Inference in non stationary asymmetric garch models

Christian Francq and Jean-Michel Zakoian

CREST, University Lille 3

1. March 2013

Inference in non stationary asymmetric garch models

Christian Francq, Jean-Michel Zakoïan

CREST and University Lille 3 (EQUIPPE)*

Abstract. This paper considers the statistical inference of the class of asymmetric power-transformed GARCH(1,1) models in presence of possible explosiveness. We study the explosive behavior of volatility when the strict stationarity condition is not met. This allows us to establish the asymptotic normality of the quasi-maximum likelihood estimator (QMLE) of the parameter, including the power but without the intercept, when strict stationarity does not hold. Two important issues can be tested in this framework: asymmetry and stationarity. The tests exploit the existence of a universal estimator of the asymptotic covariance matrix of the QMLE. By establishing the local asymptotic normality (LAN) property in this nonstationary framework, we can also study optimality issues.

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

Key words and phrases: GARCH models, Inconsistency of estimators, Local power of tests, Nonstationarity, Quasi Maximum Likelihood Estimation.

* We are grateful to the Agence Nationale de la Recherche (ANR), which supported this work via the Project ECONOM&RISK (ANR 2010 blanc 1804 03).

1. INTRODUCTION

Following more than twenty years of tremendous development of the theory of unit roots in linear time series models (see the seminal papers by Dickey and Fuller (1979), and Phillips and Perron (1988)), there has been, in the last decade, much interest in the statistical analysis of non linear time series models under non stationarity assumptions (see e.g. Karlsen and Tjøstheim (2001), Karlsen, Myklebust and Tjøstheim (2007), Ling and Li (2008), Aue and Horvàth (2011)). In the framework of GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models, Jensen and Rahbek (2004a, 2004b) were the first to establish an asymptotic theory for the quasi-maximum likelihood estimator (QMLE) of non-stationary GARCH(1,1), assuming that the intercept is fixed to an arbitrary value. Aknouche, Al-Eid and Hmeid (2011), Aknouche and Al-Eid (2012) studied the properties of weighted least-squares estimators. Francq and Zakoïan (2012) established the asymptotic properties of the standard QMLE of the complete parameter vector: they showed that, while the intercept cannot be consistently estimated, the QMLE of the remaining parameters is consistent (in the weak sense at the frontier of the stationarity region, and in the strong sense outside) and asymptotically normal with or without strict stationarity.

Financial series are well-known to present conditional asymmetry features, in the sense that large negative returns tend to have more impact on future volatilities than large positive returns of the same magnitude. This stylized fact, known as the leverage effect, was first documented by Black (1976), and led to various generalizations of the GARCH models of the first generation (see among others, Glosten, Jaganathan and Runkle (1993), Rabemananjara and Zakoïan (1993), Higgins and Bera (1992), Li and Li (1996), Francq and Zakoïan (2010)). Motivated by the Box-Cox transformation, Hwang and Kim (2004) introduced a power transformed ARCH model, and the GARCH extension was studied by Pan, Wang and Tong (2008). In this paper we consider an asymmetric power-transformed GARCH(1,1) model defined, for a given positive constant δ, by

$$(1.1)\qquad \begin{cases} \epsilon_t = h_t^{1/\delta}\eta_t \\ h_t = \omega_0 + \alpha_{0+}(\epsilon_{t-1}^+)^\delta + \alpha_{0-}(-\epsilon_{t-1}^-)^\delta + \beta_0 h_{t-1} \end{cases}$$

with initial values ϵ_0 and $h_0 \geq 0$, where $\omega_0 > 0$, $\alpha_{0+} \geq 0$, $\alpha_{0-} \geq 0$, $\beta_0 \geq 0$, and using the notation $x^+ = \max(x, 0), x^- = \min(x, 0)$. In this model, (η_t) is a sequence of independent and identically distributed (iid) variables such that

$$(1.2)\qquad E\eta_1^2 = 1 \quad \text{and} \quad P(\eta_1^2 = 1) < 1.$$

Most commonly used extensions of the standard GARCH of Engle (1982) and Bollerslev (1986) can be written in the form (1.1).

The first goal of the present paper is to derive a strict stationarity test in the framework of Model (1.1). In this model, strict stationarity is characterized by the negativity of the so-called top Lyapunov exponent (see Bougerol and Picard (1992)), which depends on the parameters (except ω) and the errors distribution. By deriving the asymptotic behavior of the QMLE of the top-Lyapunov exponent, under stationarity and non stationarity, a strict stationarity test can be derived. The second goal of the paper is to propose a test for the symmetry assumption in Model (1.1), namely $\alpha_{0+} = \alpha_{0-}$. Existing tests, to our knowledge, rely on the stationarity assumption. Our aim is to derive a test which can be used without bothering about stationarity.

The rest of the paper is organized as follows. In Section 2, we study the convergence of the volatility to infinity, in a model encompassing (1.1), when stationarity does not hold. Section 3 is devoted to the asymptotic properties of the QMLE. In Section 4, we consider strict stationarity testing and asymmetry testing. In Section 5, the LAN property is established and used to derive the local asymptotic power of the proposed tests. Local alternative allowing for an arbitrary rate of convergence with respect to ω_0 are considered. Optimality issues are discussed. Necessary and sufficient conditions on the noise density are derived for the tests to be uniformly locally asymptotically most powerful. Section 6 is devoted to the case where the power δ is unknown and is jointly estimated with the volatility coefficients. Proofs and technical lemmas are in Section 7. The possibility of extensions is discussed in Section 8. Several lemmas and proofs, along with a study of the finite sample performance of the stationarity and asymmetry tests and an empirical application, are included in appendix.

2. EXPLOSIVITY IN THE AUGMENTED GARCH(1,1)

In this section, we analyze the convergence of the volatility to infinity, for a class of augmented GARCH processes encompassing (1.1) and many GARCH(1,1) models introduced in the literature (see Hörmann, 2008). Given a sequence $(\xi_t)_{t\geq 0}$, let $(\epsilon_t)_{t\geq 1}$ be defined by

$$\left\{ \begin{array}{lcl} \epsilon_t & = & h_t^{1/\delta}\xi_t, \quad t = 1, 2, \ldots \\ h_t & = & \omega(\xi_{t-1}) + a(\xi_{t-1})h_{t-1} \end{array} \right. \tag{2.1}$$

where δ is a positive constant, $h_0 \geq 0$ is a given initial value, and the functions $\omega(\cdot)$ and $a(\cdot)$ satisfy $\omega : \mathbb{R} \to [\underline{\omega}, +\infty)$ and $a : \mathbb{R} \to [0, +\infty)$, for some $\underline{\omega} > 0$. When (ξ_t) is assumed to be a white noise, (ϵ_t) is called an augmented GARCH process. We purposely use a different notation for ξ_t in (2.1) and η_t in (1.1) because, for the moment, we only assume that (ξ_t) is stationary and ergodic. Define in $\mathbb{R} \cup \{+\infty\}$ the top Lyapunov exponent

$$\gamma = E \log a(\xi_1).$$

The following proposition is an extension of results proven for the standard GARCH(1,1) by Nelson (1990) and completed by Klüppelberg, Lindner, and Maller (2004), and Francq and Zakoïan (2012).

PROPOSITION 2.1. *For the process (ϵ_t) satisfying (2.1), the following properties hold.*

i) When $\gamma > 0$, $h_t \to \infty$ a.s. at an exponential rate: for any $\rho > e^{-\gamma}$,

$$\rho^t h_t \to \infty \quad \text{and, if } E|\log(\xi_1^2)| < \infty, \qquad \rho^t \epsilon_t^2 \to \infty \quad \text{a.s. as } t \to \infty.$$

ii) When $\gamma = 0$ and (ξ_t) is time reversible (i.e. for all k the distributions of $(\xi_t, \xi_{t-1}, \ldots, \xi_{t-k})$ and $(\xi_{t-k}, \ldots, \xi_{t-1}, \xi_t)$ are identical), the following convergences in probability hold as $t \to \infty$,

$$h_t \to \infty \quad \text{and, if } E|\log(\xi_1^2)| < \infty, \qquad \epsilon_t^2 \to \infty.$$

Moreover, if ψ is a decreasing bijection from $(0,\infty)$ to $(0,\infty)$, if $E\psi(h_1)<\infty$ (resp. $E\psi(\epsilon_1^2)<\infty$ and $E|\log(\xi_1^2)|<\infty$), then

$$\psi(h_t)\to 0 \quad \textit{(resp. } \psi(\epsilon_t^2)\to 0\textit{)} \quad \textit{in } L^1. \tag{2.2}$$

The main ideas of the proof are as follows. The a.s. convergence of h_t to infinity in the case $\gamma>0$ follows from the minoration $\log h_t\geq \log\underline{\omega}+\sum_{i=1}^{t-1}\log a(\xi_{t-i})$, and the fact that the latter sum is strictly increasing, in average, as t goes to infinity. The argument is in failure when $\gamma=0$, the expectation of the sum being equal to zero. The key argument in this case is that the sequence (h_t) is increasing *in distribution*. Indeed, taking $h_0=0$ we have $h_1=\omega(\xi_0)$ and $h_2=\omega(\xi_1)+a(\xi_0)\omega(\xi_0)\overset{d}{=}\omega(\xi_0)+a(\xi_1)\omega(\xi_1)>h_1$ under the reversibility assumption, and the same argument applies for any $t>0$.

In the rest of the paper, these results will be applied with $\xi_t=\eta_t$ to Model (1.1), for which the top Lyapunov exponent is given by

$$\gamma_0=E\log a_0(\eta_1),\qquad a_0(x)=\alpha_{0+}(x^+)^\delta+\alpha_{0-}(-x^-)^\delta+\beta_0.$$

3. ASYMPTOTIC PROPERTIES OF THE QMLE

We wish to estimate $\vartheta_0=(\alpha_{0+},\alpha_{0-},\beta_0)'$ from observations $\epsilon_t, t=1,\dots,n$, in the stationary and the explosive cases under mild assumption. Denote by $\theta=(\omega,\alpha_+,\alpha_-,\beta)'$ the parameter and define the QMLE as any measurable solution of

$$\hat\theta_n=(\hat\omega_n,\hat\alpha_{n+},\hat\alpha_{n-},\hat\beta_n)'=\arg\min_{\theta\in\Theta}\frac{1}{n}\sum_{t=1}^{n}\ell_t(\theta),\quad \ell_t(\theta)=\frac{\epsilon_t^2}{\sigma_t^2(\theta)}+\log\sigma_t^2(\theta), \tag{3.1}$$

where Θ is a compact subset of $(0,\infty)^4$ containing the true value $\theta_0=(\omega_0,\alpha_{0+},\alpha_{0-},\beta_0)'$, and $\sigma_t^\delta(\theta)=\omega+\alpha_+(\epsilon_{t-1}^+)^\delta+\alpha_-(-\epsilon_{t-1}^-)^\delta+\beta\sigma_{t-1}^\delta(\theta)$ for $t=1,\dots,n$ (with initial values for ϵ_0 and $\sigma_0^\delta(\theta)$). The rescaled residuals are defined by $\hat\eta_t=\eta_t(\hat\theta_n)$ where $\eta_t(\theta)=\epsilon_t/\sigma_t(\theta)$ for $t=1,\dots,n$.

Write $\vartheta=(\alpha_+,\alpha_-,\beta)'$ and let $\hat\vartheta_n=\left(\hat\alpha_{n+},\hat\alpha_{n-},\hat\beta_n\right)'$.

3.1 Consistency and asymptotic normality of $\hat\vartheta_n$

The following theorem extends, to the non stationary framework, results obtained for the stationary case (see Hamadeh and Zakoïan (2011) and the references therein), which we recall for convenience. We introduce the assumptions:

A1: The support of (η_t) contains at least 3 points and is not concentrated on the positive or the negative line.

A2: When t tends to infinity,

$$E\left\{1+\sum_{i=1}^{t-1}a_0(\eta_1)\dots a_0(\eta_i)\right\}^{-1}=o\left(\frac{1}{\sqrt{t}}\right).$$

Note that **A2**, which is only required in the case $\gamma_0=0$, is obviously satisfied in the degenerate case when $a(\eta_t)=1, a.s.$, since the expectation is then equal to $1/t$.

To handle initial values we introduce the following notation. For any asymptotically stationary process $(X_t)_{t\geq 0}$ let $E_\infty(X_t) = \lim_{t\to\infty} E(X_t)$ provided this limit exists. Let also $\overset{\circ}{\Theta}$ denote the interior of Θ.

THEOREM 3.1. *Let (1.1)-(1.2) and* **A1** *hold. Then the QMLE defined in (3.1) satisfies the following properties.*

i) **Stationary case.** *When $\gamma_0 < 0$, and $\beta < 1$ for all $\theta \in \Theta$,*

$$\hat{\theta}_n \to \theta_0, \qquad a.s.\ as\ n \to \infty.$$

If, in addition, $\kappa_\eta = E\eta_1^4 \in (1, \infty)$ and $\theta_0 \in \overset{\circ}{\Theta}$, we have

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\mathcal{J}^{-1}\right\}, \qquad as\ n \to \infty, \tag{3.2}$$

where

$$\mathcal{J} = \frac{4}{\delta^2} E_\infty \left(\frac{1}{\sigma_t^{2\delta}} \frac{\partial \sigma_t^\delta}{\partial \theta} \frac{\partial \sigma_t^\delta}{\partial \theta'}(\theta_0) \right). \tag{3.3}$$

ii) **Explosive case.** *When $\gamma_0 > 0$, if $P(\eta_1 = 0) = 0$,*

$$\hat{\vartheta}_n \to \vartheta_0, \qquad a.s.\ as\ n \to \infty.$$

If, in addition, $\kappa_\eta \in (1, \infty)$, $E|\log \eta_1^2| < \infty$ and $\theta_0 \in \overset{\circ}{\Theta}$,

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\mathcal{I}^{-1}\right\}, \tag{3.4}$$

as $n \to \infty$, where $\mathcal{I}$ is a positive definite matrix.

iii) **At the boundary of the stationarity region.** *When $\gamma_0 = 0$, if $P(\eta_1 = 0) = 0$, and $\forall \theta \in \Theta$, $\beta < \|1/a_0(\eta_1)\|_p^{-1}$ for some $p > 1$,*

$$\hat{\vartheta}_n \to \vartheta_0, \qquad in\ probability\ as\ n \to \infty.$$

If, in addition, $\theta_0 \in \overset{\circ}{\Theta}$, $\kappa_\eta \in (1, \infty)$, $E|\log \eta_1^2| < \infty$ and **A2** *is satisfied, then (3.4) holds.*

The key ideas of the proof can be summarized as follows. First, we note that $\hat{\theta}_n$ can be equivalently defined as the minimizer of $\frac{1}{n}\sum_{t=1}^n \{\ell_t(\theta) - \ell_t(\theta_0)\}$, where $\ell_t(\theta) - \ell_t(\theta_0)$ is a function of η_t^2 and the ratio $\sigma_t^\delta(\theta)/h_t$. While the numerator and the denominator explode to infinity as t increases, the ratio is close to a stationary process for t sufficiently large. For instance in the symmetric ARCH(1) case ($\alpha_+ = \alpha_- = \alpha$ and $\beta = 0$), we have $\sigma_t^\delta(\theta)/h_t \to \alpha/\alpha_0$, a.s. in the strictly explosive case (in probability in the case $\gamma = 0$). The situation is much more intricate when $\beta \neq 0$ but we can show that, when $\gamma > 0$,

$$\left| \frac{\sigma_t^\delta(\theta)}{h_t} - v_t(\vartheta) \right| \to 0 \text{ a.s. as } t \to \infty.$$

uniformly on some compact set included in Θ, where $(v_t(\vartheta))$ is a strictly stationary and ergodic process. The a.s. convergence is replaced by a L^p convergence in the case $\gamma = 0$. The consistency results are established by showing that the criterion in which $\sigma_t^\delta(\theta)/h_t$ is replaced by $v_t(\vartheta)$ produces an estimator which is consistent to ϑ_0. Similar arguments are used to prove the asymptotic normality results, but we now show that

$$\left\| \frac{1}{\sigma_t^\delta(\theta)} \frac{\partial \sigma_t^\delta}{\partial \vartheta}(\theta_0) - d_t \right\| \to 0 \text{ in } L^p \text{ as } t \to \infty,$$

for some strictly stationary and ergodic process d_t.

An explicit expression of $\mathcal{I}$ is given in appendix. To conclude the section, it can be noted that no asymptotically valid inference on ω_0 can be done in the nonstationary case (see Propositions 2.1 and 3.1 in Francq and Zakoïan (2012), denoted hereafter FZ, for the standard GARCH(1,1) model).

