STATIONARITY AND MEMORY OF *ARCH* MODELS (∞)

by

Paolo Zaffaroni*
Banca d'Italia, Roma

Contents:

Discussion Paper
No. EM/00/383
March 2000

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7405 7686

---

* I thank Peter M Robinson for useful comments and suggestions.

Abstract

Sufficient conditions for strict stationarity of ARCH(∞) are established, without imposing covariance stationarity and for any specification of the conditional second moment coefficients. GARCH(*p,q*) as well as the case of hyperbolically decaying coefficients are included, such as the autoregressive coefficients of ARFIMA(*p,d,q*), once the non-negativity constraints are imposed. Second, we show the necessary and sufficient conditions for covariance stationarity of ARCH(∞), both for the levels and the squares. These prove to be much stronger than the strict stationarity conditions. The covariance stationarity condition for the levels rules out long memory in the squares.

Keywords: ARCH(∞); GARCH(*p,q*); nonlinear moving average representation; strict and weak stationarity; memory.

JEL No.: C22

1 Introduction

Introduced by the seminal work of Engle (1982), ARCH represent with no doubt the most popular class of nonlinear time series models, in particular thanks to the GARCH(p, q) development (Bollerslev 1986) defined by

$$\epsilon_t = z_t \sigma_t, \tag{1}$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + .. + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + .. + \beta_p \sigma_{t-p}^2, \;\; a.s., \tag{2}$$

where $\omega > 0$, $\beta_i \geq 0, \alpha_j > 0 \, (i = 1, .., p, \; j = 1, .., q)$ for integers $p \geq 0, \; q > 0$ and a.s. means almost surely. The minimum conditions, which we assume hereafter, imposed on the rescaled innovation z_t are i.i.d.-ness, that $|z_t| < \infty$ a.s. and that the z_t^2 are not degenerate. When $p = 0$ one gets the ARCH(q).

However, for long time, the probabilistic properties of ARCH(p) and GARCH(p, q) and of the related statistical inference methods remained unknown. The only exception is Weiss (1986) who established the asymptotic properties of the Gaussian pseudo maximum likelihood estimator (PMLE) and of the least squares estimator for ARCH(p). However severe restrictions were imposed, in particular a bounded fourth moment of the ϵ_t was required. This is both theoretically unduly restrictive (e.g. for ARCH(1) it implies that α_1 be smaller than $1/\sqrt{3}$) as well as in contrast with empirical findings suggesting unbounded kurtosis of asset returns' distribution (see e.g. Loretan and Phillips (1994)).

The breakthrough development was made by Nelson (1990) who established necessary and sufficient conditions on the GARCH coefficients and on the rescaled innovation for strict stationarity and ergodicity of GARCH$(1, 1)$ allowing for unbounded second moment of the ϵ_t, including the IGARCH$(1, 1)$ case. In particular when

$$E[ln(\beta_1 + \alpha_1 z_t^2)] < 0, \tag{3}$$

then σ_t^2 is strictly stationary and ergodic. The strength of (3) is that nowhere covariance stationarity of the ϵ_t is implied and, indeed, it allows for mildly explosive behaviour, e.g. $\alpha_1 + \beta_1 > 1$ when $E(z_t^2) = 1$.

The appeal of Nelson (1990)'s strict stationarity condition (3) consists in its simplicity, in particular as a primitive function of the GARCH coefficients and of the rescaled innovation distribution.

Indeed Lumsdaine (1996) and Lee and Hansen (1994) crucially rely on Nelson (1990) result in order to establish the asymptotic properties of the Gaussian PMLE for GARCH/IGARCH$(1,1)$ without imposing a bounded second, and thus fourth, moment condition.

Sufficient conditions for strict stationarity of more general GARCH(p,q) were obtained by Bougerol and Picard (1992). They rely on the representation of GARCH as solutions of multivariate stochastic recurrence relations and give the essential condition in terms of the top Lyapunov exponent of certain random matrices (see eq. (2) and (3) in Bougerol and Picard (1992)). These conditions collapse to (3) when $p=q=1$. Further results on the probabilistic properties of GARCH(p,q) have then been developed, in particular on the strong mixing (with geometric rate) and regular variation property, imposing suitable smoothing assumptions on the distribution of the rescaled innovation; see e.g. Davis, Mikosch, and Basrak (1999). However, these results on GARCH(p,q) suggest that in general is very difficult to express the required conditions in terms of more primitive conditions, in contrast to Nelson's GARCH$(1,1)$ result. Moreover, they crucially rely on the multivariate Markovian structure of GARCH(p,q).

Giraitis, Kokoszka, and Leipus (1998) found sufficient conditions for strict stationarity of the ARCH(∞). This represents the most general formulation of GARCH processes, generalizing (2) to:

$$\sigma_t^2 = \tau + \sum_{k=1}^{\infty} \psi_k \epsilon_{t-k}^2, \quad a.s., \quad \sum_{k=1}^{\infty} \psi_k < \infty, \tag{4}$$

where $\tau \geq 0$ and $\psi_k \geq 0$. The ARCH(∞) has been introduced by Robinson (1991). The GARCH(p,q) model is obtained choosing exponentially behaving ψ_k, e.g. the GARCH$(1,1)$ follows setting $\tau = \omega/(1-\beta_1)$ and $\psi_j = \alpha_1\, \beta_1^{j-1}$, and the ARCH$(p)$ model when $\tau=\omega$ and $\psi_j = 0$ for $j > p$. Although an hyperbolic behaviour in the ARCH coefficients ψ_j is allowed for, Giraitis, Kokoszka, and Leipus (1998) strict stationarity condition

$$E(z_t^2) \sum_{k=1}^{\infty} \psi_k < 1, \tag{5}$$

implies a bounded second moment of the ϵ_t. Indeed, for GARCH$(1,1)$ (5) equals $E(\beta_1 + \alpha_1 z_t^2) < 1$, the well-known covariance stationarity condition for the ϵ_t (Bollerslev 1986), much stronger than (3).

