_{t=1}^n \frac{\epsilon_t^2}{\sigma_t^2(\omega,\alpha)} + \log\sigma_t^2(\omega,\alpha) \\ \to \quad L(\alpha) &= E\left\{\frac{\omega_0+\alpha_0\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2} + \log\left(\omega+\alpha\epsilon_{t-1}^2\right)\right\}\end{aligned}$$

as $n\to\infty$. The dominated convergence theorem implies that

$$\begin{aligned}L'(\alpha) &= E\frac{\partial}{\partial\alpha}\left\{\frac{\omega_0+\alpha_0\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2} + \log\left(\omega+\alpha\epsilon_{t-1}^2\right)\right\} \\ &= E\left\{\frac{\epsilon_{t-1}^2}{\left(\omega+\alpha\epsilon_{t-1}^2\right)^2}\{(\omega-\omega_0)+(\alpha-\alpha_0)\epsilon_{t-1}^2\}\right\}.\end{aligned}$$

First suppose that $\omega < \omega_0$. Then $L'(\alpha) < 0$ for $\alpha \le \alpha_0$. The intermediate values theorem shows that the function $L(\cdot)$ has a minimum at a point $\alpha^* > \alpha_0$ and that $L(\alpha^*) < L(\alpha_0)$. Now suppose that $\omega > \omega_0$. Then $L'(\alpha) > 0$ for $\alpha \ge \alpha_0$. This shows that $L(\cdot)$ has a minimum at a point $\alpha^* \in [0, \alpha_0)$, with $L(\alpha^*) < L(\alpha_0)$. Thus, we have shown that for any $\omega \ne \omega_0$, the function $L(\cdot)$ has a minimum at a point $\alpha^* \ne \alpha_0$ and $L(\alpha^*) < L(\alpha_0)$.

A Taylor expansion of $L_n(\cdot)$ yields

$$L_n\left\{\hat{\alpha}_n^c(\omega)\right\} = L_n(\alpha_0) + L_n'(\tilde{\alpha}_n)\{\hat{\alpha}_n^c(\omega) - \alpha_0\} \tag{7.30}$$

where $\tilde{\alpha}_n$ is between $\hat{\alpha}_n^c(\omega)$ and α_0. Note that since $E\epsilon_t^4 < \infty$, almost surely,

$$\limsup_{n\to\infty} \sup_{\alpha} \left|L_n'(\alpha)\right| \le \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left(1 + \frac{\epsilon_t^2}{\omega}\right) \frac{\epsilon_{t-1}^2}{\omega} < \infty.$$

Now suppose that

$$\hat{\alpha}_n^c(\omega) \to \alpha_0, \quad \text{in probability as } n \to \infty. \tag{7.31}$$

Then, it follows from (7.30) that

$$L_n\left\{\hat{\alpha}_n^c(\omega)\right\} \to L(\alpha_0), \quad \text{in probability as } n \to \infty.$$

Then, taking the limit in probability in the following inequality

$$L_n\{\hat{\alpha}_n^c(\omega)\} \le L_n(\alpha^*)$$

we find that $L(\alpha_0) \le L(\alpha^*)$, which is in contradiction with the definition of $\alpha^* \ne \alpha_0$. Thus (7.31) cannot be true. □

7.5. *Stationarity test.* **Proof of Theorem 3.1.** First consider the case $\gamma_0 < 0$. Let $\zeta_n = n^{-1}\sum_{t=1}^{n} \log \eta_t^2$. Note that $\hat{\zeta}_n = \zeta_n(\hat{\theta}_n)$ and $\zeta_n = \zeta_n(\theta_0)$ with $\zeta_n(\theta) = n^{-1}\sum_{t=1}^{n} \log \eta_t^2(\theta)$ and $\eta_t(\theta) = \epsilon_t/\sigma_t(\theta)$. A Taylor expansion thus gives

$$\hat{\zeta}_n = \zeta_n + \frac{\partial \zeta_n(\theta_0)}{\partial \theta'}(\hat{\theta}_n - \theta_0) + o_P(n^{-1/2}) \tag{7.32}$$

with

$$\frac{\partial \zeta_n(\theta_0)}{\partial \theta'} = \frac{-1}{n} \sum_{t=1}^{n} \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'}.$$