3.2 A universal estimator of the asymptotic variance of $\hat{\vartheta}_n$

In view of (3.2)-(3.3), when $\gamma_0 < 0$ the asymptotic distribution of the QMLE $\hat{\vartheta}_n$ of ϑ_0 (the parameter without ω_0) is given by

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\mathcal{I}_*^{-1}\right\}, \qquad \text{as } n \to \infty, \tag{3.5}$$

with

$$\mathcal{I}_* = \mathcal{J}_{\vartheta,\vartheta} - \mathcal{J}_{\vartheta,\omega}\mathcal{J}_{\omega,\omega}^{-1}\mathcal{J}_{\omega,\vartheta}, \tag{3.6}$$

$\mathcal{J}_{\omega,\omega} = \frac{4}{\delta^2} E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^\delta}{\partial\omega}\frac{\partial\sigma_t^\delta}{\partial\omega}(\theta_0)\right)$, $\mathcal{J}_{\vartheta,\vartheta} = \frac{4}{\delta^2} E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^\delta}{\partial\vartheta}\frac{\partial\sigma_t^\delta}{\partial\vartheta'}(\theta_0)\right)$ and $\mathcal{J}_{\omega,\vartheta} = \mathcal{J}_{\vartheta,\omega}' = \frac{4}{\delta^2} E_\infty\left(\frac{1}{h_t^2}\frac{\partial\sigma_t^\delta}{\partial\omega}\frac{\partial\sigma_t^\delta}{\partial\vartheta'}(\theta_0)\right)$. Letting

$$\widehat{\mathcal{J}}_{\vartheta,\vartheta} = \frac{4}{\delta^2}\frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma_t^{2\delta}(\hat{\theta}_n)}\frac{\partial\sigma_t^\delta}{\partial\vartheta}\frac{\partial\sigma_t^\delta}{\partial\vartheta'}(\hat{\theta}_n),$$

and defining $\widehat{\mathcal{J}}_{\vartheta,\omega}$, $\widehat{\mathcal{J}}_{\omega,\omega}$ and $\widehat{\mathcal{J}}_{\omega,\vartheta}$ accordingly, it can be shown that

$$\widehat{\mathcal{I}}_* = \widehat{\mathcal{J}}_{\vartheta,\vartheta} - \widehat{\mathcal{J}}_{\vartheta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\vartheta},$$

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

Theorem 3.2. *Let the assumptions required for the consistency results in Theorem 3.1 hold, assume $\kappa_\eta \in (1, \infty)$ 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 $\widehat{\mathcal{I}}_* \to \mathcal{I}_*$ a.s as $n \to \infty$.
ii) When $\gamma_0 > 0$, we have $\hat{\kappa}_\eta \to \kappa_\eta$ and $\widehat{\mathcal{I}}_* \to \mathcal{I}$ a.s.
*iii) When $\gamma_0 = 0$, we have $\hat{\kappa}_\eta \to \kappa_\eta$ and, if **A2** is satisfied, $\widehat{\mathcal{I}}_* \to \mathcal{I}$ in probability.*

In any case, $(\hat{\kappa}_\eta - 1)\widehat{\mathcal{I}}_*^{-1}$ is a consistent estimator of the asymptotic variance of the QMLE of ϑ_0.

It follows that asymptotically valid confidence intervals for the parameter ϑ_0 can be constructed without knowing if the underlying process is stationary or not. This theorem also has interesting applications for testing problems, which we now consider.

4. TESTING

In this section we consider testing stationarity and testing asymmetry.

4.1 Strict stationarity testing

Consider the strict stationarity testing problems

$$H_0 : \gamma_0 < 0 \qquad \text{against} \qquad H_1 : \gamma_0 \geq 0, \tag{4.1}$$

and

$$H_0 : \gamma_0 \geq 0 \qquad \text{against} \qquad H_1 : \gamma_0 < 0. \tag{4.2}$$

Let $\hat{\gamma}_n = \gamma_n(\hat{\theta}_n)$ be the empirical estimator of γ_0, with for any $\theta \in \Theta$,

$$\gamma_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} \log\left[\alpha_+ \left\{\eta_t^+(\theta)\right\}^\delta + \alpha_- \left\{-\eta_t^-(\theta)\right\}^\delta + \beta\right], \tag{4.3}$$

where $\eta_t(\theta) = \epsilon_t/\sigma_t(\theta)$. The following result shows that the asymptotic distribution of $\hat{\gamma}_n$ is particularly simple in the nonstationarity case.

Theorem 4.1. *Let $u_t = \log a_0(\eta_t) - \gamma_0$, and $\sigma_u^2 = Eu_t^2$. Then, under the assumptions of Theorem 3.1,*

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}\left(0, \sigma_\gamma^2\right) \qquad \text{as } n \to \infty \tag{4.4}$$

where

$$\sigma_\gamma^2 = \begin{cases} \sigma_u^2 + (\kappa_\eta - 1)\{a'\mathcal{J}^{-1}a - (1-\nu_1)^2\} & \text{when } \gamma_0 < 0, \\ \sigma_u^2 & \text{when } \gamma_0 \geq 0, \end{cases}$$

with $a = (0, \tilde{\nu}_{1,+}, \tilde{\nu}_{1,-}, \nu_1/\beta_0)'$ and

$$\tilde{\nu}_{1+} = E\left\{\frac{(\eta_1^+)^\delta}{a_0(\eta_1)}\right\}, \quad \tilde{\nu}_{1-} = E\left\{\frac{(-\eta_1^-)^\delta}{a_0(\eta_1)}\right\}, \quad \nu_1 = E\left\{\frac{\beta_0}{a_0(\eta_1)}\right\}.$$

Let $\hat{\sigma}_u^2$ be the empirical variance of $\log\left\{\hat{\alpha}_{n+}\left(\hat{\eta}_t^+\right)^\delta + \hat{\alpha}_{n-}\left(-\hat{\eta}_t^-\right)^\delta + \hat{\beta}_n\right\}$, for $t = 1, \ldots, n$. Under the assumptions of Theorem 4.1, it can be shown that $\hat{\sigma}_u^2$ is a weakly consistent estimator of σ_u^2. The statistics

$$T_n = \sqrt{n}\hat{\gamma}_n/\hat{\sigma}_u$$

is thus asymptotically $\mathcal{N}(0,1)$ distributed when $\gamma_0 = 0$. For the testing problem (4.1) (resp. (4.2)), at the asymptotic significance level $\underline{\alpha}$, this leads to consider the critical region

$$\mathrm{C}^{\mathrm{ST}} = \left\{T_n > \Phi^{-1}(1-\underline{\alpha})\right\} \qquad (\text{resp. } \mathrm{C}^{\mathrm{NS}} = \{T_n < \Phi^{-1}(\underline{\alpha})\}). \tag{4.5}$$

4.2 Asymmetry testing

It is of particular interest to test the existence of a leverage effect in stock market returns. In the framework of model (1.1), this testing problem is of the form

$$H_0 : \alpha_{0+} = \alpha_{0-} \qquad \text{against} \qquad H_1 : \alpha_{0+} \neq \alpha_{0-}. \tag{4.6}$$

Consider the test statistic for symmetry

$$T_n^{\mathrm{S}} := \frac{\sqrt{n}(\hat{\alpha}_{n+} - \hat{\alpha}_{n-})}{\hat{\sigma}_{T^{\mathrm{S}}}}, \qquad \hat{\sigma}_{T^{\mathrm{S}}} = \sqrt{(\hat{\kappa}_\eta - 1)\boldsymbol{e}'\widehat{\mathcal{I}}_*^{-1}\boldsymbol{e}}.$$

with $\boldsymbol{e}' = (1, -1, 0)$. The following result is a direct consequence of (3.4), (3.5) and Theorem 3.1.

Corollary 4.1. *Assume that $\theta_0 \in \overset{\circ}{\Theta}$ and the assumptions of Theorem 3.1 hold. For the testing problem (4.6), the test defined by the critical region*

$$\mathrm{C}^{\mathrm{S}} = \left\{ |T_n^{\mathrm{S}}| > \Phi^{-1}(1 - \underline{\alpha}/2) \right\} \tag{4.7}$$

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

We emphasize the fact that this test for symmetry does not require any stationarity assumption. The somewhat surprising output is that the usual Wald test, based on the asymptotic theory for the stationary case, also works in the non stationary situation.[1]

5. ASYMPTOTIC LOCAL POWERS

The section investigates the asymptotic behavior under local alternatives of the asymmetry test (4.7) and of the strict stationarity test (4.5). We first establish the LAN of the power-transformed GARCH 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).

5.1 LAN without stationarity constraint

Assume that η_t has a density f with third-order derivatives such that

$$\lim_{|y|\to\infty} yf(y) = 0 \quad \text{and} \quad \lim_{|y|\to\infty} y^2 f'(y) = 0, \tag{5.1}$$

and that, for some positive constants K and δ,

$$|y| \left| \frac{f'}{f}(y) \right| + y^2 \left| \left(\frac{f'}{f}\right)'(y) \right| + y^2 \left| \left(\frac{f'}{f}\right)''(y) \right| \leq K\left(1 + |y|^\delta\right), \tag{5.2}$$

[1]For instance in ARMA models, Wald tests on the parameters are not the same in the stationary and non stationary cases.

$$E\left|\eta_1\right|^{2\delta} < \infty. \tag{5.3}$$

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

$$\theta_n = \theta_0 + \boldsymbol{\tau}_n/\sqrt{n}, \tag{5.4}$$

where $(\boldsymbol{\tau}_n)$ is a bounded sequence of $\mathbb{R}^4$. 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, for standard GARCH, 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

$$\Lambda_{n,f}(\theta_n, \theta_0) = \boldsymbol{\tau}_n' S_{n,f}(\theta_0) - \frac{1}{2}\boldsymbol{\tau}_n' \mathfrak{I}_f \boldsymbol{\tau}_n + o_{P_{\theta_0}}(1), \tag{5.5}$$

where $S_{n,f}(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \mathfrak{I}_f\right\}$ under P_{θ_0} as $n \to \infty$. The following proposition shows that (5.5) holds regardless of γ_0.

Proposition 5.1. *When $\theta_0 \in \overset{\circ}{\Theta}$, under (5.1)-(5.3) we have the LAN property (5.5). When $\gamma_0 < 0$, we have $\mathfrak{I}_f = \frac{\iota_f}{4}\mathcal{J}$, where $\mathcal{J}$ is defined in (3.3). When $\gamma_0 \geq 0$, the Fisher information is the degenerate matrix*

$$\mathfrak{I}_f = \frac{\iota_f}{4}\begin{pmatrix} 0 & 0_3' \\ 0_3 & \mathcal{I} \end{pmatrix}, \tag{5.6}$$

where $\mathcal{I}$ is the positive definite matrix introduced in (3.4).

5.2 Near-global alternatives with respect to ω_0

We now show that, in the non stationary case, LAN continues to hold when the local alternative allows for an arbitrary rate of convergence with respect to ω_0. To this aim we assume that

$$\theta_n = \theta_0 + \upsilon_n \boldsymbol{e}_1 + \frac{\boldsymbol{\tau}_n}{\sqrt{n}} \tag{5.7}$$

where $\boldsymbol{e}_1 = (1,0,0,0)'$, $(\boldsymbol{\tau}_n)$ is as in (5.4), and (υ_n) is a deterministic sequence converging to zero. The next result shows that, in the non stationary case, (5.5) which was established under (5.4), continues to hold under the more general alternatives (5.7). For simplicity, take $\boldsymbol{\tau}_n = \boldsymbol{\tau} = (\tau_1, \tilde{\boldsymbol{\tau}}')'$ and $\tilde{\boldsymbol{\tau}}' = (\tau_2, \tau_3, \tau_4)$.

Proposition 5.2. *Let $\theta_0 \in \overset{\circ}{\Theta}$ with $\gamma_0 \geq 0$. Then, under (5.1)-(5.3) and (5.7), we have the LAN property*

$$\Lambda_{n,f}(\theta_n, \theta_0) \xrightarrow{d} \mathcal{N}\left(-\frac{\iota_f}{8}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}}, \frac{\iota_f}{4}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}}\right), \quad \textit{under } P_{\theta_0} \textit{ as } n \to \infty.$$

Note that this Gaussian law is the distribution of the log-likelihood ratio in the statistical model $\mathcal{N}\left\{\tilde{\boldsymbol{\tau}}, 4\mathcal{I}^{-1}/\iota_f\right\}$ of parameter $\tilde{\boldsymbol{\tau}}$, or equivalently in the statistical model $\mathcal{N}\left\{\iota_f\mathcal{I}\tilde{\boldsymbol{\tau}}/4, \iota_f\mathcal{I}/4\right\}$. To interpret this result in terms of convergence of statistical experiments (see van der Vaart (1998) for details), assume that $\upsilon_n = \upsilon\nu_n$ where $\upsilon \in \mathbb{R}$ and (ν_n) is a given sequence converging to zero as $n \to \infty$. Denoting by $\mathcal{T}$ a subset of $\mathbb{R}^4$ containing a neighborhood of $\mathbf{0}$, the so-called local experiments $\{L_{n,f}(\theta_0 + \upsilon\nu_n\boldsymbol{e}_1 + (0, \tilde{\boldsymbol{\tau}}')/\sqrt{n}), (\upsilon, \tilde{\boldsymbol{\tau}}') \in \mathcal{T}\}$ converge to the gaussian experiment $\{\mathcal{N}\left(\tilde{\boldsymbol{\tau}}, 4\mathcal{I}^{-1}/\iota_f\right), (\upsilon, \tilde{\boldsymbol{\tau}}') \in \mathcal{T}\}$.

Interestingly, the parameter υ vanishes in the limiting experiment. Consequently, in the limit experiment there exists no test on the parameter υ (except of trivial power equal to the level). On the other hand, the limit of any converging sequence of power functions in the local experiments is a power function in the Gaussian limit experiment, by the asymptotic representation theorem. We can conclude that there exists no test with a non trivial asymptotic power, for local alternatives on the parameter υ at the rate $1/\nu_n$. Given that the rate of convergence of ν_n to zero is arbitrary, the LAN approach shows that no asymptotically valid inference can be made on the parameter ω_0.[2]

5.3 Local asymptotic powers of the tests

The LAN property, with the help of Le Cam's third lemma, allows to easily compute local asymptotic powers of tests. In view of Theorem 4.1,

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

when θ_0 is such that $\gamma_0 = 0$. For $\boldsymbol{\tau}$ such that $\theta_0 + \boldsymbol{\tau}/\sqrt{n} \in \Theta$, we denote by $P_{n,\boldsymbol{\tau}}$ the distribution of the observations $(\epsilon_1, \dots, \epsilon_n)$ when the parameter is $\theta_0 + \boldsymbol{\tau}/\sqrt{n}$. 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. Let

$$a_{\boldsymbol{\tau}}(\eta_1) = \left(\alpha_{0+} + \frac{\tau_2}{\sqrt{n}}\right)\left(\eta_1^+\right)^\delta + \left(\alpha_{0-} + \frac{\tau_3}{\sqrt{n}}\right)\left(-\eta_1^-\right)^\delta + \beta_0 + \frac{\tau_4}{\sqrt{n}}.$$

Local alternatives for the C^{ST}-test (resp. the C^{NS}-test) are obtained for $\boldsymbol{\tau}$ such that $E\log a_{\boldsymbol{\tau}}(\eta_1) > 0$ (resp. $E\log a_{\boldsymbol{\tau}}(\eta_1) < 0$).

Proposition 5.3. *Under the assumptions of Theorem 3.1 and Proposition 5.1, the local asymptotic powers of the strict stationarity tests (4.5) are given by*

$$\lim_{n\to\infty} P_{n,\boldsymbol{\tau}}\left(\mathrm{C}^{\mathrm{ST}}\right) = \Phi\left\{c_f(\theta_0) - \Phi^{-1}(1-\underline{\alpha})\right\} \tag{5.8}$$

and, using the notations of Theorem 4.1,

$$\lim_{n\to\infty} P_{n,\boldsymbol{\tau}}\left(\mathrm{C}^{\mathrm{NS}}\right) = \Phi\left\{\Phi^{-1}(\underline{\alpha}) - c_f(\theta_0)\right\},$$

where

$$c_f(\theta_0) = \frac{(\tau_2\tilde{\nu}_{1+} + \tau_3\tilde{\nu}_{1-} + \tau_4\nu_1/\beta_0)\, E\log a_0(\eta_1)\left\{1 + \eta_1\frac{f'(\eta_1)}{f(\eta_1)}\right\}}{\delta\sigma_u(1-\nu_1)}.$$

[2]This is in accordance with the observation that, at least in the explosive case, the Fisher information with respect to ω_0 is bounded as n increases. A proof is available from the authors.

We now compute the local asymptotic power of the asymmetry test defined by (4.7). We thus consider a sequence of local parameters of the form $\theta_n = \theta_0 + \boldsymbol{\tau}/\sqrt{n}$ where $\theta_0 = (\omega_0, \alpha_0, \alpha_0, \beta_0)'$ and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3, \tau_4)'$ (with $\tau_2 \neq \tau_3$ under a local alternative). We denote by $P^{\mathrm{S}}_{n,\boldsymbol{\tau}}$ the distribution of the observations under the assumption that the parameter is θ_n.