Based on a suitable nonlinear moving average representation of ARCH(∞) we newly establish, the first contribution of this paper is to prove strict stationarity and ergodicity of the ARCH(∞) model using Nelson (1990)-type conditions, i.e. expressed as primitive functions of the GARCH coefficients and of the rescaled innovation distribution, without imposing the existence of the second moment of the ϵ_t. We discuss implications for hyperbolically decaying specifications of the ψ_i and provide a stronger result for exponentially decaying ψ_i, which includes the GARCH(p, q) case. This is developed in section 2.

Exploiting the well known linear ARMA(m, p) (with $m = max[p, q]$) representation of GARCH(p, q), introduced by Bollerslev (1986), the autocovariance function (ACF) of the squares ϵ_t^2 can be readily shown to be proportional to the ACF of an ARMA(m, p) once the bounded fourth moment conditions of the ϵ_t are imposed. The critical aspect is precisely calculating this constant of proportionality, given by $E[z_t^2 - E(z_t^2)]^2 E(\sigma_t^4)$. For general GARCH($p, q$) the solution has been recently given by Karanasos (1999). In section 3 we propose an alternative way of evaluating $E(\sigma_t^4)$, valid for ARCH(∞). Both the case of exponentially (e.g. the GARCH(p, q) case) and hyperbolically decaying coefficients ψ_j are accounted for. Necessary and sufficient conditions for weak stationarity of the levels ϵ_t and the squares ϵ_t^2 are provided. Finally, we discuss the implications of the covariance stationarity conditions on the memory of the squares. It follows that covariance stationarity of the ϵ_t rules out the possibility of long memory in the ϵ_t^2.

The proofs for the results of both sections are reported in the Appendix.

2 Strict stationarity of ARCH(∞)

For ARCH(∞), it is well known that σ_t^2 admits the following representation, readily obtained by using (1) and (4) recursively (cf. Giraitis, Kokoszka, and Leipus (1998)),

$$\sigma_t^2 = \tau \sum_{l=0}^{\infty} M_l(t), \quad a.s., \tag{6}$$

with $M_0(t) := 1$ and

$$M_l(t) := \sum_{j_1,..,j_l=1}^{\infty} \psi_{j_1}...\psi_{j_l} z_{t-j_1}^2 ... z_{t-j_1-...-j_l}^2, \qquad l \geq 1. \tag{7}$$

Note that each $M_l(t) = M_l(z_{t-1}, z_{t-2}, ...)$ $(l \geq 1)$, i.e. each $M_l(t)$ is a function of all the lagged values of z_t.

However, σ_t^2 admits other 'nonlinear moving average' representations, in contrast to linear processes. Nelson (1990), who focused on GARCH$(1,1)$, considered the following

$$\sigma_t^2 = \omega \sum_{k=0}^{\infty} N_k(t), \quad a.s., \tag{8}$$

with

$$N_k(t) := \prod_{j=1}^{k} (\beta_1 + \alpha_1 z_{t-j}^2), \qquad k \geq 0, \tag{9}$$

with $\prod_{i=a}^{b} c_i = 1$ when $a > b$ for any sequence $\{c_i\}$. In contrast to the $M_l(t)$, note that $N_k(t) = N_k(z_{t-1}, ..., z_{t-k})$ $(k \geq 1)$, i.e. each $N_k(t)$ is a function of at most the k-th lagged value of the rescaled innovation.

The equivalent representation of the ARCH(∞) is given in the following result, generalizing (8)-(9).

Theorem 1 *For ARCH(∞) (4), given $\sum_{j=1}^{\infty} \psi_j < \infty$,*

$$\sigma_t^2 = \omega \sum_{k=0}^{\infty} N_k(t), \quad a.s., \tag{10}$$

setting $\omega := \tau / (\sum_{i=1}^{\infty} \psi_i)$ and

$$N_k(t) := \tag{11}$$
$$\psi_{k+1} + \sum_{l=1}^{k} \left(\sum_{j_1=1}^{k-l+1} \sum_{j_2=1}^{k-l+2-j_1} \cdots \sum_{j_l=1}^{k-j_1-...-j_{l-1}} \psi_{j_1}\psi_{j_2}...\psi_{j_l}\psi_{k-j_1-...j_l+1} z_{t-j_1}^2 ... z_{t-j_1-...-j_l}^2 \right), \; k \geq 0,$$

where $\sum_{i=a}^{b} c_i = 0$ when $a > b$ for any sequence $\{c_i\}$.

Note that, equivalently, when $\psi_1 > 0$, one could set $N_0(t) := 1, \omega := \psi_1 \tau / (\sum_{i=1}^{\infty} \psi_i)$ and divide each $N_k(t)\,(k \geq 1)$ in (11) by ψ_1. This is the parameterization used in (9) for GARCH$(1,1)$.