Moreover the QMLE satisfies

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -J^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} v_t \frac{1}{h_t} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} + o_P(1). \tag{7.33}$$

In view of (7.32) and (7.33), we have

$$\sqrt{n}(\hat{\zeta}_n - \zeta) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t + \Omega' J^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} v_t \frac{1}{h_t}\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} + o_P(1),$$

where $\Omega = Eh_t^{-1}\partial\sigma_t^2(\theta_0)/\partial\theta$. Note that

$$\mathrm{Cov}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t,\ \Omega' J^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} v_t \frac{1}{h_t}\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta}\right) = \sigma_{uv}\Omega' J^{-1}\Omega.$$

The Slutsky lemma and the central limit theorem for martingale differences thus entail

$$\sqrt{n}\begin{pmatrix} \hat{\zeta}_n - \zeta \\ \hat{\theta}_n - \theta_0 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{0, \Sigma := \begin{pmatrix} \sigma_u^2 + (\sigma_v^2 + 2\sigma_{uv})\Omega' J^{-1}\Omega & -(\sigma_v^2 + \sigma_{uv})\Omega' J^{-1} \\ -(\sigma_v^2 + \sigma_{uv}) J^{-1}\Omega & \sigma_v^2 J^{-1} \end{pmatrix}\right\}.$$

Noting that $\theta_0'\partial\sigma_t^2(\theta_0)/\partial\theta = h_t$ almost surely, we have

$$E\left\{\frac{1}{h_t}\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta}\left(1 - \frac{1}{h_t}\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'}\theta_0\right)\right\} = 0,$$

which entails $J\theta_0 = \Omega$. Thus $J^{-1}\Omega = \theta_0$ and $\Omega' J^{-1}\Omega = 1$, and (3.5) follows.

The convergence in probability in (3.6) is a straightforward consequence of (3.5). By direct application of the delta method (see *e.g.* Theorem 3.1 in van der Vaart, 1998), in the case $\gamma_0 < 0$

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}\left(0, L\Sigma L'\right) \quad \text{where} \quad L = \left(1,\, 0,\, \alpha_0^{-1}\right).$$

It is easy to verify that $L\Sigma L' = \sigma_u^2 + \sigma_v^2\left\{\frac{\xi}{\alpha_0^2} - 1\right\}$.

Now consider the case $\gamma_0 \geq 0$. Note that,

$$\frac{\partial^2 \zeta_n(\theta)}{\partial\theta\partial\theta'} = J_n(\theta) := \frac{1}{n}\sum_{t=1}^{n}\frac{1}{\sigma_t^4(\theta)}\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \to J(\theta) := \begin{pmatrix} 0 & 0 \\ 0 & \alpha^{-2} \end{pmatrix}$$

a.s. (resp. in probability) as $n \to \infty$ when $\gamma_0 > 0$ (resp. when $\gamma_0 = 0$), uniformly in $\theta \in \Theta$, by Proposition 7.1. Moreover the matrix $\Lambda_n J_n(\theta)\Lambda_n$ converges to the same limit, where Λ_n is the diagonal matrix with elements $n^{1/4}$ and 1. Thus

$$\hat{\zeta}_n = \zeta_n + \frac{\partial \zeta_n(\theta_0)}{\partial \theta'}(\hat{\theta}_n - \theta_0) + \frac{1}{2}(\hat{\theta}_n - \theta_0)'\Lambda_n^{-1}\Lambda_n J_n(\theta^*)\Lambda_n\Lambda_n^{-1}(\hat{\theta}_n - \theta_0). \tag{7.34}$$