Proposition 5.4. *Let the assumptions of Proposition 5.1 and Theorem 3.1 be satisfied. For testing (4.6), the test defined by the rejection region (4.7) has the local asymptotic power*

$$\begin{aligned}\lim_{n\to\infty} P^{\mathrm{S}}_{n,\boldsymbol{\tau}}\left(\mathrm{C}^{\mathrm{S}}\right) &= 1 - \Phi\left\{\Phi^{-1}\left(1-\frac{\alpha}{2}\right) - \frac{\tau_2-\tau_3}{\sigma_{T^{\mathrm{S}}}}\right\} \\ &\quad + \Phi\left\{-\Phi^{-1}\left(\frac{\alpha}{2}\right) - \frac{\tau_2-\tau_3}{\sigma_{T^{\mathrm{S}}}}\right\},\end{aligned}$$

where, recalling the notation $\boldsymbol{e}' = (1,-1,0)$,

$$\sigma^2_{T^{\mathrm{S}}} = \begin{cases} (\kappa_\eta - 1)\boldsymbol{e}'\mathcal{I}_*^{-1}\boldsymbol{e} & \text{when} \quad \gamma_0 < 0 \\ (\kappa_\eta - 1)\boldsymbol{e}'\mathcal{I}^{-1}\boldsymbol{e} & \text{when} \quad \gamma_0 \geq 0. \end{cases}$$

5.4 Optimality issues

We discuss, in this section, the optimality of the symmetry test defined in (4.7). Let $\theta_0 = (\omega_0, \alpha_0, \alpha_0, \beta_0)'$ be a parameter value corresponding to a symmetric GARCH. Assume that, at this point, $\gamma_0 \geq 0$. If $\gamma_0 < 0$, it suffices to replace $\mathcal{I}$ by $\mathcal{I}_*$ in the sequel. A sequence of local alternatives to this symmetric parameter is defined by $\theta_0 + \boldsymbol{\tau}/\sqrt{n}$ where $\boldsymbol{\tau}' = (\tau_1, \tau_2, \tau_3, \tau_4)'$ is such that $\tau_2 \neq \tau_3$. The relations (5.5)-(5.6) imply that

$$\Lambda_{n,f}(\theta_0 + \boldsymbol{\tau}/\sqrt{n}, \theta_0) \xrightarrow{d} \mathcal{N}\left(-\frac{\iota_f}{8}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}}, \frac{\iota_f}{4}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}}\right) \qquad \text{under } P_{\theta_0},$$

with $\tilde{\boldsymbol{\tau}} = (\tau_2, \tau_3, \tau_4)'$, which is the distribution of the log-likelihood ratio in the statistical model $\mathcal{N}\left\{\tilde{\boldsymbol{\tau}}, 4\mathcal{I}^{-1}/\iota_f\right\}$ of parameter $\tilde{\boldsymbol{\tau}}$. In other words, denoting by $\tilde{\mathcal{T}}$ a subset of $\mathbb{R}^3$ containing a neighborhood of $\mathbf{0}$, for any τ_1, the so-called local experiments $\left\{L_{n,f}(\theta_0 + (\tau_1, \tilde{\boldsymbol{\tau}}')/\sqrt{n}), \tilde{\boldsymbol{\tau}} \in \tilde{\mathcal{T}}\right\}$ converge to the gaussian experiment $\left\{\mathcal{N}\left(\tilde{\boldsymbol{\tau}}, 4\mathcal{I}^{-1}/\iota_f\right), \tilde{\boldsymbol{\tau}} \in \tilde{\mathcal{T}}\right\}$.

The asymmetry test (4.6) corresponds to the test

$$\boldsymbol{e}'\tilde{\boldsymbol{\tau}} = 0 \quad \text{against} \quad \boldsymbol{e}'\tilde{\boldsymbol{\tau}} \neq 0$$

in the limiting experiment. The uniformly most powerful unbiased (UMPU) test based on $\boldsymbol{X} \sim \mathcal{N}\left(\tilde{\boldsymbol{\tau}}, 4\mathcal{I}^{-1}/\iota_f\right)$ is the test of rejection region

$$C = \left\{|\boldsymbol{e}'\boldsymbol{X}|/\sqrt{4\boldsymbol{e}'\mathcal{I}^{-1}\boldsymbol{e}/\iota_f} > \Phi^{-1}(1-\underline{\alpha}/2)\right\}.$$

This UMPU test has the power

$$P_{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}}(C) = 1 - \Phi\left\{\Phi^{-1}\left(1-\frac{\alpha}{2}\right) - c_{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}}\right\} + \Phi\left\{-\Phi^{-1}\left(\frac{\alpha}{2}\right) - c_{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}}\right\}, \tag{5.9}$$

with $c_{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}} = \frac{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}\sqrt{\iota_f}}{2\sqrt{\boldsymbol{e}'\mathcal{I}^{-1}\boldsymbol{e}}}$. A test of (4.6) whose level converges to $\underline{\alpha}$, which is asymptotically unbiased, and whose power converges to the bound in (5.9) will be called asymptotically locally UMPU.

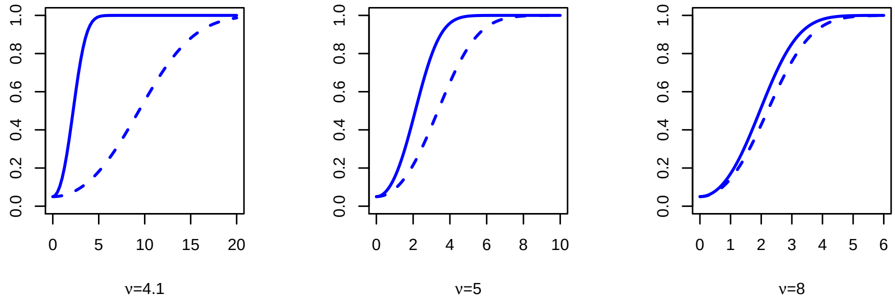

FIGURE 1. *Optimal asymptotic power (5.9) (in full line) and local asymptotic power of the asymmetry test (4.7) (in dotted line) when η_t follows a standardized Student distribution with ν degrees of freedom. The horizontal axis correspond to the local parameter $\mathbf{e}'\boldsymbol{\tau}$.*

PROPOSITION 5.5. *Under the assumptions of Proposition 5.3, the test (4.7) is asymptotically locally UMPU for the testing problem (4.6) if and only if the density of η_t has the form*

$$f(y) = \frac{a^a}{\Gamma(a)} e^{-ay^2} |y|^{2a-1}, \quad a > 0, \quad \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt. \tag{5.10}$$

A figure displaying the density (5.10) for different values of a is in appendix. Note that the gaussian density is obtained for $a = 1/2$. The result was expected because the C$^{\mathrm{S}}$-test is based on the QMLE of θ_0, and the QMLE is obviously efficient in the gaussian case. It can be shown that when the distribution of η_t is of the form (5.10), the MLE does not depend on a. The QMLE is then equal to the MLE, which makes obvious the "if part" of Proposition 5.5. The "only if" part of the proposition shows that there is necessarily an efficiency loss when the test is not based on the MLE of θ_0.

This point is illustrated by Figure 1, in which the local asymptotic power of the asymmetry test (in dotted lines) is compared to the optimal asymptotic power given by (5.9). In this figure, the noise η_t is assumed to satisfy a Student distribution with $\nu > 2$ degrees of freedom, standardized in such a way that $E\eta_t^2 = 1$. The parameters of the model under the null are $\alpha_{0+} = \alpha_{0-} = 0.2$, $\beta_0 = 0.9$ and $\delta = 1$, which corresponds to a nonstationary model with $\gamma_0 = 0.045$. In the figure, it can be seen that the local asymptotic power is far from the optimal power when ν is small, but, as expected, the discrepancy decreases as ν increases.

6. ESTIMATION WHEN THE POWER δ IS UNKNOWN

In this section, we consider the case where the power δ, now denoted δ_0, is unknown and is jointly estimated with θ_0. We rewrite the vector of parameters as $\zeta := (\delta, \theta')'$, which is assumed to belong to a compact parameter space $\Upsilon \subset (0,\infty)^2 \times [0,\infty)^3$. The true parameters value is denoted by $\zeta_0 := (\delta_0, \theta_0')'$. A QMLE of ζ is defined as any measurable solution $\hat{\zeta}_n$ of

$$\hat{\zeta}_n = (\hat{\delta}_n, \hat{\theta}_n')' = \arg\min_{\zeta \in \Upsilon} \frac{1}{n} \sum_{t=1}^{n} \ell_t(\zeta), \quad \ell_t(\zeta) = \frac{\epsilon_t^2}{\sigma_t^2(\zeta)} + \log \sigma_t^2(\zeta), \tag{6.1}$$

where

$$\sigma_t = \sigma_t(\zeta) = \left(\omega + \alpha_+(\epsilon_{t-1}^+)^\delta + \alpha_-(-\epsilon_{t-1}^-)^\delta + \beta\sigma_{t-1}^\delta(\zeta)\right)^{1/\delta}, \tag{6.2}$$

for $t = 1, \dots, n$ (with initial values for ϵ_0 and $\sigma_0(\zeta)$). The rescaled residuals are defined by $\hat{\eta}_t = \eta_t(\hat{\zeta}_n)$ where $\eta_t(\zeta) = \epsilon_t / \sigma_t(\zeta)$ for $t = 1, \dots, n$. For identifiability reasons, we need to slightly reinforce assumption **A1** as follows.

A3: The support of η_t contains at least three points of the same sign, and at least two points of opposite signs.

We also introduce the following technical assumption to handle the derivatives of ℓ_t with respect to the exponent δ.

A4: $\forall \zeta \in \Upsilon$, $\beta < \|1/a_0^2(\eta_1)\|_p^{-1}$ and $\||\eta_1|^\delta \log|\eta_1|\|_p < \infty$ for some $p > 1$.

For brevity, we only present results for the non stationary cases.

THEOREM 6.1. *Let (1.1)-(1.2) and* **A3** *hold. Then, the QMLE defined in (6.1) satisfies the following properties.*

i) **Explosive case.** *When $\gamma_0 > 0$, if $P(\eta_1 = 0) = 0$*

$$(\delta_n, \hat{\vartheta}_n') \to (\delta_0, \vartheta_0'), \qquad a.s.\ as\ n \to \infty.$$

If, in addition, $\kappa_\eta \in (1, \infty)$, $E|\log \eta_1^2| < \infty$, $\zeta_0 \in \overset{\circ}{\Upsilon}$, and **A4** *holds, then*

$$\sqrt{n}\left((\hat{\delta}_n, \hat{\vartheta}_n') - (\delta_0, \vartheta_0')\right)' \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\mathcal{I}_\delta^{-1}\right\}, \tag{6.3}$$

as $n \to \infty$, where $\mathcal{I}_\delta$ is a positive definite matrix (see Lemma D.4).

ii) **At the boundary of the stationarity region.** *When $\gamma_0 = 0$, if $P(\eta_1 = 0) = 0$, and $\forall \zeta \in \Upsilon$, $\beta < \|1/a_0(\eta_1)\|_p^{-1}$ for some $p > 1$,*

$$(\delta_n, \hat{\vartheta}_n') \to (\delta_0, \vartheta_0'), \qquad in\ probability\ as\ n \to \infty.$$

If, in addition, $\zeta_0 \in \overset{\circ}{\Upsilon}$, $\kappa_\eta \in (1, \infty)$, $E|\log \eta_1^2| < \infty$ and **A2** *and* **A4** *are satisfied, then (6.3) holds.*

The presence of parameter δ induces specific difficulties. It turns out that the derivative of the criterion with respect to δ involves the process $(\partial\sigma_t^\delta/\partial\delta - \log\sigma_t)$. A strictly stationary approximation to this process can then be obtained, but in a more complicated way than for the other parameters. To save place, the proofs of this section are given in appendix.

Obviously, stationarity and symmetry tests could be derived as in Sections 4 and 5. Other tests concerning the exponent δ (for instance testing the TARCH model ($\delta = 1$) against the GJR model ($\delta = 2$)) could be considered as well, but we leave this for further investigation.

7. PROOFS AND COMPLEMENTARY RESULTS

7.1 Proof of Proposition 2.1

Writing $\omega_t = \omega(\xi_t)$ and $a_t = a(\xi_t)$, we have, for all $t > 1$ and $1 \le k < t$,

$$h_t = \omega_{t-1} + \sum_{j=1}^{k} \omega_{t-j-1} \prod_{i=1}^{j} a_{t-i} + h_{t-k-1} \prod_{i=1}^{k+1} a_{t-i}. \tag{7.1}$$

We begin by showing *i)*. Since all the random variables involved in (7.1) are positive, $h_t \ge \underline{\omega} \prod_{i=1}^{t-1} a_{t-i}$. For any constant $\rho > e^{-\gamma}$, we thus have, a.s.

$$\liminf_{t\to\infty} \frac{1}{t} \log \rho^t h_t \quad \ge \quad \log\rho + \lim_{t\to\infty} \frac{1}{t}\left\{\log\underline{\omega} + \sum_{i=1}^{t-1} \log a_i\right\} = \log\rho + \gamma > 0,$$

by the ergodic theorem. It follows that $\log \rho^t h_t$, and hence $\rho^t h_t$, tend to $+\infty$ a.s as $n \to \infty$. The second convergence is shown right in the same way, arguing that $E|\log \xi_1^2| < \infty$ entails $\log \xi_t^2/t \to 0$ a.s. as $t \to \infty$.

To show *ii)*, first consider the case where $h_0 = 0$. Note that, for all t, the distribution of $h_t = h_t(\xi_0, \dots, \xi_{t-1})$ is equal to that of

$$h_t^* := h_t(\xi_t, \dots, \xi_1) = \omega_1 + \sum_{j=1}^{t-1} \omega_{j+1} \prod_{i=1}^{j} a_i. \tag{7.2}$$

Note that, contrary to (h_t), the sequence (h_t^*) increases with t. The Chung-Fuchs theorem applied to the random walk $\sum_{i=1}^t \log a_i$ entails that $\limsup_{t\to\infty} \prod_{i=1}^t a_i = +\infty$ a.s. It follows that $h_t^* \to +\infty$ as $t \to \infty$. We thus have $P(h_t \geq A) = P(h_t^* \geq A) \to 1$ for all $A > 0$, from which the first part of *ii)* easily follows. To prove the first convergence of (2.2), note that the dominated convergence theorem entails

$$E\psi(h_t) = \int_0^\infty P\left\{h_t^* < \psi^{-1}(u)\right\} du \to \int_0^\infty \lim_{t\to\infty} P\left\{h_t^* < \psi^{-1}(u)\right\} du = 0.$$

The second convergence is shown similarly. Now consider the case where the initial value is not equal to zero. It is clear from (7.1), with $k = t-1$, that h_t is an increasing function of h_0. So the convergences to infinity obtained when $h_0 = 0$, and the convergences in (2.2), hold a fortiori when $h_0 > 0$. □

7.2 Asymptotic behavior of the QMLE of ϑ_0

Define the $[0, \infty]$-valued process

$$v_t(\vartheta) = \sum_{j=1}^{\infty} \frac{\{\alpha_+(\eta_{t-j}^+)^\delta + \alpha_-(-\eta_{t-j}^-)^\delta\}}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta}{a_0(\eta_{t-k})}$$

with the convention $\prod_{k=1}^{j-1} = 1$ when $j \leq 1$. Let $\Theta_0 = \{\theta \in \Theta : \beta < e^{\gamma_0}\}$ and $\Theta_p = \{\theta \in [0,\infty)^4 : \beta < \|1/a_0(\eta_1)\|_p^{-1}\}$.

Lemma 7.1. *i) When $\gamma_0 > 0$, for any $\theta \in \Theta_0$ the process $v_t(\vartheta)$ is stationary and ergodic. Moreover, for any compact $\Theta_0^* \subset \Theta_0$,*

$$\sup_{\theta\in\Theta_0^*} \left| \frac{\sigma_t^\delta(\theta)}{h_t} - v_t(\vartheta) \right| \to 0 \text{ a.s. as } t \to \infty.$$

Finally, for any $\theta \notin \Theta_0$ it holds that $\sigma_t^\delta(\theta)/h_t \to \infty$ a.s.

ii) When $\gamma_0 = 0$, for any $\theta \in \Theta_p$ with $p \geq 1$, the process $v_t(\vartheta)$ is stationary and ergodic. Moreover, for any compact $\Theta_p^* \subset \Theta_p$,

$$\sup_{\theta\in\Theta_p^*} \left| \frac{\sigma_t^\delta(\theta)}{h_t} - v_t(\vartheta) \right| \to 0 \text{ in } L^p.$$

Proof. Assuming, with no generality loss, that $\sigma_0(\theta) = 0$, we have $\sigma_t^\delta(\theta) = \sum_{j=1}^t \beta^{j-1} z_{t-j}$ where $z_t = \omega + \alpha_+(\epsilon_t^+)^\delta + \alpha_-(-\epsilon_t^-)^\delta$ and

$$\frac{\sigma_t^\delta(\theta)}{h_t} = \sum_{j=1}^{t} \beta^{j-1} \left\{ \prod_{k=1}^{j} \frac{h_{t-k}}{h_{t-k+1}} \right\} \frac{z_{t-j}}{h_{t-j}}. \tag{7.3}$$

Noting that

$$\frac{h_{t-k}}{h_{t-k+1}} = \frac{h_{t-k}}{\omega_0 + a_0(\eta_{t-k})h_{t-k}} \leq \frac{1}{a_0(\eta_{t-k})}, \tag{7.4}$$

the rest of the proof follows from arguments similar to those used in the proof of Lemma A.1 in FZ. Therefore is it omitted. □

Lemma 7.2. *If $\theta \in \Theta_0$, we have* $\quad v_t(\vartheta) = 1,$ *a.s.* $\quad$ *iff* $\quad \vartheta = \vartheta_0$.

Proof. Straightforward algebra shows that

$$v_t(\vartheta)a_0(\eta_{t-1}) = \beta v_{t-1}(\vartheta) + \alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta. \tag{7.5}$$

Hence

$$\{v_t(\vartheta) - 1\}a_0(\eta_{t-1}) = \beta v_{t-1}(\vartheta) - \beta_0 + (\alpha_+ - \alpha_{0+})(\eta_{t-1}^+)^\delta + (\alpha_- - \alpha_{0-})(-\eta_{t-1}^-)^\delta.$$

It follows that $v_t(\vartheta) = 1 \quad$ a.s. iff

$$\beta - \beta_0 + (\alpha_+ - \alpha_{0+})(\eta_{t-1}^+)^\delta + (\alpha_- - \alpha_{0-})(-\eta_{t-1}^-)^\delta = 0.$$

Thus, if $\vartheta \neq \vartheta_0$, η_t takes at most two values of different signs, in contradiction with Assumption **A1**. The conclusion follows. □

Let $\underline{\omega} = \inf\{\omega \mid \theta \in \Theta\}$, $\underline{\alpha} = \inf\{\alpha_+, \alpha_- \mid \theta \in \Theta\}$, $\underline{\beta} = \inf\{\beta \mid \theta \in \Theta\}$, $\overline{\omega} = \sup\{\omega \mid \theta \in \Theta\}$, $\overline{\alpha} = \sup\{\alpha_+, \alpha_- \mid \theta \in \Theta\}$, $\overline{\beta} = \sup\{\beta \mid \theta \in \Theta\}$. Denote by K any constant whose value is unimportant and can change throughout the proofs. Let $\check{\Theta}$ be the compact set of the ϑ's such that $(\omega, \vartheta')' \in \Theta$.