Expanding (11) for $k = 0, 1, 2, 3, ...$ yields

$$\begin{aligned}
N_0(t) &= \psi_1, \\
N_1(t) &= \psi_2 + \psi_1^2 z_{t-1}^2, \\
N_2(t) &= \psi_3 + \psi_1\psi_2(z_{t-1}^2 + z_{t-2}^2) + \psi_1^3 z_{t-1}^2 z_{t-2}^2, \\
N_3(t) &= \psi_4 + (\psi_1\psi_3 z_{t-1}^2 + \psi_2^2 z_{t-2}^2 + \psi_1\psi_3 z_{t-3}^2) \\
&\quad +\psi_1^2\psi_2(z_{t-1}^2 z_{t-2}^2 + z_{t-1}^2 z_{t-3}^2 + z_{t-2}^2 z_{t-3}^2) + \psi_1^4 z_{t-1}^2 z_{t-2}^2 z_{t-3}^2, \\
N_4(t) &= \psi_5 + ...
\end{aligned}$$

Summing up terms across the $N_k(t)$ involving, in turn, no z_{t-i}^2, the singles z_{t-i}^2, the couples $z_{t-i}^2 z_{t-j}^2$ $(i \neq j)$, and so forth, and dividing by $\sum_{j=1}^{\infty} \psi_j$, yields, respectively, the $M_0(t)$, $M_1(t)$, $M_2(t), ...$ as in (7). Only for ARCH(p) the two representations (6) and (11) coincides as, in this case, $\psi_j = 0$ for $j > p$ yielding $M_l(t) = N_l(t)$ $(l \geq 1)$.

Using the suitable 'nonlinear moving average' representation prove to be crucial when looking at the probabilistic properties of the ARCH(∞), as shown by Nelson (1990) with respect to the GARCH$(1, 1)$ case. Using (11) we now establish conditions which imply strict stationarity and ergodicity of ARCH(∞).

Assume that $\gamma := E(ln z_t^2)$ is well defined (eventually unbounded). Set

$$\lambda := \begin{cases} \frac{\gamma}{2}, & \gamma < 0, \\ \frac{3(\gamma+\delta)}{2}, & \gamma \geq 0, \end{cases} \tag{12}$$

for an arbitrary constant $\delta > 0$.

Theorem 2 *Let $0 < \tau < \infty$ and $\psi_k > 0$ for at least one $k \geq 1$. If*

$$e^{\lambda} \sum_{i=1}^{\infty} \psi_i < 1, \tag{13}$$

then for the ARCH(∞) model (4), for any t,

$$\tau \leq \sigma_t^2 < \infty \quad a.s.,$$

and σ_t^2 is strictly stationary and ergodic, with a well-defined nondegenerate probability measure on $[\tau, \infty)$.

Remark 2.1 Unlike GARCH(p, q), a multivariate Markovian representation of ARCH(∞) does not always exist and thus one cannot rely on it to establish strict stationarity. Indeed, our result allows for hyperbolically decaying coefficients such as $\psi_j \sim c\, j^{-\delta}$ when $j \to \infty$ with $0 < c, (\delta - 1) < \infty$, where $a(x) \sim b(x)$ as $x \to x_0$ means that $a(x)/b(x) \to 1$.

For instance, when the ψ_j are chosen to be the AR(∞) coefficients of ARFIMA(p, d, q), then $\sum_{j=1}^{\infty} \psi_j = 1$ and thus a sufficient condition for (13) would be the standard normalization $E(z_t^2) = 1$, as this implies $\lambda < 0$. Note that, relying on this parameterization, $E(z_t^2) = 1$ is ruled out by the Giraitis, Kokoszka, and Leipus (1998) condition (5). Indeed, it is shown below (cf. Theorem 4) that (5) expresses the necessary and sufficient condition for covariance stationarity of the ϵ_t, much stronger than (13).

Remark 2.2 Condition (13) is a sufficient condition for strict stationarity of ARCH(∞) although we conjecture to be very close to the necessary one. For instance, for GARCH$(1, 1)$, (13) is

$$\beta_1 + \alpha_1 e^{\lambda} < 1,$$

slightly stronger than (3), yet still allowing $\alpha_1 + \beta_1 > 1$ when $E(z_t^2) = 1$.

Remark 2.3 When $\tau = 0$, Giraitis, Kokoszka, and Leipus (1998) show that $\sigma_t^2 = 0$ a.s. is the unique solution of (4).

Remark 2.4 In analogy with the asymptotic results obtained relatively to the Gaussian PMLE of GARCH$(1, 1)$ (Lumsdaine 1996) (Lee and Hansen 1994), condition (13) appears as the minimal regularity condition required in order to establish the asymptotic distribution theory of the Gaussian PMLE of ARCH(∞) for some finite-dimensional parameterization of the $\psi_j = \psi_j(\xi)$ where ξ is a $p \times 1$ vector. This is the topic of forthcoming research.

For exponentially decaying ψ_j, as e.g. for GARCH(p, q), when $\psi_j \leq A\, \rho^{j-1}\, (j \geq 1)$ for some constants $0 < A < \infty$, $0 < \rho < 1$, (13) becomes

$$\rho + Ae^{\lambda} < 1. \tag{14}$$

However, we can exploit the special structure of exponentially decaying ψ_j and obtain weaker strict stationarity conditions than (14) as follows.