Noting that $\Lambda_n^{-1}(\hat{\theta}_n - \theta_0)$ tends to zero (a.s. when $\gamma_0 > 0$, in probability when $\gamma_0 = 0$), we conclude that (7.32) still holds. By the same arguments,

$$\frac{1}{n}\sum_{t=1}^{n}\frac{1}{h_t}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta'} \to \begin{pmatrix} 0 \\ \alpha_0^{-1} \end{pmatrix} \quad \text{a.s. (in probability when } \gamma_0 = 0).$$

Therefore

$$\sqrt{n}(\hat{\zeta}_n - \zeta) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t - \alpha_0^{-1}\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) + o_P(1). \tag{7.35}$$

From the proof of (2.3) and Theorem 2.2, it can be seen that

$$\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) = -\alpha_0 n^{-1/2}\sum_{t=1}^{n} v_t + o_P(1), \tag{7.36}$$

and (3.8) follows. The rest of the proof is as in the case $\gamma_0 < 0$. □

Proof of Corollary 3.2. By arguments used in the proof of Theorem 2.3, $\hat{\sigma}_u^2$ converges almost surely to σ_u^2 when $\gamma < 0$ or $\gamma \geq 0$. Therefore $T_n = \sqrt{n}(\hat{\gamma}_n - \gamma_0)/\hat{\sigma}_u + \sqrt{n}\gamma_0/\hat{\sigma}_u$ converges in probability to $-\infty$ when $\gamma < 0$, to $+\infty$ when $\gamma > 0$, and in distribution to the $\mathcal{N}(0,1)$ when $\gamma_0 = 0$. □

Proof of Proposition 4.1. We consider the case $\gamma_0 \geq 0$ because the LAN of GARCH models has already be established in the stationary case (see Drost and Klaassen (1997), Lee and Taniguchi (2005)). A Taylor expansion of $\tau_n \mapsto \Lambda_{n,f}(\theta_0 + \tau_n/\sqrt{n}, \theta_0)$ around 0 yields

$$\Lambda_{n,f}(\theta_0 + \tau_n/\sqrt{n}, \theta_0) = \tau_n' S_{n,f}(\theta_0) - \frac{1}{2}\tau_n'\mathfrak{I}_n(\theta_n^*)\tau_n, \tag{7.37}$$

where $\theta_n^* = \theta_0 + \tau_n^*/\sqrt{n}$ with τ_n^* between 0 and τ_n,

$$S_{n,f}(\theta_0) = \frac{-1}{\sqrt{n}}\sum_{t=1}^{n}\left\{1 + \eta_t\frac{f'(\eta_t)}{f(\eta_t)}\right\}\frac{1}{2\sigma_t^2(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta} \tag{7.38}$$

and, introducing the function $g(y) = 1 + 2y(f'/f)(y) + y^2(f'/f)'(y)$,

$$\mathfrak{I}_n(\theta) = -(1/4n)\sum_{t=1}^{n} g\left(\frac{\epsilon_t}{\sigma_t(\theta)}\right)\Delta_t(\theta), \qquad \Delta_t(\theta) = \frac{1}{\sigma_t^4(\theta)}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'}.$$

As in the proof of Theorem 2.2, the Lindeberg central limit theorem for martingale differences shows that

$$S_{n,f}(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \mathfrak{I}_f\right\}. \tag{7.39}$$

Let the matrix norm defined by $\|A\| = \sum |a_{ij}|$ with standard notations. We have

$$\begin{aligned}\|\Delta_t(\theta) - \Delta_t(\theta_0)\| &\leq |\omega_0 - \omega| \left(\frac{2}{\omega\omega_0} + \frac{1}{\alpha\omega_0} + \frac{1}{\alpha_0\omega} \right) \\ &\quad + |\alpha_0 - \alpha| \left(\frac{2}{\alpha\alpha_0} + \frac{1}{\alpha\omega_0} + \frac{1}{\alpha_0\omega} \right) \\ &= O(\|\theta - \theta_0\|).\end{aligned}$$