Lemma 7.3. *Suppose that $P(\eta_t = 0) = 0$. Then, for any $k > 0$*

$$E \sup_{\vartheta \in \check{\Theta}} \left(\frac{1}{v_t(\vartheta)}\right)^k < \infty \quad \textit{and} \quad E \sup_{\theta \in \Theta} \left(\frac{h_t}{\sigma_t^\delta(\theta)}\right)^k < \infty.$$

Proof. Let $\varepsilon > 0$ such that $p(\varepsilon) := P(|\eta_t| \leq \varepsilon) \in [0, 1)$. If $|\eta_{t-1}| > \varepsilon$, since the sum $v_t(\vartheta)$ is greater than its first term, we have,

$$\frac{1}{v_t(\vartheta)} \leq \frac{a_0(\eta_{t-1})}{\alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta} \leq \frac{\max(\alpha_{0+}, \alpha_{0-})}{\underline{\alpha}} + \frac{\beta_0}{\underline{\alpha}\varepsilon^\delta} := K(\varepsilon).$$

Iterating this method, we can write

$$\sup_{\vartheta \in \check{\Theta}} \frac{1}{v_t(\vartheta)} \quad \leq \quad K(\varepsilon) \sum_{i=1}^\infty \mathbb{1}_{|\eta_{t-1}| \leq \varepsilon} \dots \mathbb{1}_{|\eta_{t-i+1}| \leq \varepsilon} \mathbb{1}_{|\eta_{t-i}| > \varepsilon} \left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^{i-1}.$$

where $\overline{a}_0(\varepsilon) = \max(\alpha_{0+}, \alpha_{0-})\epsilon^\delta + \beta_0$. It follows that, for any integer k,

$$E \sup_{\vartheta \in \check{\Theta}} \left(\frac{1}{v_t(\vartheta)}\right)^k \quad \leq \quad \{K(\varepsilon)\}^k \{1 - p(\varepsilon)\} \sum_{i=1}^\infty p(\varepsilon)^{i-1} \left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^{k(i-1)}.$$

Noting that $\lim_{\varepsilon\to 0} p(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0} \overline{a}_0(\varepsilon) = \beta_0$ we have $p(\varepsilon)\left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^k < 1$ for ε sufficiently small. The first result of the lemma is thus proven.

Similarly, we have for $|\eta_{t-1}| > \varepsilon$,

$$\frac{h_t}{\sigma_t^\delta(\theta)} \leq \frac{\omega_0}{\underline{\omega}} + \frac{\overline{\alpha}}{\underline{\alpha}} + \frac{\beta_0}{\underline{\alpha}\varepsilon^\delta} := H(\varepsilon),$$

and for $|\eta_{t-1}| \leq \varepsilon$ and $|\eta_{t-2}| > \varepsilon$,

$$\frac{h_t}{\sigma_t^\delta(\theta)} \leq \frac{\omega_0}{\underline{\omega}} + \frac{\overline{a}_0(\varepsilon)}{\underline{\beta}} H(\varepsilon).$$

More generally,

$$\begin{aligned}\sup_{\theta\in\Theta} \frac{h_t}{\sigma_t^\delta(\theta)} \leq & \sum_{i=1}^{\infty} \mathbb{1}_{|\eta_{t-1}|\leq\varepsilon} \dots \mathbb{1}_{|\eta_{t-i+1}|\leq\varepsilon} \mathbb{1}_{|\eta_{t-i}|>\varepsilon} \\ & \times \left(\frac{\omega_0}{\underline{\omega}} \sum_{j=0}^{i-2} \left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^j + \left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^{i-1} H(\varepsilon)\right).\end{aligned}$$

The conclusion follows by the same arguments as before. □

Proof of the consistency results in the cases *ii)* **and** *iii)* **of Theorem 3.1.** Note that $(\hat{\omega}_n, \hat{\vartheta}_n') = \arg\min_{\theta\in\Theta} Q_n(\theta)$, where $Q_n(\theta) = n^{-1}\sum_{t=1}^n \{\ell_t(\theta) - \ell_t(\theta_0)\}$. We have

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2 \left\{\left(\frac{h_t}{\sigma_t^\delta(\theta)}\right)^{2/\delta} - 1\right\} + \log\left(\frac{\sigma_t^\delta(\theta)}{h_t}\right)^{2/\delta} = O_n(\vartheta) + R_n(\theta)$$

where

$$O_n(\vartheta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2 \left\{\frac{1}{v_t^{2/\delta}(\vartheta)} - 1\right\} + \log v_t^{2/\delta}(\vartheta)$$

and

$$R_n(\theta) = \frac{1}{n}\sum_{t=1}^n \eta_t^2 \left\{\left(\frac{h_t}{\sigma_t^\delta(\theta)}\right)^{2/\delta} - \frac{1}{v_t^{2/\delta}(\vartheta)}\right\} + \log\left(\frac{\sigma_t^\delta(\theta)}{h_t v_t(\vartheta)}\right)^{2/\delta}.$$

It suffices to consider the case $\theta \in \Theta_0^*$ where Θ_0^* is an arbitrary compact subset of Θ_0, because by Lemma 7.1 *i)* $Q_n(\theta) \to \infty$ a.s. if $\theta \notin \Theta_0$. We have by stationarity and ergodicity of $v_t(\vartheta)$, a.s.

$$\lim_{n\to\infty} O_n(\vartheta) = E\left\{\frac{1}{v_1^{2/\delta}(\vartheta)} - 1 + \log v_1^{2/\delta}(\vartheta)\right\} \geq 0$$

because $\log x \leq x - 1$ for $x > 0$. The inequality is strict except when $v_1(\vartheta) = 1$ a.s. By Lemma 7.2 we thus have $E\{O_n(\vartheta)\} \geq 0$, with equality only if $\vartheta = \vartheta_0$.

By Lemma 7.3 we prove, as in FZ, that

$$\lim_{n\to\infty} \sup_{\theta\in\Theta_0^*} |R_n(\theta)| = 0 \quad \text{a.s. (resp. } \lim_{n\to\infty} \sup_{\theta\in\Theta_p^*} |R_n(\theta)| = 0 \quad \text{in } L^1), \tag{7.6}$$

when $\gamma_0 > 0$ (resp. $\gamma_0 = 0$) and Θ_0^*, Θ_p^* are defined in Lemma 7.1, which completes the proof. $\square$

We now need to introduce new $[0, \infty]$-valued processes. Let $a(\eta_t) = \alpha_+(\eta_t^+)^\delta + \alpha_-(-\eta_t^-)^\delta + \beta$ and

$$d_t^{\alpha+} = \sum_{j=1}^{\infty} \frac{(\eta_{t-j}^+)^\delta}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})}, \qquad d_t^{\alpha-} = \sum_{j=1}^{\infty} \frac{(-\eta_{t-j}^-)^\delta}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})}$$

$$d_t^{\beta} = \sum_{j=2}^{\infty} \frac{(j-1)\{\alpha_{0+}(\eta_{t-j}^+)^\delta + \alpha_{0-}(-\eta_{t-j}^-)^\delta\}}{\beta_0 a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})}.$$

Lemma 7.4. *Assume $\gamma_0 \geq 0$ and $E\eta_t^4 < \infty$. We have*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \vartheta}(\theta_0) \xrightarrow{d} \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{I}\} \qquad \text{as } n \to \infty,$$

where $\mathcal{I} = \frac{4}{\delta^2} E d_1 d_1'$ and $d_t' = \left(d_t^{\alpha+}, d_t^{\alpha-}, d_t^{\beta}\right)$. Moreover, $\mathcal{I}$ is non singular.

Proof. Since $E \log \beta_0 / a_0(\eta_1) < 0$, by the Cauchy root test, the processes $d_t^{\alpha+}, d_t^{\alpha-}$ and d_t^{β} are stationary and ergodic. Still assuming $\sigma_0^2 = 0$, we have

$$\frac{\partial \sigma_t^\delta}{\partial(\alpha_+, \alpha_-)}(\theta) = \sum_{j=1}^{t} \beta^{j-1}(\{\epsilon_{t-j}^+\}^\delta, \{-\epsilon_{t-j}^-\}^\delta), \qquad \frac{\partial \sigma_t^2}{\partial \beta}(\theta) = \sum_{j=2}^{t}(j-1)\beta^{j-2} z_{t-j}.$$

Thus, using a direct extension of (7.4)

$$\begin{aligned}
\frac{1}{\sigma_t^\delta(\theta_0)} \frac{\partial \sigma_t^\delta}{\partial(\alpha_+, \alpha_-)}(\theta_0) &= \sum_{j=1}^{t} \beta^{j-1} \left\{ \prod_{k=1}^{j} \frac{\sigma_{t-k}^\delta(\theta_0)}{\sigma_{t-k+1}^\delta(\theta_0)} \right\} \frac{\{(\epsilon_{t-j}^+)^\delta, (-\epsilon_{t-j}^-)^\delta\}}{\sigma_{t-j}^\delta(\theta_0)} \\
&\leq (d_t^{\alpha+}(\vartheta_0), d_t^{\alpha-}(\vartheta_0)), \\
\frac{1}{\sigma_t^\delta(\theta_0)} \frac{\partial \sigma_t^\delta}{\partial \beta}(\theta_0) &= \sum_{j=2}^{t}(j-1)\beta_0^{j-2} \left\{ \prod_{k=1}^{j} \frac{\sigma_{t-k}^\delta(\theta_0)}{\sigma_{t-k+1}^\delta(\theta_0)} \right\} \frac{z_{t-j}}{\sigma_{t-j}^\delta(\theta_0)} \\
&\leq d_t^{\beta}(\vartheta_0),
\end{aligned}$$

where the first inequality stands componentwise. Moreover, we have

$$0 \leq d_t^{\alpha+}(\vartheta_0) - \frac{1}{\sigma_t^\delta} \frac{\partial \sigma_t^\delta}{\partial \alpha^+}(\theta_0) \leq s_{t_0} + r_{t_0},$$

where

$$\begin{aligned}
s_{t_0} &= \sum_{j=1}^{t_0} \frac{(\eta_{t-j}^+)^\delta}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})} - \frac{(\epsilon_{t-j}^+)^\delta}{\beta_0 \sigma_{t-j}^\delta(\theta_0)} \prod_{k=1}^{j} \frac{\beta_0 \sigma_{t-k}^\delta(\theta_0)}{\sigma_{t-k+1}^\delta(\theta_0)}, \\
r_{t_0} &= \sum_{j=t_0+1}^{\infty} \frac{(\eta_{t-j}^+)^\delta}{a_0(\eta_{t-j})} \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})}.
\end{aligned}$$

For all $p \geq 1$, $\|r_{t_0}\|_p \to 0$ as $t_0 \to \infty$ because $\|\beta_0/a_0(\eta_1)\|_p < 1$ and $\|(\eta_1^+)^\delta/a_0(\eta_1)\|_p < 1/\alpha_{0+}$. Since, in addition, $\|\beta_0\sigma_{t-1}^\delta(\theta_0)/\sigma_t^\delta(\theta_0)\|_p < 1$, and

$$\left\| \frac{\beta_0}{a_0(\eta_{t-1})} - \frac{\beta_0\sigma_{t-1}^\delta(\theta_0)}{\sigma_t^\delta(\theta_0)} \right\|_p = \left\| \frac{\beta_0\omega_0}{a_0(\eta_{t-1})\sigma_t^\delta(\theta_0)} \right\|_p \to 0$$

as $t \to \infty$ by the dominated convergence theorem, $s_{t_0} = s_{t_0}(t)$ converges to 0 in L^p as $t \to \infty$. The same derivations hold true when $d_t^{\alpha+}$ is replaced by $d_t^{\alpha-}$ and d_t^β. Therefore, $d_t^{\alpha+}, d_t^{\alpha-}$ and d_t^β have moments of any order, and

$$\left\| \frac{1}{\sigma_t^\delta} \frac{\partial \sigma_t^\delta}{\partial \vartheta}(\theta_0) - d_t \right\| \to 0 \tag{7.7}$$

in L^p for any $p \geq 1$.

Using (7.7) and the ergodic theorem, we thus have, as $n \to \infty$,

$$\mathrm{Var}\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial}{\partial\vartheta}\ell_t(\theta_0) = \frac{4}{\delta^2}\frac{\kappa_\eta - 1}{n}\sum_{t=1}^n E(d_t d_t') + o(1) \to (\kappa_\eta - 1)\mathcal{I}.$$

Moreover, it can be shown as in FZ that the Lindeberg condition is satisfied, allowing to apply the Lindeberg central limit theorem for martingale differences (see Billingsley, 1995, p. 476).

Now we show that $\mathcal{I}$ is nonsingular. Suppose there exists $x = (x_1, x_2, x_3)' \in \mathbb{R}^3$ such that $x'\mathcal{I}x = 0$. Then we get $x'd_t = 0$, that is,

$$\sum_{j=1}^\infty \left(x_1 \frac{(\eta_{t-j}^+)^\delta}{a(\eta_{t-j})} + x_2 \frac{(-\eta_{t-j}^-)^\delta}{a(\eta_{t-j})} + x_3(j-1)\frac{\alpha_+(\eta_{t-j}^+)^\delta + \alpha_-(-\eta_{t-j}^-)^\delta}{\beta a(\eta_{t-j})} \right) \times \prod_{k=1}^{j-1} \frac{\beta}{a(\eta_{t-k})} = 0, \quad a.s.$$

It follows that $x_1(\eta_{t-1}^+)^\delta + x_2(-\eta_{t-1}^-)^\delta = z_{t-2}$, $\quad a.s.$ where z_{t-2} is a measurable function of the η_{t-j} with $j > 1$. Because η_{t-1} is independent of z_{t-2}, this variable must be a.s. constant. In view of Assumption **A1**, this entails $x_1 = x_2 = 0$ and then $x_3 = 0$. Therefore, $\mathcal{I}$ is nonsingular. $\square$

LEMMA 7.5. *Let ϖ be an arbitrary compact subset of $[0, \infty)$. Assume that $E\log\eta_1^2 < \infty$. When $\gamma_0 > 0$ we have, a.s.*

$$\sum_{t=1}^\infty \sup_{\theta\in\Theta_0}\left|\frac{\partial}{\partial\omega}\ell_t(\theta)\right| < \infty, \quad \sum_{t=1}^\infty \sup_{\theta\in\Theta_0}\left\|\frac{\partial^2}{\partial\omega\partial\theta}\ell_t(\theta)\right\| < \infty,$$

$$\sup_{\omega\in\varpi}\left|\frac{1}{n}\sum_{t=1}^n \frac{\partial^2\ell_t(\omega,\vartheta_0)}{\partial\theta_{i+1}\partial\theta_{j+1}} - \mathcal{I}_{ij}\right| = o(1) \quad \text{for all } i,j \in \{1,2,3\},$$

$$\frac{1}{n}\sum_{t=1}^n \sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}\ell_t(\theta)\right| = O(1) \quad \text{for all } i,j,k \in \{2,3,4\}.$$

When $\gamma_0 = 0$ we have, for all $i,j,k \in \{2,3,4\}$,

$$\sup_{\omega\in\varpi}\left|\frac{1}{n}\sum_{t=1}^n \frac{\partial^2\ell_t(\omega,\alpha_0,\beta_0)}{\partial\theta_{i+1}\partial\theta_{j+1}} - \mathcal{I}_{ij}\right| = o_P(1), \tag{7.8}$$

$$\frac{1}{n}\sum_{t=1}^n \sup_{\theta\in\Theta_4}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}\ell_t(\theta)\right| = O_P(1). \tag{7.9}$$

Proof. This is similar to that of Lemma A.5. in FZ, therefore is it omitted. □

Proof of the asymptotic normality in the case *ii)* **of Theorem 3.1.** An expansion of the criterion derivative gives

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\omega}\ell_t(\hat{\theta}_n) \\ 0 \end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0) + \mathcal{J}_n\sqrt{n}(\hat{\theta}_n - \theta_0) \tag{7.10}$$

where $\mathcal{J}_n$ is a 4×4 matrix whose elements have the form

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

where $\theta_i^* = (\omega_i^*, \alpha_{i+}^*, \alpha_{i-}^*, \beta_i^*)'$ is between $\hat{\theta}_n$ and θ_0. Moreover, it can be shown that, for $i,j = 1,2,3$,

$$\mathcal{J}_n(i+1,1) = o(1/\sqrt{n}), \quad \mathcal{J}_n(i+1,j+1) \to \mathcal{I}(i,j) \quad \text{a.s.} \tag{7.11}$$

The conclusion follows from the last rows of (7.10) and Lemma 7.4. □

Proof of the asymptotic normality in the case *iii)* **of Theorem 3.1.** Note that (7.10) continues to hold. In view of (7.8)-(7.9), we have

$$\mathcal{J}_n(i+1,j+1) \to \mathcal{I}(i,j) \quad \text{in probability as } n\to\infty.$$

To conclude, by the arguments used in the case *ii)*, it suffices to show that,

$$\text{for } i = 2,3,4, \quad E|\mathcal{J}_n(i,1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \to 0 \quad \text{as } n\to\infty. \tag{7.12}$$

Noting that

$$\frac{1}{\sigma_t^\delta(\theta)}\sum_{j=1}^{t}\beta^{j-1}(\epsilon_{t-j}^+)^\delta \le \frac{1}{\alpha_+}, \tag{7.13}$$

and $\beta_2^* < 1$ for n large enough, and using the compactness of Θ, we obtain

$$\begin{aligned} &|\mathcal{J}_n(2,1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \\ \le\ & \frac{K}{\sqrt{n}}\sum_{t=1}^{n}\left(\frac{2h_t^{2/\delta}\eta_t^2}{\sigma_t^2(\theta_2^*)} + 1\right)\frac{\left\{\sum_{j=1}^{t}(\beta_2^*)^{j-1}(\epsilon_{t-j}^+)^\delta\right\}\left\{\sum_{j=1}^{t}(\beta_2^*)^{j-1}\right\}}{\sigma_t^{2\delta}(\theta_2^*)} \\ \le\ & \frac{K}{\sqrt{n}}\sum_{t=1}^{n}\left(\frac{2h_t^{2/\delta}\eta_t^2}{\sigma_t^2(\theta_2^*)} + 1\right)\frac{h_t}{\sigma_t^\delta(\theta_2^*)}\frac{1}{h_t}. \end{aligned}$$

Hence, by Lemma 7.3 and Hölder's inequality

$$E|\mathcal{J}_n(2,1)\sqrt{n}(\hat{\omega}_n - \omega_0)| \le \frac{K}{\sqrt{n}}\sum_{t=1}^{n}E\frac{1}{h_t^{1+\tau}},$$

for any $\tau > 0$. The same bound is obtained when $\mathcal{J}_n(2,1)$ is replaced by $\mathcal{J}_n(3,1)$ and $\mathcal{J}_n(4,1)$. Moreover,

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

Hence

$$\frac{1}{h_t^{1+\tau}} \le \frac{1}{\omega_0^{1+\tau}(1 + Z_{t-1} + Z_{t-1}Z_{t-2} + \cdots + Z_{t-1}\ldots Z_1)}$$

By Assumption **A2**, the conclusion follows. □

Proof of Theorem 3.2. It is displayed in appendix.