Theorem 3 *Let* $0 < \tau < \infty$ *and* $\psi_j \leq A\rho^{j-1}\,(j \geq 1)$ *for some* $0 < \alpha \leq A < \infty$, $0 < \rho < 1$ *with* $\psi_j \sim \alpha\rho^{j-1}$ *as* $j \to \infty$. *If*

$$\rho + Ae^{\lambda} < 2, \tag{15}$$

and

$$E[ln(\rho + \alpha z_t^2)] < 0, \tag{16}$$

then for the ARCH(∞*) model (4), for any* t,

$$\tau \leq \sigma_t^2 < \infty \quad a.s.,$$

and σ_t^2 *is strictly stationary and ergodic, with a well-defined non-degenerate probability measure on* $[\tau, \infty)$.

Remark 3.1 Unlike the case of exponentially decaying coefficients of Theorem 3, full knowledge of the ψ_j for the hyperbolically decaying case, such as e.g. $\psi_j = cj^{-\delta}$, $0 < c, (\delta - 1) < \infty\,(j \geq 1)$, or the autoregressive coefficients of ARFIMA(p, d, q), does not help in finding weaker conditions than (13).

Remark 3.2 In the GARCH$(1, 1)$ case, where $A = \alpha$, it follows that (15) is redundant, in agreement with Nelson (1990, Theorem 2).

Let us focus on the GARCH(p, q) case (cf. (2)). Setting

$$\beta(L) := \beta_1 L + ... + \beta_p L^p,$$

L being the lag operator $(L\, z_t = z_{t-1})$, we assume that for complex valued z

$$|\, 1 - \beta(z)\, | \neq 0, |z| \leq 1, \tag{17}$$

i.e. the p roots of $|\, 1 - \beta(z)\, | = 0$ all lie outside the unit circle in the complex plane. Note that, although (17) implies

$$\beta(1) < 1,$$

it is not the covariance stationarity condition of the ϵ_t, given instead by (cf. Bollerslev (1986))

$$E(z_t^2)\alpha(1) + \beta(1) < 1, \tag{18}$$

with $\alpha(L) := \alpha_1 L + .. + \alpha_q L^q$. More importantly, note that (17) is implied by Nelson (1990)'s condition (3) for otherwise, if $\beta_1 \geq 1$,

$$E[ln(\beta_1 + \alpha_1 z_t^2)] \geq ln\beta_1 \geq 0,$$

in contrast to (3).

Assume that the roots are all distinct, real and distinct from the roots of $\alpha(L) = 0$. The possibility of complex roots is tightly linked to the non-negativity constraints on the coefficients in the ARCH(∞) representation of GARCH(p, q). Nelson and Cao (1992, Theorem 2) show that for the GARCH($2, q$) case, complex roots are not allowed although the conditions $\beta_1 \geq 0$, $\beta_2 \geq 0$, $\alpha_i \geq 0\,(i = 1, .., q)$ can be substantially relaxed. However, for general values of p it is unclear whether oscillatory behaviours in the ψ_j, induced by complex roots, are always ruled out by the non-negativity constraints.

By standard arguments one obtains

$$\frac{\alpha(L)}{1-\beta(L)} = \alpha(L)\left(\frac{A_1}{1-\rho_1 L} + .. + \frac{A_p}{1-\rho_p L}\right),$$

where ρ_i $(i = 1, .., p)$ define the inverse of the roots of $|1 - \beta(z)| = 0$ and

$$A_i := \frac{1}{(1-\rho_1/\rho_i)...(1-\rho_{i-1}/\rho_i)(1-\rho_{i+1}/\rho_i)..(1-\rho_p/\rho_i)}, \quad i = 1, ..., p,$$

where $A_p = A_1 = 1$ for $p = 1$. It follows that the ARCH(∞) (4) representation of GARCH(p, q) is readily obtained setting

$$\psi_j = \begin{cases} \alpha_1\mu_{j-1} + .. + \alpha_j\mu_0, & j = 1, .., q, \\ \alpha_1\mu_{j-1} + .. + \alpha_q\mu_{j-q}, & j > q \end{cases} \qquad (19)$$
$$\tau = \omega/(1 - \beta_1 - .. - \beta_p),$$

with $\mu_j := A_1\,(\rho_1)^j + .. + A_p\,(\rho_p)^j$, $j \geq 0$. Let $\rho := max\{\rho_1, \rho_2, ..., \rho_p\}$, $A_{(\rho)}$ the corresponding A_i and set $A := (\alpha_1 + .. + \alpha_q)\sum_{i=1}^{p} |A_i|$. Finally set $\alpha := (\alpha_1 + \alpha_2\rho^{-1} + ... + \alpha_q\rho^{1-q})A_{(\rho)}$, yielding $\psi_j \sim \alpha\,\rho^{j-1}$ as $j \to \infty$.

Hence, with these definitions for A, α and ρ, Theorem 3 generalizes Nelson (1990, Theorem 2) to the GARCH(p, q) case.