Note that (4.2)-(4.3) implies $E|g(\eta_1)| < \infty$. The law of large numbers then entails

$$\begin{aligned}&\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{t=1}^{n} g(\eta_t) \left\{ \Delta_t(\theta) - \Delta_t(\theta_0) \right\} \right\| \\ \leq \; & O(\|\theta - \theta_0\|) \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} |g(\eta_t)| = O(\|\theta - \theta_0\|) \quad a.s.\end{aligned}$$

and

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{t=1}^{n} g(\eta_t) \left\{ \Delta_t(\theta_n^*) - \Delta_t(\theta_0) \right\} \right\| = 0 \quad a.s.$$

Noting that

$$\left| \frac{\epsilon_t}{\sigma_t(\theta)} \right| = \left| \frac{\sigma_t(\theta_0)}{\sigma_t(\theta)} \eta_t \right| \leq \left(\sqrt{\frac{\omega_0}{\omega}} + \sqrt{\frac{\alpha_0}{\alpha}} \right) |\eta_t|,$$

and that $\sup_\theta \|\Delta_t(\theta)\| = O(1)$, Assumption (4.2) and the mean value theorem entail

$$\left\| \frac{1}{n} \sum_{t=1}^{n} \left\{ g\left(\frac{\epsilon_t}{\sigma_t(\theta)} \right) - g(\eta_t) \right\} \Delta_t(\theta) \right\| \leq O(\|\theta - \theta_0\|) \frac{1}{n} \sum_{t=1}^{n} |\eta_t|^{\delta+1}.$$

The Hölder inequality shows that $E\,|\eta_t|^{\delta+1} \leq \sqrt{E\,|\eta_t|^{2\delta}\, E\eta_t^2}$, which is finite under (4.3). It follows that

$$\frac{1}{n} \sum_{t=1}^{n} \left\| \left\{ g\left(\frac{\epsilon_t}{\sigma_t(\theta_n^*)} \right) - g(\eta_t) \right\} \Delta_t(\theta_n^*) \right\| \to 0 \quad a.s.$$

We thus have shown that, as $n \to \infty$,

$$\|\mathfrak{I}_n(\theta_n^*) - \mathfrak{I}_n(\theta_0)\| \to 0 \quad a.s. \tag{7.40}$$

Integrations by parts show that, under (4.1), we have $\int y^2 f''(y)dy = -2\int y f'(y)dy = 2$. It follows that $Eg(\eta_1) = -\iota_f$. Proposition 7.1 *iii)* shows that

$$E\left\| \frac{1}{n}\sum_{t=1}^{n} g(\eta_t)\left\{\Delta_t(\theta_0) - \begin{pmatrix} 0 \\ 1/\alpha_0 \end{pmatrix}\right\}\right\|$$

$$\leq \quad E|g(\eta_1)|E\left\| \frac{1}{n}\sum_{t=1}^{n} \begin{pmatrix} \frac{1}{\omega_0+\alpha_0\epsilon_{t-1}^2} \\ \frac{\omega_0}{\omega_0+\alpha_0\epsilon_{t-1}^2} \end{pmatrix}\right\| \to 0,$$

and thus, by the law of large numbers

(7.41)

$$\mathfrak{I}_n(\theta_0) = \frac{1}{n}\sum_{t=1}^{n} \frac{-g(\eta_t)}{4}\begin{pmatrix} 0 \\ 1/\alpha_0 \end{pmatrix} + O_{P_{\theta_0}}(1) \to \mathfrak{I}_f \text{ in probability as } n \to \infty.$$