7.3 Stationarity test

Proof of Theorem 4.1. In the stationary case $\gamma_0 < 0$, standard arguments show that

$$\hat{\gamma}_n = \gamma_n(\theta_0) + \frac{\partial \gamma_n(\theta_0)}{\partial \theta'}(\hat{\theta}_n - \theta_0) + o_P(n^{-1/2}), \tag{7.14}$$

with

$$\begin{aligned} \frac{\partial \gamma_n(\theta_0)}{\partial \theta} &= \frac{-1}{n}\sum_{t=1}^{n} \frac{1}{a_0(\eta_t)} \left[\{a_0(\eta_t) - \beta_0\} \frac{1}{h_t} \frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} - \begin{pmatrix} 0 \\ (\eta_t^+)^\delta \\ (-\eta_t^-)^\delta \\ 1 \end{pmatrix} \right] \\ &= -\Psi + o_P(1), \end{aligned} \tag{7.15}$$

where $\Psi = (1-\nu_1)\Omega - a$ and $\Omega = E_\infty \frac{1}{h_t}\frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta}$. Moreover the QMLE satisfies

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -\mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1-\eta_t^2) \frac{2}{\delta h_t} \frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} + o_P(1). \tag{7.16}$$

In view of (7.14), (7.15) and (7.16), we have

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t + \Psi'\mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1-\eta_t^2) \frac{2}{\delta h_t} \frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} + o_P(1).$$

Note that

$$\mathrm{Cov}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t,\ \Psi'\mathcal{J}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (1-\eta_t^2) \frac{2}{\delta h_t} \frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} \right) = \frac{2c}{\delta} \Omega' \mathcal{J}^{-1} \Psi,$$

where $c = \mathrm{Cov}(u_t, 1-\eta_t^2)$. The Slutsky lemma and the central limit theorem for martingale differences thus entail

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}\left(0, \sigma_u^2 + 4\frac{c}{\delta}\Omega'\mathcal{J}^{-1}\Psi + (\kappa_\eta - 1)\Psi'\mathcal{J}^{-1}\Psi\right).$$

Now let $\overline{\theta}_0 = (\omega_0, \alpha_{0+}, \alpha_{0-}, 0)'$. Noting that $\overline{\theta}_0' \partial \sigma_t^\delta(\theta_0)/\partial \theta = h_t$ almost surely, we have

$$E\left\{ \frac{1}{h_t}\frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} \left(1 - \frac{1}{h_t}\frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta'} \overline{\theta}_0 \right) \right\} = 0,$$

which entails $\frac{\delta^2}{4}\mathcal{J}\overline{\theta}_0 = \Omega$ and $\Omega'\mathcal{J}^{-1}\Omega = \frac{\delta^2}{4}$. It follows that

$$\Omega'\mathcal{J}^{-1}\Psi = (1-\nu_1)\frac{\delta^2}{4} - \frac{\delta^2}{4}\overline{\theta}_0' a = \frac{\delta^2}{4}(1 - \nu_1 - \alpha_{0+}\tilde{\nu}_{1+} - \alpha_{0-}\tilde{\nu}_{1-}) = 0.$$

We also have $\Psi'\mathcal{J}^{-1}\Psi = a'\mathcal{J}^{-1}a - (1-\nu_1)^2$, which completes the proof of the asymptotic distribution (4.4) in the case $\gamma_0 < 0$.

Now consider the case $\gamma_0 \geq 0$. Let θ_n^* be a sequence such that $\|\theta_n^* - \theta_0\| \leq \|\hat{\theta}_n - \theta_0\|$. By Proposition 2.1 (using Assumption **A2** when $\gamma_0 = 0$), we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{1}{\sigma_t^\delta(\theta_n^*)} \frac{\partial \sigma_t^\delta(\theta_n^*)}{\partial \omega} = o(1), \text{ a.s. (resp. in probability) as } n \to \infty$$

when $\gamma_0 > 0$ (resp. when $\gamma_0 = 0$). It can be deduced that, under the same conditions, $\sqrt{n}\frac{\partial^2\gamma_n(\theta_n^*)}{\partial\omega\partial\theta} = o(1)$, and $\sqrt{n}(\hat{\theta}-\theta_0)'\frac{\partial^2\gamma_n(\theta_n^*)}{\partial\theta\partial\theta'}(\hat{\theta}-\theta_0) = o(1)$, which entails that (7.14) still holds. The previous arguments show that (7.15) holds with

$$\Omega = E\begin{pmatrix} 0 \\ d_t^{\alpha+}(\theta_0) \\ d_t^{\alpha-}(\theta_0) \\ d_t^{\beta}(\theta_0) \end{pmatrix} = \frac{1}{1-\nu_1}\begin{pmatrix} 0 \\ \tilde{\nu}_{1+} \\ \tilde{\nu}_{1-} \\ \nu_1/\beta \end{pmatrix} \quad \text{and} \quad \Psi = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The conclusion follows. □

7.4 Asymptotic local powers

Proof of Proposition 5.1. The LAN of GARCH models has already been established in the stationary case (see Drost and Klaassen (1997), Lee and Taniguchi (2005)). The non stationary case will be studied under more general assumptions in the proof of Proposition 5.2. □

Proof of Proposition 5.2. Let the functions

$$g_1(y) = 1 + y\frac{f'}{f}(y) \quad \text{and} \quad g_2(y) = 1 + 2y\frac{f'}{f}(y) + y^2\left(\frac{f'}{f}\right)'(y).$$

Introduce also the notations

$$\Delta_{1,t}(\theta) = \frac{1}{\sigma_t(\theta)}\frac{\partial^2\sigma_t(\theta)}{\partial\theta\partial\theta'}, \quad \Delta_{2,t}(\theta) = \frac{1}{\delta^2\sigma_t^{2\delta}(\theta)}\frac{\partial\sigma_t^\delta(\theta)}{\partial\theta}\frac{\partial\sigma_t^\delta(\theta)}{\partial\theta'}.$$

A Taylor expansion of $\theta_n \mapsto \Lambda_{n,f}(\theta_n,\theta_0)$ around θ_0 yields

$$\Lambda_{n,f}(\theta_n,\theta_0) = \boldsymbol{\tau}'S_{n,f}(\theta_0) - \frac{1}{2}\boldsymbol{\tau}'\mathfrak{I}_n(\theta_n^*)\boldsymbol{\tau} + \mathcal{R}_n, \tag{7.17}$$

where θ_n^* is between θ_0 and θ_n,

$$S_{n,f}(\theta_0) = \frac{-1}{\sqrt{n}}\sum_{t=1}^n g_1(\eta_t)\frac{1}{\delta h_t}\frac{\partial\sigma_t^\delta(\theta_0)}{\partial\theta}, \tag{7.18}$$

$$\mathfrak{I}_n(\theta) = \frac{1}{n}\sum_{t=1}^n g_1\left(\frac{\epsilon_t}{\sigma_t(\theta)}\right)\Delta_{1,t}(\theta) - \frac{1}{n}\sum_{t=1}^n g_2\left(\frac{\epsilon_t}{\sigma_t(\theta)}\right)\Delta_{2,t}(\theta),$$

and $\mathcal{R}_n$ is a reminder which is displayed below. As in the proof of Lemma 7.4, it can be seen that

$$S_{n,f}(\theta_0) = \frac{-1}{\delta\sqrt{n}}\sum_{t=1}^n g_1(\eta_t)d_t(\vartheta_0) + o_P(1), \quad d_t(\vartheta) = \begin{pmatrix} 0 \\ d_t^{\alpha+} \\ d_t^{\alpha-} \\ d_t^{\beta} \end{pmatrix}.$$

Using (5.1), it is easy to see that $Eg_1(\eta_1) = 0$, and thus $Eg_1^2(\eta_1) = \iota_f$. The Lindeberg central limit theorem for martingale differences then shows that

$$S_{n,f}(\theta_0) \xrightarrow{d} \mathcal{N}\left(0, \mathfrak{I}_f\right). \tag{7.19}$$

Turning to the second term of (7.17) we first note that, similarly to (7.7),

$$\left| \frac{1}{h_t} \frac{\partial \sigma_t^\delta(\theta_0)}{\partial \theta} - d_t(\vartheta_0) \right| \to 0 \quad \text{in } L^2 \text{ as} \quad t \to \infty.$$

Moreover, integrations by parts show that, under (5.1), $\int y^2 f''(y)dy = -2\int y f'(y)dy = 2$. It follows that $Eg_2(\eta_1) = -\iota_f$. We thus have, using $Eg_1(\eta_1) = 0$,

$$\mathfrak{I}_n(\theta_0) = \frac{1}{n}\sum_{t=1}^n \frac{-g_2(\eta_t)}{\delta^2} d_t(\vartheta_0)d_t'(\vartheta_0) + o_{P_{\theta_0}}(1) \to \mathfrak{I}_f \text{ in probability as } n \to \infty.$$

Next, it can be shown that, as $n \to \infty$,

$$\|\mathfrak{I}_n(\theta_n^*) - \mathfrak{I}_n(\theta_0)\| \to 0 \quad \text{in probability.} \tag{7.20}$$

Finally, we show the convergence in probability to zero of

$$\mathcal{R}_n = \upsilon_n \sum_{t=1}^n g_1(\eta_t)\frac{1}{\delta h_t}\frac{\partial \sigma_t^\delta(\theta_0)}{\partial \omega} - \upsilon_n\sqrt{n}\boldsymbol{\tau}'\mathfrak{I}_n(\theta_n^*)\boldsymbol{e}_1' - \frac{1}{2}n\upsilon_n^2\boldsymbol{e}_1\mathfrak{I}_n(\theta_n^*)\boldsymbol{e}_1'.$$

Noting that $\partial\sigma_t^\delta(\theta_0)/\partial\omega$ is constant and that $1/h_t$ converges to 0 in L^2 by Proposition 2.1, the first term in the right-hand side converges to zero in probability. The two other terms can be handled similarly. The conclusion then follows from (7.17)–(7.20). □

Proof of Proposition 5.3. For simplicity, write P instead of $P_{n,0}$. In the proof of Theorem 4.1 we have seen that

$$T_n = \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{u_t}{\sigma_u} + o_P(1).$$

By (5.5) and (7.18), it follows that under P

$$\begin{pmatrix} T_n \\ \Lambda_{n,f}(\theta_0 + \boldsymbol{\tau}/\sqrt{n}, \theta_0) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{ \begin{pmatrix} 0 \\ -\frac{\iota_f}{8}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}} \end{pmatrix}, \begin{pmatrix} 1 & c \\ c & \frac{\iota_f}{4}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}} \end{pmatrix} \right\},$$

where $\tilde{\boldsymbol{\tau}}' = (\tau_2, \tau_3, \tau_4)$, $c = -\frac{\boldsymbol{\tau}' Ed_1(\vartheta_0)}{\delta\sigma_u}Eu_1g_1(\eta_1) = c_f(\theta_0)$. Le Cam's third lemma (see *e.g.* van der Vaart, 1998, page 90) shows that

$$T_n \xrightarrow{d} \mathcal{N}\left(c_f(\theta_0), 1\right), \qquad \text{under } P_{n,\boldsymbol{\tau}}.$$

The conclusion easily follows. □

Proof of Proposition 5.4. First consider the case $\gamma_0 \geq 0$. In the proof of (3.4) it has been shown that

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) = -\frac{2}{\delta}\mathcal{I}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^n (1-\eta_t^2)d_t + o_P(1).$$

Moreover

$$\Lambda_{n,f}(\theta_0 + \boldsymbol{\tau}/\sqrt{n}, \theta_0) = -\frac{1}{\delta\sqrt{n}}\sum_{t=1}^n \left\{1 + \eta_t\frac{f'(\eta_t)}{f(\eta_t)}\right\}\tilde{\boldsymbol{\tau}}'d_t - \frac{\iota_f}{8}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}} + o_P(1)$$

with $\tilde{\boldsymbol{\tau}}' = (\tau_2, \tau_3, \tau_4)$. Note also that, since $E\eta_1^4 < \infty$ implies $y^3 f(y) \to 0$ as $|y| \to \infty$, we have

$$E(1-\eta_t^2)\left\{1+\eta_t \frac{f'(\eta_t)}{f(\eta_t)}\right\} = 2. \tag{7.21}$$

It follows that under $P^{\mathrm{S}}_{n,0}$

$$\begin{pmatrix} \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \\ \Lambda_{n,f}\left(\theta_0 + \frac{\boldsymbol{\tau}}{\sqrt{n}}, \theta_0\right) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left\{ \begin{pmatrix} 0_3 \\ \frac{-\iota_f}{8}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}} \end{pmatrix}, \begin{pmatrix} (\kappa_\eta - 1)\mathcal{I}^{-1} & \tilde{\boldsymbol{\tau}} \\ \tilde{\boldsymbol{\tau}}' & \frac{\iota_f}{4}\tilde{\boldsymbol{\tau}}'\mathcal{I}\tilde{\boldsymbol{\tau}} \end{pmatrix} \right\}.$$

Le Cam's third lemma (see *e.g.* van der Vaart, 1998, page 90) shows that

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} \mathcal{N}\left(\tilde{\boldsymbol{\tau}}, (\kappa_\eta - 1)\mathcal{I}^{-1}\right), \qquad \text{under } P^{\mathrm{S}}_{n,\boldsymbol{\tau}}.$$

We thus have shown that, in the case $\gamma_0 > 0$, $\hat{\vartheta}_n$ is a regular estimator of ϑ_0, in the sense that $\sqrt{n}\left(\hat{\vartheta}_n - \vartheta_0 - \tilde{\boldsymbol{\tau}}/\sqrt{n}\right)$ converges to a distribution which does not depend on $\tilde{\boldsymbol{\tau}}$. More precisely

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta_0 - \tilde{\boldsymbol{\tau}}/\sqrt{n}\right) \xrightarrow{d} \mathcal{N}\left(0, (\kappa_\eta - 1)\mathcal{I}^{-1}\right), \qquad \text{under } P^{\mathrm{S}}_{n,\boldsymbol{\tau}}. \tag{7.22}$$

When $\gamma_0 \leq 0$, the same arguments show that $\hat{\theta}_n$ is a regular estimator of θ_0

$$\sqrt{n}(\hat{\theta}_n - \theta_0 - \boldsymbol{\tau}/\sqrt{n}) \xrightarrow{d} \mathcal{N}\left(0, (\kappa_\eta - 1)\mathcal{J}^{-1}\right), \qquad \text{under } P^{\mathrm{S}}_{n,\boldsymbol{\tau}}.$$

In the case $\gamma_0 \leq 0$, we thus have (7.22) with $\mathcal{I}$ replaced by $\mathcal{I}_*$. Now, noting that $T_n^{\mathrm{S}} = \frac{\boldsymbol{e}'\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)}{\hat{\sigma}_{T^{\mathrm{S}}}}$, and by the same arguments, it follows that $T_n^{\mathrm{S}} \xrightarrow{d} \mathcal{N}(0,1)$, under $P^{\mathrm{S}}_{n,0}$ and more generally $T_n^{\mathrm{S}} \xrightarrow{d} \mathcal{N}(c_{\boldsymbol{\tau}}, 1)$, under $P^{\mathrm{S}}_{n,\boldsymbol{\tau}}$, where $c_{\boldsymbol{\tau}} = (0, 1, -1, 0)\boldsymbol{\tau}/\sigma_{T^{\mathrm{S}}}$. The conclusion easily follows. □

Proof of Proposition 5.5. Recall that we assume $\gamma_0 \geq 0$. The case $\gamma_0 < 0$ is obtained similarly, replacing $\mathcal{I}$ by $\mathcal{I}_*$. In view of Proposition 5.4 and (5.9), the C^{S}-test is asymptotically locally UMPU if and only if $c_{\boldsymbol{e}'\tilde{\boldsymbol{\tau}}} = \boldsymbol{e}'\tilde{\boldsymbol{\tau}}/\sigma_{T^{\mathrm{S}}}$, which is equivalent to $(\kappa_\eta - 1)\iota_f = 4$. By Corollary 1 in Francq and Zakoïan (2006), the solutions of this equation are given by (5.10). □

8. CONCLUDING REMARKS

Our framework covers the most widely used GARCH models in financial applications. Strictly stationary models are a special case but symmetry tests, and asymptotically valid confidence intervals for the parameters (except the intercept) can be built without this assumption. Surprisingly, while the asymptotic covariance matrix of the estimators is sensitive to the stationarity of the underlying process, an estimator which converges to the appropriate covariance matrix in every situation can be built. Nevertheless, if the interest is on the whole parameter vector, including the intercept, it is important to know whether the observations come from a stationary process or not. To this aim we derived strict stationarity/non stationarity tests which are very easy to implement.