3 Weak stationarity and memory of ARCH(∞)

The minimal condition for covariance stationarity of the ϵ_t (and thus of the ϵ_t^2), is clearly $\kappa := E(z_t^2) < \infty$, by (1). Following Robinson (1991), who considered case $\kappa = 1$, setting $\psi(L) := 1 - \kappa \sum_{j=1}^{\infty} \psi_j L^j$, we can re-write (4) as

$$\psi(L)\epsilon_t^2 = \kappa\tau + \nu_t, \tag{20}$$

setting $\nu_t := \epsilon_t^2 - \kappa\sigma_t^2$. By (1) and i.i.d.-ness of the z_t, $E(\nu_t \mid \mathcal{F}_{t-1}) = 0$, where $\mathcal{F}_t$ is the σ-field of events induced by the ϵ_s $(s \leq t)$; see Loève (1978, section 27.2, *Extension*) for the definition of conditional expectations when the corresponding unconditional expectations might not exist.

Assume that, for complexed valued z,

$$\exists \quad \delta(z) = \sum_{j=0}^{\infty} \delta_j z^j := \psi^{-1}(z), \;\; \delta_0 = 1, \;\text{ s.t. } \sum_{j=0}^{\infty} \delta_j^2 < \infty. \tag{21}$$

A sufficient condition for (21) is

$$|\psi(z)| \neq 0, \; |z| \leq 1,$$

implying $\psi(1) > 0$ but we want to allow for the possibility that $\psi(1) = 0$.

Given (21), (20) can be re-written as

$$\epsilon_t^2 = \kappa\tau\delta(1) + \sum_{j=0}^{\infty} \delta_j \nu_{t-j}. \tag{22}$$

These simple manipulations suggest that τ, $\kappa < \infty$ and $\kappa \sum_{i=1}^{\infty} \psi_i < 1$ imply covariance stationarity of the ϵ_t. Indeed, it turns out that these are the necessary and sufficient conditions for $E(\epsilon_t^2) < \infty$. This is developed in the following theorem where the necessary and sufficient conditions for covariance stationarity of the squares ϵ_t^2 are also established.

Theorem 4 *Assume that the conditions of Theorem 2 and (21) hold.*
(i) Necessary and sufficient conditions for $E(\sigma_t^2) < \infty$ *are*

$$\tau < \infty, \tag{23}$$
$$\kappa < \infty, \tag{24}$$
$$\kappa \sum_{i=1}^{\infty} \psi_i < 1. \tag{25}$$

Under these conditions

$$\phi := E(\epsilon_t^2) = \kappa\tau/(1 - \kappa \sum_{i=1}^{\infty} \psi_i) \quad < \infty.$$

(ii) Necessary and sufficient conditions for $E(\sigma_t^4) < \infty$ *are*

$$\tau < \infty, \tag{26}$$
$$\theta := E(z_t^2 - \kappa)^2 < \infty, \tag{27}$$
$$\left(\theta \sum_{u=-\infty}^{\infty} \chi_\delta(u)\chi_{\tilde{\psi}}(u)\right) < 1, \tag{28}$$

setting $\tilde{\psi}_0 = 0$, $\tilde{\psi}_k = \psi_k$ $(k \geq 1)$, *and* $\chi_c(u) := \sum_{k=0}^{\infty} c_k c_{k+u}$, $u = 0, \pm 1, ...,$ *for any square integrable sequence* c_i.

Under these conditions the ϵ_t^2 *are covariance stationary with ACF*

$$cov(\epsilon_t^2, \epsilon_{t+u}^2) = E(\nu_t^2)\chi_\delta(u), \;\; u = 0, \pm 1, ...,$$

where

$$E(\nu_t^2) = \theta\,(\phi/\kappa)^2 \left(1 - \theta \sum_{u=-\infty}^{\infty} \chi_\delta(u)\chi_{\tilde{\psi}}(u)\right)^{-1} < \infty.$$

Remark 4.1 The ACF of GARCH(p, q) follows setting

$$\phi = \kappa\omega/(1 - \kappa\alpha(1) - \beta(1)),$$

deriving the $\chi_{\tilde{\psi}}(u)$ and the $\chi_\delta(u)$, and thus $E(\nu_t^2)$, using the ψ_j from (19) and deriving the δ_i from (21), respectively. For example, for GARCH$(1, 1)$, setting $\kappa = 1$ and $\pi_1 := \alpha_1 + \beta_1$, i.e. the 'persistence' parameter, one obtains

$$\chi_\delta(0) = 1 + \frac{\alpha_1^2}{1-\pi_1^2},\ \chi_\delta(u) = \alpha_1 \pi_1^{|u|-1}\left(1 + \frac{\alpha_1\pi_1}{1-\pi_1^2}\right),$$
$$\chi_{\tilde{\psi}}(0) = \frac{\alpha_1^2}{1-\beta_1^2},\ \chi_{\tilde{\psi}}(u) = \frac{\alpha_1^2\beta_1^{|u|}}{1-\beta_1^2},\ u = \pm 1, ...$$

When $\theta = 2$, by means of simple manipulations, (28) yields the well-known covariance stationarity conditions for ϵ_t^2 (cf. Bollerslev (1986, section 3))

$$3\alpha_1^2 + \beta_1^2 + 2\alpha_1\beta_1 < 1, \tag{29}$$

which, in turn, implies (25) (cf. (18)). In remark 4.4 below, we show that, indeed, (28) strictly implies (25) without specifying any particular parameterization of the ψ_j.

Remark 4.2 An alternative proof of sufficiency of (23)-(25) for $E(\sigma_t^2) < \infty$ is in Giraitis, Kokoszka, and Leipus (1998, Theorem 2.1).