The conclusion follows from (7.37)–(7.41). □

Proof of Proposition 4.2. For simplicity, write P instead of $P_{n,0}$. By the delta method and (7.36) we obtain

$$\sqrt{n}\,(\log \hat{\alpha}_n - \log \alpha_0) = \frac{1}{\alpha_0}\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) + o_P(1) = -\frac{1}{\sqrt{n}}\sum_{t=1}^{n} v_t + o_P(1).$$

In view of (7.35), and noting $\zeta + \log \alpha_0 = 0$, we have

$$T_n = \sqrt{n}\frac{\hat{\zeta}_n - \zeta + \log \hat{\alpha}_n - \log \alpha_0}{\hat{\sigma}_u} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{u_t}{\sigma_u} + o_P(1).$$

By (4.4) and (7.38), it follows that under P

$$\begin{pmatrix} T_n \\ \Lambda_{n,f}(\theta_0 + \tau/\sqrt{n}, \theta_0) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{\begin{pmatrix} 0 \\ -\frac{\tau_2^2 \iota_f}{8\alpha_0^2} \end{pmatrix}, \begin{pmatrix} 1 & c \\ c & \frac{\tau_2^2\iota_f}{4\alpha_0^2} \end{pmatrix}\right\}, \tag{7.42}$$

where

$$c = -\frac{\tau_2}{2\alpha_0\sigma_u}Eu_1\left\{1 + \eta_1\frac{f'(\eta_1)}{f(\eta_1)}\right\} = \frac{\tau_2}{\alpha_0\sigma_u}. \tag{7.43}$$

For the last equality, we used an integration by parts and we noted that $E\eta_1^2 < \infty$ entails $\lim_{x\to\pm\infty} x \log x^2 f(x) = 0$. Le Cam's third lemma (see e.g. van der Vaart, 1998, page 90) shows that

$$T_n \xrightarrow{d} \mathcal{N}\left(\frac{\tau_2}{\alpha_0\sigma_u}, 1\right), \quad \text{under } P_{n,\tau}.$$

The conclusion easily follows. □

Proof of Proposition 4.3. Distinguishing the cases $\gamma_0 < 0$ and $\gamma_0 \geq 0$, and reasoning as in the proof of Proposition 4.2, it can be shown that $\hat{\alpha}_n$ is a regular estimator of α^*, in the sense that

$$\frac{\sqrt{n}(\hat{\alpha}_n - \alpha^* - \tau_2/\sqrt{n})}{\sqrt{(\hat{\kappa}_\eta - 1)\hat{\xi}_n}} \xrightarrow{d} \mathcal{N}(0,1), \quad \text{under} \quad P^{\alpha^*}_{n,\tau} \quad \text{as} \quad n \to \infty.$$

The conclusion easily follows. □

Proof of Proposition 4.4. In view of (4.6) and (4.8), the C$^{\text{ST}}$-test is asymptotically optimal if and only if $\tau_2/\alpha_0\sigma_u = \tau_2/\sqrt{4\alpha_0^2/\iota_f}$, which is equivalent to $\sigma_u^2\iota_f = 4$. In the second equality of (7.43), we have seen that

$$\int (\log y^2 - \zeta)\left(1 + \frac{f'(y)}{f(y)}y\right) f(y)dy \quad = \quad -2.$$

Thus, the Cauchy-Schwarz inequality yields

$$4 \quad \leq \quad \int (\log y^2 - \zeta)^2 f(y)dy \int \left(1 + \frac{f'(y)}{f(y)}y\right)^2 f(y)dy = \sigma_u^2\iota_f$$

with equality iff there exists $a \neq 0$ such that $1 + \eta_t f'(\eta_t)/f(\eta_t) = a\left(\log \eta_t^2 - \zeta\right)$ a.s. Such densities f must satisfy the differential equation $f'(y)/f(y) = 2a \log|y|/y - (1+\zeta a)/y$ for almost every y. Setting $a = -1/2\sigma^2$ and $b = \zeta/4\sigma^2$, the solutions are two-sided generalized log-normal densities of the form

$$f(y) = \frac{1}{2\sigma\sqrt{2\pi}e^{2b^2\sigma^2}} e^{-\frac{(\log|y|)^2}{2\sigma^2}} |y|^{-1+2b}, \quad \sigma^2 > 0,\, b \in \mathbb{R}.$$