Are our results extendable to higher-order models? It seems likely that for particular extensions involving *univariate* stochastic recurrence equations for the volatility, the asymptotic theory derived in this paper can also be established. One key problem, to show consistency, is to find stationary approximations to ϵ_{t-j}^2/h_t for $j = 1, 2, \ldots$. For an ARCH-type model of order q it suffices to take $j \leq q$. Consider standard symmetric GARCH models for simplicity. In the GARCH(1,1) case, the problem can be circumvented because

$$\frac{\epsilon_{t-j}^2}{h_t} = \frac{h_{t-1}}{h_t} \cdots \frac{h_{t-j}}{h_{t-j+1}} \eta_{t-j}^2$$

can be approximated by a stationary process, in view of

$$\frac{h_{t-i}}{h_{t-i+1}} \approx \frac{1}{\alpha \eta_{t-i}^2 + \beta} \qquad \text{for large } t.$$

To have a glimpse of the considerable difficulties encountered when the orders increase, consider a standard ARCH(2) model

$$\epsilon_t = \sqrt{h_t}\eta_t, \qquad h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2.$$

We have, neglecting ω and for t large enough $h_t/\epsilon_{t-1}^2 \approx X_t$ and $h_t/\epsilon_{t-2}^2 \approx Y_t$ where

$$X_t = \alpha_1 + \frac{\alpha_2}{X_{t-1}} \frac{1}{\eta_{t-1}^2}, \qquad Y_t = \alpha_2 + \alpha_1 \eta_{t-1}^2 X_{t-1}.$$

It is not difficult to show that the first stochastic recurrence equation admits a strictly stationary solution (X_t) under mild assumptions on the density of η_t, whatever the values of α_1 and α_2. From this solution we deduce a strictly stationary solution (Y_t) to the second equation. We thus believe that, at least for the consistency, the ARCH(2) model is amenable to a treatment similar to that developed in this paper, but at the price of increasing technical difficulties. To summarize, the ratio h_t/h_{t-1} is, for large t, close to i) a constant in the ARCH(1) case, ii) an iid process in the GARCH(1,1) case, iii) the stationary solution of a nonlinear times series model in the ARCH(2) case. Whether or not this approach based on the resolution of nonlinear stochastic recurrence equations could be extended, is left for further investigation.

References

Aknouche, A., Al-Eid, E.M. and A.M. Hmeid (2011) Offline and online weighted least squares estimation of nonstationary power ARCH processes. *Statistics and Probability Letters* 81, 1535–1540.

Aknouche, A. and E.M. Al-Eid (2012) Asymptotic inference of unstable periodic ARCH processes. *Statistical Inference for Stochastic Processes* 15, 61–79.

Aue A., and L. Horváth (2011) Quasi-likelihood Estimation in Stationary and Nonstationary Autoregressive Models With Random Coefficients. *Statistica Sinica*, 21, 973–999.

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.

Bougerol, P. and Picard, N. (1992) Strict stationarity of generalized autoregressive processes. *Annals of Probability* 20, 1714–1729.

Dickey, D. A. and W. A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

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

Francq, C. and J-M. Zakoïan (2010) *GARCH models: structure, statistical inference and financial applications.* Chichester: John Wiley.

Francq, C. and J.M. Zakoïan (2012) Strict stationarity testing and estimation of explosive and stationary GARCH models. *Econometrica* 80, 821–861.

Glosten, L.R., Jaganathan, R. and D. Runkle (1993) On the relation between the expected values and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.

Hamadeh, T., and J.M. Zakoïan (2011) Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH Processes. *Journal of Statistical Planning and Inference*, 141, 488–507.

Higgins, M.L. and A.K. Bera (1992) A class of nonlinear ARCH models. *International Economic Review*, 33, 137-158.

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

Hwang, S.Y. and T.Y. Kim (2004) Power transformation and threshold modeling for ARCH innovations with applications to test for ARCH structure. *Stochastic Processes and Their Applications* 110, 295–314.

Karlsen H.A., Myklebust, T. and D. Tjøstheim (2007) Nonparametric estimation in a nonlinear cointegration type model. *Annals of Statistics* 35, 252–299.

Karlsen H.A. and D. Tjøstheim (2001) Nonparametric estimation in null recurrent time series. *Annals of Statistics* 29, 372–416.

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.

Li, C.W. and W.K. Li (1996) On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics* 11, 253–274.

Ling, S. and D. Li (2008) Asymptotic inference for a nonstationary double AR(1) model. *Biometrika* 95, 257–263.

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

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

Pan, J., Wang, H., and H. Tong (2008) Estimation and tests for power-transformed and threshold GARCH models. *Journal of Econometrics*, 142, 352-378.

Phillips, P.C.B. and P. Perron (1988) Testing for a Unit Root in Time Series Regression *Biometrika*, 75, 335–346.

Rabemananjara, R. and J-M. Zakoïan (1993) Threshold ARCH Models and Asymmetries in Volatility. *Journal of Applied Econometrics* 8, 31–49.

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

Appendices

APPENDIX A: DENSITIES ENSURING LOCAL OPTIMALITY OF THE ASYMMETRY TEST

Figure 2 shows densities (5.10) for different values of a. For such densities, the test (4.7) is asymptotically locally UMPU for the testing problem (4.6).

APPENDIX B: NUMERICAL ILLUSTRATIONS

We first simulated $N = 1,000$ independent trajectories of size $n = 500$, $n = 2,000$ and $n = 4,000$ of an asymmetric GARCH(1,1) model, with a parameter of the form $\theta_0 = (0.1, \alpha_{0+}, \alpha_{0-}, 0.8)$ and the standardized Student distribution with 7 degrees of freedom for η_t. With such parameters, we have $\gamma_0 = 0$ for $\alpha_{0+} = \alpha_{0-} = 0.2575$ or for $\alpha_{0+} = 0.22$ and $\alpha_{0-} = 0.2971$, in particular. Table 1 studies the finite sample properties of the stationarity and asymmetry tests based respectively on the statistics T_n and T_n^S. The first rows of the table concern tests applied to a symmetric model $\alpha_{0+} = \alpha_{0-}$, and the last rows concern the asymmetric model.

The empirical size of a test of theoretical level $\alpha = 5\%$, over $N = 1,000$ independent replications of the null hypothesis, belongs to the interval [3.6%, 6.4%] with probability 95%. In view of this basic result, one can consider that the error of first kind of the asymmetry test is well controlled. Indeed, except for three values of α_{0-} when $n = 500$ (displayed in bold-face in the table), in the first part of Table 1 the relative frequency of rejection of H_0^S is always between the significance bounds 3.6% and 6.4%, regardless of the value of γ_0. The empirical frequency of rejection of the stationarity test under the null, displayed in the rows H_0^γ of the two gray columns, is also quite satisfactory when $n > 500$. Looking at the second part of Table 1, one can note that the asymmetry test behaves as expected under the alternative. Indeed the frequency of rejection of H_0^S increases when α_{0-} moves away from $\alpha_{0+} = 0.22$ and when the sample size n increases. As expected from the asymptotic study, as $n \to \infty$, the frequency of rejection of H_0^γ tends to zero when $\gamma_0 < 0$ (on the left of the gray columns) and tends to one when $\gamma_0 > 0$ (on the right of the gray columns).

We then considered applications to financial series. The strict stationarity and asymmetry tests have been applied to daily returns of various daily stock indices. We do not report the detailed results of the tests, since they always lead to the same conclusions for this kind of series. The strict stationarity can never be rejected because the value of T_n is always very small when computed on the returns of stock indices. Moreover, a significant negative value of T_n^S, indicating

TABLE 1

Finite sample behaviors of the stationarity test based on T_n and of the asymmetry test based on T_n^S: relative frequency of rejection of the hypotheses $H_0^\gamma : \gamma_0 < 0$ and $H_0^S : \alpha_{0+} = \alpha_{0-}$ at the nominal level $\underline{\alpha} = 5\%$ for a symmetric GARCH(1,1) (for which $\alpha_{0+} = \alpha_{0+} = 0.2575$ corresponds to $\gamma_0 = 0$) and for an asymmetric GARCH(1,1) (with $\alpha_{0+} = 0.22$, and for which $\alpha_{0-} = 0.2971$ corresponds to $\gamma_0 = 0$). The unexpected frequencies of rejection are displayed in bold.

Model with $\alpha_{0+} = \alpha_{0+}$								
					α_{0-}			
n	Null	0.18	0.20	0.22	0.2575	0.28	0.30	0.31
500	H_0^γ	0.0	0.0	0.1	**7.3**	28.4	62.4	76.2
	H_0^S	4.5	6.0	**6.6**	5.6	**8.0**	**7.7**	6.1
$2,000$	H_0^γ	0.0	0.0	0.0	5.8	67.7	99.1	100.0
	H_0^S	6.0	5.1	4.5	5.3	4.4	5.2	5.0
$4,000$	H_0^γ	0.0	0.0	0.0	4.0	91.8	100.0	100.0
	H_0^S	6.0	5.3	4.5	5.2	5.7	6.4	4.2
Model with $\alpha_{0+} < \alpha_{0-}$								
					α_{0-}			
n	Null	0.23	0.26	0.29	0.2971	0.3	0.32	0.34
500	H_0^γ	0.1	0.9	4.8	**7.8**	7.0	15.0	25.5
	H_0^S	5.1	7.4	13.9	14.0	16.2	18.4	21.0
$2,000$	H_0^γ	0.0	0.0	2.8	6.3	7.0	26.8	59.0
	H_0^S	7.1	13.8	29.8	39.0	39.1	52.5	66.2
$4,000$	H_0^γ	0.0	0.0	1.9	4.1	5.6	42.4	83.7
	H_0^S	6.6	26.2	55.2	59.5	65.9	82.1	91.7

TABLE 2

Test statistic T_n with the $p-$value of the nonstationarity test, and asymmetry test statistic T_n^S with its $p-$value for stock returns.

	$\hat{\alpha}_{n+}$	$\hat{\alpha}_{n-}$	$\hat{\beta}_n$	T_n	p-val	T_n^S	p-val
ICGN	0.010 (0.022)	0.397 (0.190)	0.870 (0.052)	-2.412	0.008	-2.083	0.037
MCBF	0.030 (0.031)	0.021 (0.021)	0.977 (0.011)	0.039	0.515	0.189	0.850
KVA	0.007 (0.042)	0.278 (0.133)	0.928 (0.026)	0.547	0.708	-1.946	0.052
BTC	0.188 (0.183)	0.812 (0.425)	0.771 (0.075)	-0.653	0.257	-1.392	0.164
CCME	0.492 (0.155)	0.364 (0.138)	0.744 (0.046)	0.283	0.611	0.670	0.503

the presence of a leverage effect, is often observed. Different conclusions can be obtained for individual stock returns. For comparison purposes, we took the series considered in Table VII of FZ. We estimated asymmetric GARCH(1,1) models on the daily series of Icagen (NasdaqGM: ICGN), Monarch Community Bancorp (NasdaqCM: MCBF), KV Pharmaceutical (NYSE: KV-A), Community Bankers Trust (AMEX: BTC) and China MediaExpress (NasdaqGS: CCME)[3]. The stationarity results, shown in Table 2, are in accordance with those obtained in this paper: we cannot reject explosiveness for four out of five assets. Interestingly, the symmetry assumption cannot be rejected at the 5% level for these (possibly) explosive assets. This is very different from the conclusion generally obtained for stationary series (the leverage effect). For the (probably) stationary asset, ICGN, the leverage effect is present.

[3]The data range from May 31, 2007, August 28, 2007, March 31, 2006, June 29, 2007, and March 31, 2009, respectively, to February 7, 2011.

APPENDIX C: AN EXPLICIT EXPRESSION FOR $\mathcal{I}$

To derive the explicit form of $\mathcal{I}$ in (3.4), we introduce additional notations. Let $a_{0+}(\eta_t) = \alpha_{0+}(\eta_t^+)^\delta + \beta_0$, $\quad a_{0-}(\eta_t) = \alpha_{0-}(-\eta_t^-)^\delta + \beta_0$, and for $i = 1, 2$,

$$\nu_i = E\left(\frac{\beta_0}{a_0(\eta_t)}\right)^i, \qquad \nu_{i+} = E\left(\frac{\beta_0}{a_{0+}(\eta_t)}\right)^i, \qquad \nu_{i-} = E\left(\frac{\beta_0}{a_{0-}(\eta_t)}\right)^i.$$

LEMMA C.1. *Under the assumptions of Theorem 3.1, we have $\mathcal{I} = (\mathcal{I}_{ij})$ where,*

$$\begin{aligned}
\mathcal{I}_{11} &= \frac{4}{\delta^2}\frac{(1-2\nu_{1+}+\nu_{2+})(1-\nu_1)+2(\nu_{1+}-\nu_{2+})(1-\nu_{1+})}{\alpha_{0+}^2(1-\nu_1)(1-\nu_2)},\\
\mathcal{I}_{12} &= \frac{4}{\delta^2}\frac{(\nu_{1+}-\nu_{2+})(1-\nu_{1-})+(\nu_{1-}-\nu_{2-})(1-\nu_{1+})}{\alpha_{0+}\alpha_{0-}(1-\nu_1)(1-\nu_2)} = \mathcal{I}_{21},\\
\mathcal{I}_{13} &= \frac{4}{\delta^2}\frac{\nu_2(1-\nu_{1+})+\nu_{1+}-\nu_{2+}}{\beta_0\alpha_{0+}(1-\nu_2)(1-\nu_1)} = \mathcal{I}_{31},\\
\mathcal{I}_{33} &= \frac{4}{\delta^2}\frac{\nu_2(1+\nu_1)}{\beta_0^2(1-\nu_2)(1-\nu_1)},
\end{aligned}$$

and $\mathcal{I}_{22}$ (resp. $\mathcal{I}_{23} = \mathcal{I}_{32}$) is obtained by replacing α_{0+} by α_{0-} and the ν_{i+} by ν_{i-} in $\mathcal{I}_{11}$ (resp. $\mathcal{I}_{13}$).