When

$$(E(z_t^4))^{1/2} \sum_{j=1}^{\infty} \psi_j < 1, \tag{30}$$

Giraitis, Kokoszka, and Leipus (1998, Theorem 2.1) show that $E(\sigma_t^4) < \infty$. However, (30) is much stronger than (28). For instance, for GARCH(1,1) with $\theta = 2\kappa = 2$, their condition becomes

$$3^{1/2}\alpha_1 + \beta_1 < 1,$$

strictly implying (29), unless $\beta_1 = 0$, the ARCH(1).

Remark 4.3 Hyperbolically decaying specifications of the ψ_i and hence of the δ_i are allowed for. Note that from (25), $\kappa < 1/(\sum_{i=1}^{\infty}\psi_i)$ and thus $\kappa = 1$ is ruled out when $\sum_{i=1}^{\infty}\psi_i = 1$.

Imposing $\kappa = 1$, a choice compatible with covariance stationary levels is obtained using the autoregressive coefficients of stationary ARFIMA(p, d, q) (cf. Robinson and Zaffaroni (1997, section 3)) as follows. Set

$$1 - \sum_{j=1}^{\infty} \bar{\psi}_j L^j := (1-L)^d \frac{a(L)}{b(L)}, \tag{31}$$

where $0 < d < 1/2$, $a(L)$, $b(L)$ are finite order polynomials, all of whose roots are outside the unit circle in the complex plane. Assuming that the non-negativity constraints on the coefficients hold, i.e. $\bar{\psi}_i \geq 0$, set $\psi_i = \bar{\psi}_i$ $(i \geq 2)$ and $\psi_1 = \bar{\psi}_1 \epsilon$ for some given $0 < \epsilon < 1$ (e.g. $\epsilon = 9/10$). Condition (28) is more involved but, again, by a suitable modification of the first L (say) coefficients $\bar{\psi}_j$ $(j = 1, ..., L)$, a feasible sequence of coefficients can be obtained from (31).

Remark 4.4 An alternative, frequency domain, characterization of (28) is

$$\theta \int_{-\pi}^{\pi} f_\delta(\lambda) f_\psi(\lambda) d\lambda < 2\pi, \tag{32}$$

setting

$$f_\delta(\lambda) := |\delta(e^{i\lambda})|^2, \; f_\psi(\lambda) := (1/\kappa^2)\; |1 - \psi(e^{i\lambda})|^2, \;\; -\pi \leq \lambda < \pi,$$

where $\kappa^2 \, f_\psi(\lambda) = 1 + f_\delta^{-1}(\lambda) - 2 \; Re(\delta^{-1}(e^{i\lambda}))$, $-\pi \leq \lambda < \pi$ and $Re(\cdot)$ indicates real part of its argument. This frequency domain specification seems much easier to be computed and might be relevant when imposing covariance stationarity of the ϵ_t^2 in practical estimation, e.g. when estimating ARCH(∞) using the Whittle estimator (Giraitis and Robinson 1998).

The equivalent (32) representation of (28) could be used to show that (28) strictly implies (25). In fact, setting $\theta = 2\kappa = 2$ for simplicity's sake, (32) can be re-written as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_\delta(\lambda) \left(1 - 2 \; Re(\delta^{-1}(e^{i\lambda}))\right) d\lambda < -1/2.$$

However, $f_\delta(\lambda) \geq 0 \, (-\pi \leq \lambda < \pi)$ but it is arbitrary, requiring necessarily

$$1 - 2 \; Re(\delta^{-1}(e^{i\lambda})) = -1 + 2 \sum_{j=1}^{\infty} \psi_j cos(j\lambda) < 0, \; -\pi \leq \lambda < \pi.$$

Given the non-negativity of the ψ_i and $cos(j\lambda) \leq 1$, with the equality achieved for $\lambda = 0 \, (mod. \, 2\pi)$, this is equivalent to

$$-1 + 2 \sum_{j=1}^{\infty} \psi_j < 0,$$

strictly implying $\psi(1) > 0$.

Finally, exploiting the relation between the ψ_i $(i \geq 1)$ and the δ_j $(j \geq 0)$ (cf. (21)), (32) allows to express the time domain characterization (28) more simply as

$$2\kappa \sum_{j=1}^{\infty} \psi_j \chi_\delta(j) < \kappa^2/\theta + \chi_\delta(0) - 1.$$

Remark 4.5 The original formulation of ARCH(∞), analogous to Robinson (1991) but allowing $\kappa \neq 1$, is

$$\sigma_t^2 = \tilde{\tau} + \sum_{k=1}^{\infty} \psi_k(\epsilon_{t-k}^2 - \kappa\tilde{\tau}), \quad a.s., \tag{33}$$

for some $0 < \tilde{\tau} < \infty$. The re-parameterization (33) is clearly permitted only for covariance stationary ϵ_t, i.e. when $\psi(1) > 0$, given that $\tilde{\tau} = \tau/\psi(1)$. Indeed, assume to start directly from (33) rather than from (4). By the non-negativity constraint $\kappa \sum_{i=1}^{\infty} \psi_i \leq 1$ as $\psi(1) < 0$ is not allowed. Imposing $\psi(1) = 0$ and assuming (21), the linear moving average representation (22) for the ϵ_t^2 would then be

$$\epsilon_t^2 = \zeta + \sum_{j=0}^{\infty} \delta_j \nu_{t-j},$$

for any ζ, given that $\psi(L)\epsilon_t^2 = \psi(L)\epsilon_t^2 - \psi(1)\zeta = \psi(L)(\epsilon_t^2 - \zeta)$, which is meaningless.