Direct computations show that $\int \log y^2 f(y)dy = 4b\sigma^2$ and $\int y^2 f(y)dy = \exp\{2\sigma^2(2b+1)\}$, which shows that the test is optimal iff f is given by (4.11).

We now give the proof of the first result. In view of (4.7) and (4.9), the C$^{\alpha^*}$-test is asymptotically optimal if and only if $(\kappa_\eta - 1)\iota_f = 4$. By Corollary 1 in Francq and Zakoïan (2006b), the solutions of this equation are given by (4.11). □

Proof of Proposition 5.1. We start by considering the case $\Gamma > 0$. By the arguments given in the proof of Proposition 7.1 i), $h_t \to \infty$ and $\epsilon_t^2 \to \infty$ a.s. at an exponential rate as $t \to \infty$. Let

$$\alpha_0^* = E\left(\frac{a(\eta_t)}{\eta_t^2}\right).$$

Following the lines of the proof of Theorem 2.1, note that $(\hat{\omega}_n, \hat{\alpha}_n) = \arg\min_{\theta\in\Theta} Q_n(\theta)$, where

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{\epsilon_t^2}{\omega+\alpha\epsilon_{t-1}^2} + \log(\omega+\alpha\epsilon_{t-1}^2) - \frac{\epsilon_t^2}{\alpha_0^*\epsilon_{t-1}^2} - \log(\alpha_0^*\epsilon_{t-1}^2)\right\}.$$

Letting

$$O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{a(\eta_{t-1})}{\eta_{t-1}^2}\frac{(\alpha_0^*-\alpha)}{\alpha\alpha_0^*} + \log\frac{\alpha}{\alpha_0^*}$$

we have (7.10) by arguments already used. Noting that

$$\arg\min_{\alpha} O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{a(\eta_{t-1})}{\eta_{t-1}^2} \to \alpha_0^*, \quad a.s.$$

we conclude, as in the proof of Proposition 7.1 ii), that $\hat{\alpha}_n \to \alpha_0^*, \quad a.s.$ Note that for t large enough $\omega^s(\eta_t)/h_t^s < \omega^s(\eta_t)/t$, and $\lim_{t\to\infty}\omega^s(\eta_t)/t \to 0$ a.s. We thus have

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\log\hat{\eta}_t^2 &= \frac{1}{n}\sum_{t=1}^{n}\log\frac{\epsilon_t^2}{\hat{\alpha}_n\epsilon_{t-1}^2} + o_P(1)\\ &= \frac{1}{n}\sum_{t=1}^{n}\log\frac{\eta_t^2\{\omega(\eta_{t-1})+a(\eta_{t-1})h_{t-1}\}}{\hat{\alpha}_n h_{t-1}\eta_{t-1}^2} + o_P(1)\\ &\to \Gamma - \log\alpha_0^*,\end{aligned}$$

and the first result follows. The convergence of $\hat{\sigma}_u^2$ is obtained by similar arguments. By assumption $E|\log\eta_1^2|^2 > 0$, which entails that η_1^2 has a non-degenerate distribution. It follows that the a.s. limit of $\hat{\sigma}_u^2$ is positive.