Proof. For ease of notation we will omit the index 0 for the true parameters and functions in this proof. We have

$$\alpha_+ d_t^{\alpha+} + \alpha_- d_t^{\alpha-} = \sum_{j=1}^{\infty}\left(\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})}\right)\left(1-\frac{\beta}{a(\eta_{t-j})}\right) = 1 \quad \text{a.s.} \tag{C.1}$$

Letting

$$\tilde\nu_{i+} = E\left(\frac{(\eta_t^+)^\delta}{a(\eta_t)}\right)^i = E\left(\frac{(\eta_t^+)^\delta}{a_+(\eta_t)}\right)^i, \quad \tilde\nu_{i-} = E\left(\frac{(-\eta_t^-)^\delta}{a(\eta_t)}\right)^i = E\left(\frac{(-\eta_t^-)^\delta}{a_-(\eta_t)}\right)^i$$

we obtain

$$E\left(d_t^{\alpha+}\right) = \frac{\tilde\nu_{1+}}{1-\nu_1}, \qquad E\left(d_t^{\alpha-}\right) = \frac{\tilde\nu_{1-}}{1-\nu_1}. \tag{C.2}$$

Noting that

$$\alpha_+ E\frac{(\eta_t^+)^\delta}{a^2(\eta_t)} + \frac{\nu_2^+}{\beta} = \frac{\nu_1^+}{\beta},$$

we have

$$E\frac{(\eta_t^+)^\delta}{a^2(\eta_t)} = \frac{\nu_{1+}-\nu_{2+}}{\beta\alpha_+}.$$

It follows that

$$\begin{aligned}
E\left(d_t^{\alpha+}\right)^2 &= \tilde\nu_{2+}\sum_{j=1}^{\infty}\nu_2^{j-1} + 2\sum_{j=1}^{\infty}\sum_{h=1}^{\infty}\nu_2^{j-1}\beta E\frac{(\eta_1^+)^\delta}{a^2(\eta_1)}\nu_1^{h-1}\tilde\nu_{1+}\\
&= \frac{\tilde\nu_{2+}}{1-\nu_2} + 2\frac{1}{1-\nu_2}\frac{1}{1-\nu_1}\frac{\nu_{1+}-\nu_{2+}}{\alpha_+}\tilde\nu_{1+}\\
&= \frac{\tilde\nu_{2+}\alpha_+(1-\nu_1)+2(\nu_{1+}-\nu_{2+})\tilde\nu_{1+}}{\alpha_+(1-\nu_1)(1-\nu_2)}.
\end{aligned}$$

Symmetrically

$$E\left(d_t^{\alpha_-}\right)^2 = \frac{\tilde{\nu}_{2-}\alpha_-(1-\nu_1)+2\left(\nu_{1-}-\nu_{2-}\right)\tilde{\nu}_{1-}}{\alpha_-(1-\nu_1)(1-\nu_2)}.$$

Moreover,

$$\begin{aligned} Ed_t^{\alpha_+}d_t^{\alpha_-} &= \sum_{j=1}^{\infty}\nu_2^{j-1}\beta E\frac{\left(\eta_1^+\right)^\delta}{a^2(\eta_1)}\sum_{h=1}^{\infty}\nu_1^{h-1}\tilde{\nu}_{1-}+\sum_{j=1}^{\infty}\nu_2^{j-1}\beta E\frac{\left(-\eta_1^-\right)^\delta}{a^2(\eta_1)}\sum_{h=1}^{\infty}\nu_1^{h-1}\tilde{\nu}_{1+} \\ &= \frac{\alpha_-(\nu_{1+}-\nu_{2+})\tilde{\nu}_{1-}+\alpha_+(\nu_{1-}-\nu_{2-})\tilde{\nu}_{1+}}{\alpha_+\alpha_-(1-\nu_1)(1-\nu_2)}. \end{aligned}$$

Noting that

$$\begin{aligned} \alpha_+\tilde{\nu}_{1+}+\nu_1^+ &= 1, \\ \alpha_+\tilde{\nu}_{1+}+\alpha_-\tilde{\nu}_{1-}+\nu_1 &= 1 \\ 2-\nu_{1+}-\nu_{1-} &= 1-\nu_1, \\ \alpha_+^2\tilde{\nu}_{2+}+\nu_{2+}+2(\nu_{1+}-\nu_{2+}) &= 1 \\ (\nu_{1+}-\nu_{2+})+(\nu_{1-}-\nu_{2-}) &= \nu_1-\nu_2. \end{aligned}$$

we obtain the announced formulas for $\mathcal{I}_{11}, \mathcal{I}_{12}$ and $\mathcal{I}_{22}$.

Now

$$d_t^\beta = \sum_{j=2}^{\infty}(j-1)\left(\prod_{k=1}^{j-1}\frac{\beta}{a(\eta_{t-k})}\right)\frac{\alpha_+(\eta_{t-j}^+)^\delta+\alpha_-(-\eta_{t-j}^-)^\delta}{\beta a(\eta_{t-j})}.$$

Noting that

$$E\left(\frac{\alpha_+(\eta_1^+)^\delta+\alpha_-(-\eta_1^-)^\delta}{a(\eta_1)}\right) = E\left(1-\frac{\beta}{a(\eta_1)}\right) = 1-\nu_1,$$

$$E\left(\frac{\alpha_+(\eta_1^+)^\delta+\alpha_-(-\eta_1^-)^\delta}{a(\eta_1)}\right)^2 = E\left(1-\frac{\beta}{a(\eta_1)}\right)^2 = 1+\nu_2-2\nu_1,$$

and

$$E\frac{\alpha_+(\eta_1^+)^\delta+\alpha_-(-\eta_1^-)^\delta}{a^2(\eta_1)} = \frac{\nu_1-\nu_2}{\beta},$$

we have

$$\begin{aligned} E\left(d_t^\beta\right) &= \frac{\nu_1}{(1-\nu_1)\beta}, \qquad \text{(C.3)} \\ E\left(d_t^\beta\right)^2 &= \sum_{j=2}^{\infty}(j-1)^2\nu_2^{j-1}\frac{1+\nu_2-2\nu_1}{\beta^2} \\ &\quad +2\sum_{j=2}^{\infty}\sum_{h=1}^{\infty}(j-1)(j+h-1)\nu_2^{j-1}\frac{\nu_1-\nu_2}{\beta}\nu_1^{h-1}\frac{1-\nu_1}{\beta} \\ &= \frac{(1-2\nu_1+\nu_2)\nu_2(\nu_2+1)}{\beta^2(1-\nu_2)^3} \\ &\quad +2\frac{(1-\nu_1)(\nu_1-\nu_2)}{\beta^2}\sum_{h=1}^{\infty}\frac{\nu_2(\nu_2+1+h-\nu_2 h)}{(1-\nu_2)^3}\nu_1^{h-1}, \end{aligned}$$

which, in view of $\mathcal{I} = \frac{4}{\delta^2} E d_1 d_1'$, gives the formula for $\mathcal{I}_{33}$. Noting that

$$E\frac{(\eta_1^+)^\delta\left\{\alpha_+(\eta_1^+)^\delta + \alpha_-(-\eta_1^-)^\delta\right\}}{a^2(\eta_1)} = \tilde{\nu}_{1+} - \frac{\nu_{1+}-\nu_{2+}}{\alpha_+} = \frac{1-2\nu_{1+}+\nu_{2+}}{\alpha_+},$$

we also have

$$\begin{aligned}
E\left(d_t^{\alpha_+}d_t^{\beta}\right) &= \sum_{j=2}^{\infty}(j-1)\nu_2^{j-1}\frac{1-2\nu_{1+}+\nu_{2+}}{\beta\alpha_+} \\
&\quad + \sum_{j=1}^{\infty}\sum_{h=1}^{\infty}(j+h-1)\nu_2^{j-1}\frac{\nu_{1+}-\nu_{2+}}{\alpha_+}\nu_1^{h-1}\frac{1-\nu_1}{\beta} \\
&\quad + \sum_{j=2}^{\infty}\sum_{h=1}^{\infty}(j-1)\nu_2^{j-1}\frac{\nu_1-\nu_2}{\beta}\nu_1^{h-1}\tilde{\nu}_{1+} \\
&= \frac{\nu_2(1-2\nu_{1+}+\nu_{2+})}{\beta\alpha_+(1-\nu_2)^2} \\
&\quad + \frac{(1-\nu_1)(\nu_{1+}-\nu_{2+})}{\beta\alpha_+}\left\{\sum_{h=1}^{\infty}\left(\frac{\nu_2}{(1-\nu_2)^2} + h\frac{1}{1-\nu_2}\right)\nu_1^{h-1}\right\} \\
&\quad + \frac{(\nu_1-\nu_2)\tilde{\nu}_{1+}}{\beta}\left\{\sum_{h=1}^{\infty}\frac{\nu_2}{(1-\nu_2)^2}\nu_1^{h-1}\right\}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\nu_2(1-2\nu_{1+}+\nu_{2+})}{\beta\alpha_+(1-\nu_2)^2} \\
&\quad + \frac{(1-\nu_1)(\nu_{1+}-\nu_{2+})}{\beta\alpha_+}\left\{\frac{\nu_2}{(1-\nu_2)^2}\frac{1}{(1-\nu_1)} + \frac{1}{1-\nu_2}\frac{1}{(1-\nu_1)^2}\right\} \\
&\quad + \frac{(\nu_1-\nu_2)(1-\nu_{1+})}{\beta\alpha_+}\left\{\frac{\nu_2}{(1-\nu_2)^2}\frac{1}{(1-\nu_1)}\right\},
\end{aligned}$$

which gives the formula for $\mathcal{I}_{13}$. □

APPENDIX D: PROOFS

D.1 Proof of Theorem 3.2

The convergence results in *i)* follow from Taylor expansions of the functions $\hat{\kappa}_\eta = \kappa_\eta(\hat{\theta}_n)$ and $\frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma_t^2(\hat{\theta}_n)}\frac{\partial\sigma_t^2}{\partial\theta_i}\frac{\partial\sigma_t^2}{\partial\theta_j}(\hat{\theta}_n)$ around θ_0, and the ergodic theorem together with the consistency of $\hat{\theta}_n$. š

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

$$\text{(D.1)}\quad \hat{\kappa}_\eta = \frac{1}{n}\sum_{t=1}^n \eta_t^4 - \frac{4}{\delta n}\sum_{t=1}^n \frac{\epsilon_t^4}{\sigma_t^4(\theta^*)}\frac{1}{\sigma_t^\delta(\theta^*)}\frac{\partial\sigma_t^\delta(\theta^*)}{\partial\theta'}(\hat{\theta}_n-\theta_0) := \frac{1}{n}\sum_{t=1}^n \eta_t^4 + R_n.$$

Write $d_t(\vartheta) = (d_t^{\alpha+}(\vartheta), d_t^{\alpha-}(\vartheta), d_t^{\beta}(\vartheta))'$. Let the matrix norm defined by $\|A\| = \sum|a_{ij}|$ with standard notations. By Proposition 2.1 and already given arguments, for some $\rho\in(0,1)$,

$$|R_n| \le \frac{K}{n}\sum_{t=1}^n \eta_t^4\frac{h_t^{4/\delta}}{\sigma_t^4(\theta^*)}\left(\rho^t|\hat{\omega}_n-\omega_0| + \|d_t(\vartheta^*)\|\|\hat{\vartheta}_n-\vartheta_0\|\right) = o(1), \quad \text{a.s.}$$

Hence the first part of *ii)* is proven. Now we have, using (7.13),

$$n\widehat{\mathcal{J}}_{\omega,\alpha_+} \leq \frac{4}{\delta^2}\sum_{t=1}^{n}\frac{1}{\hat{\alpha}_{n+}}\frac{h_t}{\sigma_t^\delta(\hat{\theta}_n)}\frac{\sum_{j=1}^{t}\hat{\beta}_n^{j-1}}{h_t} \leq K\sum_{t=1}^{n}\frac{h_t}{\sigma_t^\delta(\hat{\theta}_n)}\rho^t, \tag{D.2}$$

for $\rho \in (0,1)$ when n is large enough, by Proposition 2.1 *i)*. It follows that $n\widehat{\mathcal{J}}_{\omega,\alpha_+} = O(1)$ a.s. More generally $n\widehat{\mathcal{J}}_{\omega,\vartheta} = O(1)$ a.s. Moreover, we have $n\widehat{\mathcal{J}}_{\omega,\omega} \geq 1/\sigma_1^4(\hat{\theta}_n) > 0$. Thus we have shown that

$$\widehat{\mathcal{J}}_{\vartheta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\vartheta} = o(1), \quad \text{a.s.}$$

Now we turn to $\widehat{\mathcal{J}}_{\vartheta,\vartheta}$. Considering the top-left term, a Taylor expansion around θ_0 gives

$$\begin{aligned}
\widehat{\mathcal{J}}_{\alpha_+,\alpha_+} &= \frac{4}{\delta^2}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\sigma_t^\delta(\hat{\theta}_n)}\sum_{j=1}^{t}\hat{\beta}_n^{j-1}(\epsilon_{t-j}^+)^\delta\right)^2 \\
&= \frac{4}{\delta^2}\frac{1}{n}\sum_{t=1}^{n}\{d_t^{\alpha_+}(\vartheta_0)\}^2 \\
&\quad +\frac{4}{\delta^2}\frac{1}{n}\sum_{t=1}^{n}\left[\left\{\frac{1}{\sigma_t^\delta(\theta)}\frac{\partial\sigma_t^\delta}{\partial\alpha_+}(\theta_0)\right\}^2 - \{d_t^{\alpha_+}(\vartheta_0)\}^2\right] + S_n,
\end{aligned} \tag{D.3}$$

where, for θ^* such that $\|\theta_0-\theta^*\| \leq \|\theta_0-\hat{\theta}_n)\|$,

$$\begin{aligned}
&|S_n| \\
\leq\ & \frac{K}{n}\sum_{t=1}^{n}\left(\frac{\sum_{j=1}^{t}(\beta^*)^{j-1}(\epsilon_{t-j}^+)^\delta}{\sigma_t^\delta(\theta^*)}\right)^2\left\{\rho^t|\hat{\omega}_n-\omega_0| + \|d_t(\vartheta^*)\|\|\hat{\vartheta}_n-\vartheta_0\|\right\} \\
&+\frac{K}{n}\sum_{t=1}^{n}\left(\frac{\sum_{j=1}^{t}(\beta^*)^{j-1}(\epsilon_{t-j}^+)^\delta}{\sigma_t^\delta(\theta^*)}\right)\left(\frac{\sum_{j=1}^{t}(j-1)(\beta^*)^{j-2}(\epsilon_{t-j}^+)^\delta}{\sigma_t^\delta(\theta^*)}\right)|\hat{\beta}_n-\beta_0| \\
=\ & o(1), \quad a.s.
\end{aligned}$$

by already used arguments. Moreover, the second term in the right-hand side of (D.4) converges to 0 a.s. by (7.7), while the first term converges to $\mathcal{I}_{11}$. We thus have shown that $\widehat{\mathcal{J}}_{\alpha_+,\alpha_+}$ a.s. converges to $\mathcal{I}_{11}$. The other two terms in $\widehat{\mathcal{J}}_{\vartheta,\vartheta}$ can be handled similarly, which completes the proof of *ii)*.

Turning to *iii)*, we note that $\partial\sigma_t^\delta(\theta^*)/\partial\omega \leq K$ for n large enough, since $\beta_0 < 1$. Moreover, $\sigma_t^\delta(\theta^*) \geq \omega^* + \underline{\alpha}|\epsilon_{t-1}|^\delta$. Therefore (D.1) continues to hold with $|R_n|$ bounded by

$$\frac{K}{n}\sum_{t=1}^{n}\eta_t^4\left(\frac{h_t}{\sigma_t^2(\theta^*)}\right)^2\frac{1}{\omega^*+\underline{\alpha}|\epsilon_{t-1}|^\delta} + \frac{K}{n}\sum_{t=1}^{n}\eta_t^4\left(\frac{h_t}{\sigma_t^2(\theta^*)}\right)^2\|d_t(\vartheta^*)\| \times\|\hat{\vartheta}_n-\vartheta_0\|.$$

Therefore $|R_n| = o_P(1)$ by Proposition 2.1 *iii)*, the weak consistency of $\hat{\vartheta}_n$ and the existence of moments for $d_t(\vartheta^*)$ and $h_t/\sigma_t^\delta(\theta^*)$. Hence $\hat{\kappa}_\eta \to \kappa$ in probability. By arguments already used we have $n\widehat{\mathcal{J}}_{\omega,\vartheta} = O_P(1)$, $\widehat{\mathcal{J}}_{\vartheta,\omega}\widehat{\mathcal{J}}_{\omega,\omega}^{-1}\widehat{\mathcal{J}}_{\omega,\vartheta\beta} = o_P(1)$, and the right-hand side of (D.4) converges to $\mathcal{I}(1,1)$ in probability. □

D.2 Proof of Theorem 6.1

Note that $\hat{\zeta}_n = \arg\min_{\zeta\in\Upsilon} Q_n(\zeta)$, where

$$\begin{aligned} Q_n(\zeta) &= \frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\left(\frac{h_t^{\delta/\delta_0}}{\sigma_t^\delta(\zeta)}\right)^{2/\delta}-1\right\}+\log\left(\frac{\sigma_t^\delta(\zeta)}{h_t^{\delta/\delta_0}}\right)^{2/\delta} \\ &= O_n(\delta,\vartheta)+R_n(\zeta) \end{aligned}$$

where

$$O_n(\delta,\vartheta)=\frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\frac{1}{v_t^{2/\delta}(\delta,\vartheta)}-1\right\}+\log v_t^{2/\delta}(\delta,\vartheta),$$

$$v_t(\delta,\vartheta)=\sum_{j=1}^\infty \frac{\alpha_+(\eta_{t-j}^+)^\delta+\alpha_-(-\eta_{t-j}^-)^\delta}{\{a_0(\eta_{t-j})\}^{\delta/\delta_0}}\prod_{k=1}^{j-1}\frac{\beta}{\{a_0(\eta_{t-k})\}^{\delta/\delta_0}},$$

and

$$R_n(\zeta)=\frac{1}{n}\sum_{t=1}^n \eta_t^2\left\{\left(\frac{h_t^{\delta/\delta_0}}{\sigma_t^\delta(\zeta)}\right)^{2/\delta}-\frac{1}{v_t^{2/\delta}(\delta,\vartheta)}\right\}+\log\left(\frac{\sigma_t^\delta(\zeta)}{h_t^{\delta/\delta_0}v_t(\delta,\vartheta)}\right)^{2/\delta}.$$

To prove the consistency, in the cases $\gamma_0>0$ and $\gamma_0=0$, it will be sufficient to establish Lemmas D.1, D.2 and D.3 below.

Let $\Upsilon_0=\{\zeta\in\Upsilon:\beta<e^{\frac{\delta}{\delta_0}\gamma_0}\}$ and $\Upsilon_p=\{\zeta\in[0,\infty)^5:\beta<\|1/a_0^{\delta/\delta_0}(\eta_1)\|_p^{-1}\}$.