Giraitis, Kokoszka, and Leipus (1998, Proposition 3.1) show that (30) implies absolute summability of the ACF for the ϵ_t^2, ruling out long memory. However, considering that (30) is stronger than required in order to obtain covariance stationary ϵ_t^2, it seems important to assess the impact of the weaker condition (28) on the memory of the ϵ_t^2.

Insights can be obtained by looking at the linear representation (22) for ϵ_t^2. In fact, it follows that the memory of the ϵ_t^2 is expressed by the asymptotic behaviour of the δ_j as $j \to \infty$. Surprisingly, it turns out that even the much weaker covariance stationarity condition (25) for the levels ϵ_t rules out long memory in the ϵ_t^2. In fact, from (21) and (25),

$$\delta(1) = \sum_{j=0}^{\infty} \delta_j = 1/(1 - \kappa \sum_{j=1}^{\infty} \psi_j) < \infty.$$

(28) ensures that the uncorrelated ν_t have finite variance but the rate of decay of the δ_j, imposed by (25), is already quick enough to imply their absolute summability. We summarize our results on the memory of the ϵ_t^2 as follows.

Theorem 5 *Assume that condition (25) of Theorem 4 hold. Then*

$$\sum_{j=1}^{\infty} \delta_j < \infty, \qquad (34)$$

where

$$\delta_l = \kappa \psi_l + \sum_{s=2}^{l} \kappa^s \sum_{i_1=1}^{l-s+1} \ldots \sum_{i_{s-1}=1}^{l-i_1-\ldots-i_{s-2}-1} \psi_{i_1} \ldots \psi_{i_{s-1}} \psi_{l-i_1-\ldots-i_{s-1}}, \quad l \geq 1. \qquad (35)$$

When the ψ_i decay toward zero slower than exponentially, viz. $\psi_i/\zeta^i \to \infty$ as $i \to \infty$ for any $0 < \zeta < 1$, (34) implies that, as $u \to \infty$,

$$\chi_\delta(u) \sim C\, \psi_u, \qquad (36)$$

for some $0 < C < \infty$.

When $\psi_i \sim c i^{-\delta}$ as $i \to \infty$ for $0 < c, \delta - 1 < \infty$, as e.g. for the ARFIMA$(p,d,q)$ parameterization described in remark 4.3, $E(\epsilon_t^2) < \infty$ implies

$$\chi_\delta(u) \sim C\, u^{-\delta},\ u \to \infty, \qquad (37)$$

for some $0 < C < \infty$ ruling out long memory in the ϵ_t^2.

Under the same assumptions on the asymptotic behaviour of the ψ_i, the exact rate in (37) was also obtained in Giraitis, Kokoszka, and Leipus (1998, Proposition 3.2), although they impose (30), a sufficient condition for $E(\epsilon_t^4) < \infty$. However, Theorem 5 makes clear that whereas the bounded second moment conditions impart the degree of memory of the ϵ_t^2, the stronger bounded fourth moment necessary and sufficient conditions (26)-(28) ensure that the martingale difference sequence ν_t, viz. the innovations in the linear representation of the squares (22), are square integrable, but do not change the memory implications of the model. This double role of the coefficients is simply a by-product of the ARCH(∞) nonlinearity.

In the case of exponentially decaying ψ_i, i.e. when it does exist a $0 < \zeta < 1$ such that $\psi_i/\zeta^i \to c$ as $i \to \infty$ for some $0 < c < \infty$, it clearly follows that

$$\chi_\delta(u) \sim C\,\delta^u,\; u \to \infty,$$

for some $0 < \delta < 1$, $0 < C < \infty$, e.g. the GARCH(p, q) case.

A Appendix

PROOF OF THEOREM 1: Re-parameterize the coefficients

$$\psi_j = c\theta_{j-1}, \;\; j \geq 1, \tag{38}$$

for some constant $0 < c < \infty$, which is absolutely innocuous. Clearly $\sum_{k=0}^{\infty} \theta_k < \infty$ by assumption. Then, expanding $M_0(t)$, $M_1(t)$, $M_2(t)$,.. yields

$$\begin{aligned}
&M_0(t) = 1, \\
&M_1(t) = c\theta_0 z_{t-1}^2 + c\theta_1 z_{t-2}^2 + c\theta_2 z_{t-3}^2 + ... \\
&M_2(t) = c^2\theta_0^2 z_{t-1}^2 z_{t-2}^2 + c^2\theta_0\theta_1(z_{t-1}^2 z_{t-3}^2 + z_{t-2}^2 z_{t-3}^2) \\
&+c^2\theta_1^2 z_{t-2}^2 z_{t-4}^2 + c^2\theta_0\theta_2(z_{t-1}^2 z_{t-4}^2 + z_{t-3}^2 z_{t-4}^2) + ... \\
&M_3(t) = c^3\theta_0^3 z_{t-1}^2 z_{t-2}^2 z_{t-3}^2 + ...
\end{aligned}$$