Now consider the case $\Gamma < 0$. Letting $\eta_t^{*2} = \epsilon_t^2/(\omega^* + \alpha^*\epsilon_{t-1}^2)$, and using the inequality $|\log(x^*/x)| \le |x^*-x|/\min(x,x^*)$ for all $x, x^* > 0$, we have

$$\left|\frac{1}{n}\sum_{t=1}^{n}\log\left(\hat{\eta}_t^2/\eta_t^{*2}\right)\right| \le \frac{1}{n}\sum_{t=1}^{n}\frac{|\omega^*-\hat{\omega}_n+(\alpha^*-\hat{\alpha}_n)\epsilon_{t-1}^2|}{\underline{\omega}^*+\underline{\alpha}^*\epsilon_{t-1}^2} \to 0 \quad \text{a.s.}$$

with $\underline{\omega}^* > 0$ and $\underline{\alpha}^* > 0$. By the arguments used in Berkes, Horváth, and Kokoszka (2003, Lemma 2.3), we have $E|\epsilon_t|^{2s} < \infty$. Applying the ergodic theorem to $(\log\eta_t^{*2})$, it follows that

$$\Gamma^* = E\log\alpha^*\eta_t^{*2} = E\log\alpha^*\epsilon_t^2/(\omega^*+\alpha^*\epsilon_{t-1}^2) < 0.$$

The convergence of $\hat{\sigma}_u^2$ to a positive limit is obtained by arguments already used. □

References

Aue A., and L. Horváth (2009) Quasi-likelihood Estimation in Stationary and Nonstationary Autoregressive Models With Random Coefficients. *Preprint*

Berkes, I., Horváth, L. and P.S. Kokoszka (2003) GARCH processes: structure and estimation. *Bernoulli* 9, 201–227.

Billingsley P. (1995) *Probability and Measure*. John Wiley, New York.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* 31, 307–327.

Drost, F. C. and C. A. J. Klaassen (1997) Efficient estimation in semiparametric GARCH models. *Journal of Econometrics* 81, 193–221.

Drost, F.C., Klaassen, C.A.J. and B.J.M. Werker (1997) Adaptive estimation in time-series models. *Annals of Statistics* 25, 786–817.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica* 50, 987–1007.

Francq, C. and J.M. Zakoïan (2004) Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes. *Bernoulli* 10, 605–637.

Francq, C. and J-M. Zakoïan (2006a) Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22, 815–834.

Francq, C. and J.M. Zakoïan (2006b) On efficient inference in GARCH processes. *In: Bertail P, Doukhan P, Soulier P (eds) Statistics for dependent data*. Springer, New-York: 305–327

Goldie, C.M. and R.A. Maller (2000) Stability of perpetuities. *Annals of Probability* 28, 1195–1218.

Hörmann, S. (2008) Augmented GARCH sequences: dependence structure and asymptotics. *Bernoulli* 14, 543–561.

Jensen, S.T. and A. Rahbek (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* 72, 641–646.

Jensen, S.T. and A. Rahbek (2004b) Asymptotic inference for nonstationary GARCH. *Econometric Theory* 20, 1203–1226.

Klüppelberg, C., Lindner, A. and R. Maller (2004) A continuous time GARCH Process driven by a Lévy process: stationarity and second order behaviour. *Journal of Applied Probability* 41, 601–622.

Lee, S. and M. Taniguchi (2005) Asymptotic theory for ARCH-SM models: LAN and residual empirical processes. *Statistica Sinica* 15, 215–234.

Ling, S. and M. McAleer (2003) Adaptative estimation in nonstationary ARMA models with GARCH errors. *The Annals of Statistics* 31, 642–674.

Linton, O., Pan, J. and H. Wang (2009) Estimation for a non-stationary semi-strong GARCH(1,1) model with heavy-tailed errors. To appear in *Econometric Theory.*

Nelson, D.B. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.

van der Vaart, A.W. (1998) *Asymptotic statistics.* Cambridge University Press, United Kingdom.

BP 60149
59653 Villeneuve d'Ascq cedex
France
E-mail: christian.francq@univ-lille3.fr

15 Boulevard Gabriel Peri
92245 Malakoff Cedex
France
E-mail: zakoian@ensae.fr