Lemma D.1. *i) When $\gamma_0>0$, for any $\zeta\in\Upsilon_0$ the process $v_t(\delta,\vartheta)$ is stationary and ergodic. Moreover, for any compact $\Upsilon_0^*\subset\Upsilon_0$,*

$$\sup_{\zeta\in\Upsilon_0^*}\left|\frac{\sigma_t^\delta(\zeta)}{h_t^{\delta/\delta_0}}-v_t(\delta,\vartheta)\right|\to 0 \text{ a.s. as } t\to\infty.$$

Finally, for any $\zeta\notin\Upsilon_0$ it holds that $\sigma_t^\delta(\zeta)/h_t^{\delta/\delta_0}\to\infty$ a.s.

ii) When $\gamma_0=0$, for any $\zeta\in\Upsilon_p$ with $p\geq 1$, the process $v_t(\delta,\vartheta)$ is stationary and ergodic. Moreover, for any compact $\Upsilon_p^*\subset\Upsilon_p$,

$$\sup_{\zeta\in\Upsilon_p^*}\left|\frac{\sigma_t^\delta(\zeta)}{h_t^{\delta/\delta_0}}-v_t(\delta,\vartheta)\right|\to 0 \text{ in } L^p.$$

Proof. When $\gamma_0>0$, for $\zeta\in\Upsilon_0$, by the Cauchy root test, the series $v_t(\delta,\vartheta)$ in a.s. finite. As a measurable function of $\{\eta_u, u<t\}$, the process $v_t(\delta,\vartheta)$ is thus stationary and ergodic. When $\gamma_0=0$, since $\|v_t(\delta,\vartheta)\|_p<\infty$ for $\zeta\in\Upsilon_p$, $v_t(\delta,\vartheta)$ is a.s. finite and the stationarity and ergodicity follow.

We have, keeping the notation of the proof of Lemma 7.1,

$$\frac{\sigma_t^\delta(\zeta)}{h_t^{\delta/\delta_0}} = \sum_{j=1}^t \beta^{j-1}\left\{\prod_{k=1}^j\frac{h_{t-k}}{h_{t-k+1}}\right\}^{\delta/\delta_0}\frac{z_{t-j}}{h_{t-j}^{\delta/\delta_0}}.$$

In view of (7.4) we have

$$\left\{\frac{h_{t-k}}{h_{t-k+1}}\right\}^{\delta/\delta_0}\leq\frac{1}{a_0(\eta_{t-k})^{\delta/\delta_0}}, \tag{D.4}$$

and the conclusion follows from arguments already used. □

Lemma D.2. *If $\zeta \in \Upsilon_0$, we have*

$$v_t(\delta,\vartheta) = 1, \quad a.s. \quad iff \quad (\delta,\vartheta) = (\delta_0,\vartheta_0).$$

Proof. We have

$$v_t(\delta,\vartheta)a_0(\eta_{t-1})^{\delta/\delta_0} = \beta v_{t-1}(\delta,\vartheta) + \alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta. \tag{D.5}$$

Thus, $v_t(\delta,\vartheta) = 1$ a.s. iff

$$\beta + \alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta - \{\beta_0 + \alpha_{0+}(\eta_{t-1}^+)^{\delta_0} + \alpha_{0-}(-\eta_{t-1}^-)^{\delta_0}\}^{\delta/\delta_0} = 0.$$

Straightforward algebra shows that the function $x \mapsto \beta + \alpha_+ x - \{\beta_0 + \alpha_{0+}x^{\delta_0/\delta}\}^{\delta/\delta_0}$, has at most two zeroes on $(0,\infty)$, except when $\delta = \delta_0, \beta = \beta_0$ and $\alpha_+ = \alpha_{0+}$. Similarly, the function $x \mapsto \beta + \alpha_- x - \{\beta_0 + \alpha_{0-}x^{\delta_0/\delta}\}^{\delta/\delta_0}$, has at most two zeroes on $(0,\infty)$, except when $\delta = \delta_0, \beta = \beta_0$ and $\alpha_- = \alpha_{0-}$. By Assumption **A3** we can conclude that $(\delta,\vartheta) = (\delta_0,\vartheta_0)$. □

To handle $R_n(\zeta)$ we prove the following lemma. Let $\check{\Upsilon}$ be the compact set of the (δ,ϑ)'s such that $\zeta \in \Upsilon$.

Lemma D.3. *Suppose that $P(\eta_t = 0) = 0$. Then, for any $k > 0$*

$$E \sup_{(\delta,\vartheta)\in\check{\Upsilon}} \left(\frac{1}{v_t(\delta,\vartheta)}\right)^k < \infty \quad and \quad E \sup_{\zeta\in\Upsilon} \left(\frac{h_t^{\delta/\delta_0}}{\sigma_t^\delta(\zeta)}\right)^k < \infty.$$

Proof. Let $\varepsilon > 0$ such that $p(\varepsilon) := P(|\eta_t| \le \varepsilon) \in [0,1)$. If $|\eta_{t-1}| > \varepsilon$, since the sum $v_t(\delta,\vartheta)$ is greater than its first term, we have,

$$\begin{aligned} \frac{1}{v_t(\delta,\vartheta)} &\le \frac{a_0^{\delta/\delta_0}(\eta_{t-1})}{\alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta} \\ &\le \left(\frac{\max(\alpha_{0+},\alpha_{0-})}{\underline{\alpha}^{\delta_0/\delta}} + \frac{\beta_0}{\underline{\alpha}^{\delta_0/\delta}\varepsilon^{\delta_0}}\right)^{\delta/\delta_0} \\ &:= K(\varepsilon). \end{aligned}$$

Iterating this method, we can write

$$\sup_{(\delta,\vartheta)\in\check{\Upsilon}} \frac{1}{v_t(\delta,\vartheta)} \le K(\varepsilon)\sum_{i=1}^\infty \mathbb{1}_{|\eta_{t-1}|\le\varepsilon}\cdots\mathbb{1}_{|\eta_{t-i+1}|\le\varepsilon}\mathbb{1}_{|\eta_{t-i}|>\varepsilon}\left(\frac{\overline{a}_0(\varepsilon)}{\underline{\beta}}\right)^{i-1}.$$

where $\overline{a}_0(\varepsilon) = \max(\alpha_{0+},\alpha_{0-})\epsilon^{\delta_0} + \beta_0$. The first result of the lemma follows by the arguments given in the proof of Lemma 7.3.

Similarly, we have for $|\eta_{t-1}| > \varepsilon$,

$$\begin{aligned} \frac{h_t^{\delta/\delta_0}}{\sigma_t^\delta(\zeta)} &\le \frac{\{\omega_0 + a_0(\eta_{t-1})h_{t-1}\}^{\delta/\delta_0}}{\omega + h_{t-1}^{\delta/\delta_0}\{\alpha_+(\eta_{t-1}^+)^\delta + \alpha_-(-\eta_{t-1}^-)^\delta\} + \beta\sigma_{t-1}^\delta(\zeta)} \\ &\le \left(\frac{\omega_0}{\underline{\omega}^{\delta_0/\delta}} + \frac{\overline{\alpha}}{\underline{\alpha}^{\delta_0/\delta}} + \frac{\beta_0}{\underline{\alpha}^{\delta_0/\delta}\varepsilon^{\delta_0}}\right)^{\delta/\delta_0} := H(\varepsilon), \end{aligned}$$

and for $|\eta_{t-1}| \leq \varepsilon$ and $|\eta_{t-2}| > \varepsilon$,

$$\frac{h_t^{\delta/\delta_0}}{\sigma_t^\delta(\zeta)} \leq \left(\frac{\omega_0}{\underline{\omega}^{\delta_0/\delta}} + \frac{\overline{a}_0(\varepsilon)}{\underline{\beta}^{\delta_0/\delta}} H(\varepsilon) \right)^{\delta/\delta_0}.$$

The conclusion follows by the arguments used for Lemma 7.3. □

Now we turn to the asymptotic normality. Let, for $\eta \in \mathbb{R}$,

$$G(\eta) = \log|\eta| \left\{ 1 - \frac{\beta_0}{a_0(\eta)} \right\} - \frac{1}{\delta} \log\{a_0(\eta)\},$$

with by convention $G(0) = -\log(\beta_0)/\delta$, and let

$$d_t^\delta = \sum_{j=1}^\infty G(\eta_{t-j}) \prod_{k=1}^{j-1} \frac{\beta_0}{a_0(\eta_{t-k})}.$$

Since $E\log\{\beta_0/a_0(\eta_1)\} < 0$, by the Cauchy root test, the process d_t^δ is stationary and ergodic.

Lemma D.4. *Assume $\gamma_0 \geq 0$ and $E\eta_t^4 < \infty$. We have*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t}{\partial \zeta'}(\zeta_0) \quad \xrightarrow{d} \quad \mathcal{N}\{0, (\kappa_\eta - 1)\mathcal{I}_\delta\} \qquad \text{as } n \to \infty,$$

where, letting $D_t' = \left(d_t^\delta, d_t^{\alpha+}, d_t^{\alpha-}, d_t^\beta\right)$, $\mathcal{I}_\delta = \frac{4}{\delta^2} E D_1 D_1'$ is nonsingular.

Proof. We have

$$\frac{\partial \ell_t(\zeta)}{\partial \delta} = \frac{2}{\delta} \left\{ 1 - \frac{\epsilon_t^2}{\sigma_t^2} \right\} \left\{ \frac{1}{\sigma_t^\delta} \frac{\partial \sigma_t^\delta}{\partial \delta} - \log \sigma_t \right\}.$$

Moreover,

$$\frac{\partial \sigma_t^\delta}{\partial \delta} = \alpha_+ (\epsilon_{t-1}^+)^\delta \log(\epsilon_{t-1}^+) + \alpha_- (-\epsilon_{t-1}^-)^\delta \log(-\epsilon_{t-1}^-) + \beta \frac{\partial \sigma_{t-1}^\delta}{\partial \delta}, \tag{D.6}$$

with by convention $\log(0) \times 0 = 0$. Thus, computation shows that, at ζ_0,

$$Z_t := \frac{1}{\sigma_t^{\delta_0}} \frac{\partial \sigma_t^\delta}{\partial \delta} - \log \sigma_t = G(\eta_{t-1}) + \frac{\beta_0}{a_0(\eta_{t-1})} Z_{t-1} + W_t,$$

with

$$\begin{aligned} W_t &= \left(1 - \frac{\sigma_t^{\delta_0}}{a_0(\eta_{t-1})\sigma_{t-1}^{\delta_0}} \right) \frac{1}{\sigma_t^{\delta_0}} \frac{\partial \sigma_t^\delta}{\partial \delta} - \frac{1}{\delta_0} \log\left(\frac{\sigma_t^{\delta_0}}{a_0(\eta_{t-1})\sigma_{t-1}^{\delta_0}} \right) \\ &= -\frac{\omega_0}{a_0(\eta_{t-1})\sigma_{t-1}^{\delta_0}} \frac{1}{\sigma_t^{\delta_0}} \frac{\partial \sigma_t^\delta}{\partial \delta} - \frac{1}{\delta_0} \log\left(1 + \frac{\omega_0}{a_0(\eta_{t-1})\sigma_{t-1}^{\delta_0}} \right). \end{aligned} \tag{D.7}$$

We will show that

$$W_t \to 0 \text{ in } L^p \text{ as } t \to \infty. \tag{D.8}$$

First note that the second term in the right-hand side of (D.7) converges to 0 in L^p, by Proposition 2.1. Now in view of (D.6) we have

$$\frac{1}{\sigma_t^\delta}\frac{1}{\sigma_{t-1}^\delta}\frac{\partial \sigma_t^\delta}{\partial \delta}(\zeta_0) = \sum_{j=1}^t \beta_0^{j-1}\frac{h_{t-1}}{h_t}\left\{\prod_{k=2}^j \frac{h_{t-k}}{h_{t-k+1}}\right\}^2 \frac{u(\eta_{t-j})\log(|\epsilon_{t-j}|)}{h_{t-j}},$$

where $u(x) = \alpha_{0+}(x^+)^{\delta_0} + \alpha_{0-}(-x^-)^{\delta_0}$. The first term in the right-hand side of (D.7) is thus bounded, in absolute values, by

$$M_t = \frac{\omega_0}{a_0^2(\eta_{t-1})}\sum_{j=1}^\infty \left\{\prod_{k=2}^j \frac{\beta_0}{a_0^2(\eta_{t-k})}\right\}\frac{u(\eta_{t-j})\{|\log(h_{t-j})| + |\log|\eta_{t-j}||\}}{h_{t-j}}.$$

Write $M_t = \sum_{j=1}^{t_0} M_{jt} + \sum_{j=t_0+1}^\infty M_{jt}$. Note that $M_{jt} \to 0$ because $\{|\log(h_{t-j})| + |\log|\eta_{t-j}||\}/h_{t-j} \to 0$ in L^p as $t \to \infty$, by Proposition 2.1. Hence $\sum_{j=1}^{t_0} M_{jt} \to 0$ in L^p as $t \to \infty$. Moreover, it can be noted that $\{|\log(h_{t-j})| + |\log|\eta_{t-j}||\}/h_{t-j} \le K(1 + |\log|\eta_{t-j}||)$. Thus

$$\sum_{j=1}^\infty M_{jt} \le K\sum_{j=1}^\infty \left\{\prod_{k=2}^j \frac{\beta_0}{a_0^2(\eta_{t-k})}\right\} u(\eta_{t-j})(1 + |\log|\eta_{t-j}||).$$

Moreover, under **A4**,

$$\left\|\sum_{j=1}^\infty M_{jt}\right\|_p \le K\sum_{j=1}^\infty \left\|\frac{\beta_0}{a_0^2(\eta_1)}\right\|_p^{j-1} \|u(\eta_1)(1 + |\log|\eta_1||)\|_p < \infty.$$

It follows that $\sum_{j=t_0+1}^\infty M_{jt} \to 0$ in L^p as $t_0 \to \infty$. The convergence in (D.8) is thus established.

On the other hand, (d_t^δ) is the strictly stationary and non anticipative solution of the stochastic recurrence equation

$$d_t^\delta = G(\eta_{t-1}) + \frac{\beta_0}{a_0(\eta_{t-1})} d_{t-1}^\delta.$$

It follows that

$$Z_t - d_t^\delta = \frac{\beta_0}{a_0(\eta_{t-1})}(Z_{t-1} - d_{t-1}^\delta) + W_t.$$

and thus

$$Z_t - d_t^\delta = \sum_{j=1}^t \left\{\prod_{k=2}^j \frac{\beta_0}{a_0(\eta_{t-k})}\right\} W_{t-j}.$$

In view of (D.8) and arguments already used it follows that

$$Z_t - d_t^\delta \to 0 \text{ in } L^p \text{ as } t \to \infty. \tag{D.9}$$

The conclusion straightforwardly follows.

Finally, we show that $\mathcal{I}_\delta$ is nonsingular. Suppose there exists $x = (x_1, x_2, x_3, x_4)' \in \mathbb{R}^4$ such that $x'\mathcal{I}_\delta x = 0$. By arguments given in the proof of Lemma 7.4, it follows that

$$x_1 G(\eta_{t-1})a_0(\eta_{t-1}) + x_2(\eta_{t-1}^+)^\delta + x_3(-\eta_{t-1}^-)^\delta = K, \quad a.s.$$

where K is a constant. Letting $L(\eta) = x_1 G(\eta) a_0(\eta) + x_2(\eta^+)^\delta + x_3(-\eta^-)^\delta - K$, we find that for $\eta > 0$, the derivative

$$L'(\eta) = (\eta)^{\delta-1}\left\{x_2\delta + x_1\alpha_+ \log\left(\frac{(\eta)^\delta}{\beta_0 + \alpha_{0+}(\eta)^\delta}\right)\right\}$$

cancels at most once on $(0,\infty)$, except if $x_1 = x_2 = 0$. It follows that the equation $L(\eta) = 0$ has at most two solutions on $(0,\infty)$ if $(x_1, x_2) \neq (0,0)$. The same arguments apply on $(-\infty, 0)$. In view of Assumption **A3**, we conclude that $\mathcal{I}_\delta$ is nonsingular. $\square$

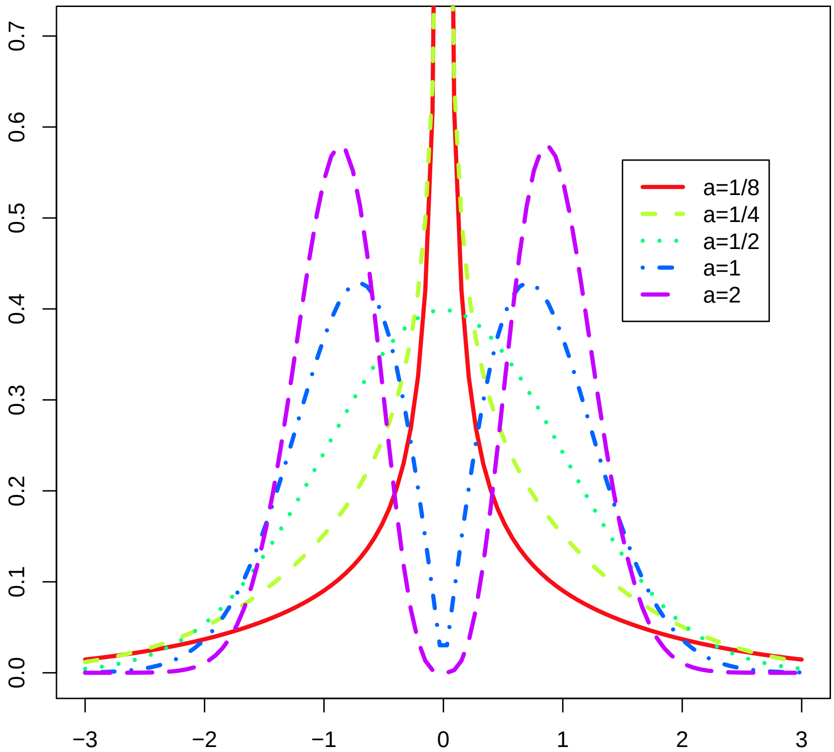

FIGURE 2. *Densities (5.10) of* η_t *for which the asymmetry test (4.7) is asymptotically optimal.*