Premultiplying each $M_l(t)\,(l \geq 0)$ by $c\sum_{k=0}^{\infty}\theta_k$, expanding terms and grouping together all terms such that their coefficients $c^{r+1}\theta_{i_1}...\theta_{i_r}$ $(r \geq 0,\, i_1,..,i_r \geq 0)$ satisfy

$$r + i_1 + ... + i_r = k, \;\; k = 0, 1, .. \tag{39}$$

and defining them as $N_k(t)$, yields

$$\begin{aligned}
N_0(t) &= c\theta_0, \\
N_1(t) &= c\theta_1 + c^2\theta_0^2 z_{t-1}^2, \\
N_2(t) &= c\theta_2 + +c^2\theta_0\theta_1 z_{t-1}^2 + c^2\theta_1\theta_0 z_{t-2}^2 + c^3\theta_0^3 z_{t-1}^2 z_{t-2}^2, \\
\vdots &= \vdots \\
N_k(t) &= c\theta_k + c^2(\theta_0\theta_{k-1} z_{t-1}^2 + \theta_1\theta_{k-2} z_{t-2}^2 + ... + \theta_{k-1}\theta_0 z_{t-k}^2) \\
&\quad +... + c^{k+1}\theta_0^k z_{t-1}^2 z_{t-2}^2 \cdots z_{t-k}^2, \\
\vdots &= \vdots
\end{aligned}$$

Substituting back the $c\theta_{j-1}$ in terms of the ψ_j and setting $\omega = \tau/(\sum_{i=1}^{\infty}\psi_i)$ concludes. The re-parameterization (38) is used as a device in that it allows to identify, within each summand, the number of factors z_t^2 and, as a consequence, to characterize all the elements corresponding to each $N_k(t)$. For instance, with respect to $N_1(t)$, both $c\theta_1$ and $c^2\theta_0^2$ satisfy (39) for $k = 1$, for $N_2(t)$, $c\theta_2$, $c^2\theta_0\theta_1$, $c^2\theta_1\theta_0$ and $c^3\theta_0^3$ satisfy (39) for $k = 2$ and so on.

Note that the only condition required for deriving the nonlinear moving average representation (11) is $\sum_{i=1}^{\infty} \psi_i < \infty$. □

PROOF OF THEOREM 2: Take an arbitrary constant integer $M > 0$ with $M < k$. Then, for any $k \geq 1$, split the RHS of (11) as

$$N_k(t) = \sum_{l=0}^{[k/M]} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t) + \sum_{l=[k/M]+1}^{k} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t), \quad (40)$$

where $[\cdot]$ is the integer part of its argument and

$$B_l(t) := \psi_{j_1}\psi_{j_2}\ldots\psi_{j_l}\psi_{k-j_1-\ldots j_l+1}\, z_{t-j_1}^2 \ldots z_{t-j_1-\ldots-j_l}^2. \quad (41)$$

Let us dispose of the first sum on the RHS of (40). As $l \leq [k/M]$, one obtains

$$\sum_{l=0}^{[k/M]} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t)$$
$$= \sum_{l=0}^{[k/M]} \left\{ \left(\sum_{j_1=1}^{M} + \sum_{j_1=M+1}^{k-l+1} \right) \cdots \left(\sum_{j_l=1}^{M} + \sum_{j_l=M+1}^{k-j_1-\ldots-j_{l-1}} \right) B_l(t) \right\} \quad (42)$$

The term in the $\{\,\}$-brackets on the RHS of (42) is made by the sum of three terms. One involving the sums $\sum_{j_i=1}^{M}$ $(i = 1, .., l)$ only:

$$\sum_{l=0}^{[k/M]} \left\{ \sum_{j_1=1}^{M} \cdots \sum_{j_l=1}^{M} \psi_{j_1}\psi_{j_2}\ldots\psi_{j_l}\psi_{k-j_1-\ldots j_l+1}\, z_{t-j_1}^2 \ldots z_{t-j_1-\ldots-j_l}^2 \right\}. \quad (43)$$

One involving the sums $\sum_{j_i=M+1}^{k-l+1-j_1-..-j_{i-1}+(i-1)}$ $(i = 1, ..., l)$ only:

$$\sum_{l=0}^{[k/M]} \times \quad (44)$$
$$\left\{ \sum_{j_1=M+1}^{k-l+1} \cdots \sum_{j_l=M+1}^{k-j_1-\ldots-j_{l-1}} \psi_{j_1}\psi_{j_2}\ldots\psi_{j_l}\psi_{k-j_1-\ldots j_l+1}\, z_{t-j_1}^2 \ldots z_{t-j_1-\ldots-j_l}^2 \right\}.$$

and the 'mixed' term

$$\sum_{l=0}^{[k/M]} \sum_{s=1}^{l-1} \widetilde{\sum}_{l,s} \left\{ \sum_{j_{i_1}=M+1}^{k-l+i_1-j_1-..-j_{i_1-1}} \cdots \sum_{j_{i_{l-s}}=M+1}^{k-l+i_{l-s}-j_1-..-j_{i_{l-s}-1}} \sum_{j_{r_1}=1}^{M} \cdots \sum_{j_{r_s}=1}^{M} \times \right.$$
$$\left. \psi_{j_1}\psi_{j_2}\ldots\psi_{j_l}\psi_{k-j_1-\ldots j_l+1}\, z_{t-j_1}^2 z_{t-j_1-j_2}^2 .. z_{t-j_1-..-j_l}^2 \right\}, \quad (45)$$

where $\widetilde{\sum}_{l,s} := \sum_{\substack{\{r_1,...,r_s\},\subset\{1,..,l\},\\ \{r_1,..,r_s\}\cup\{i_1,..,i_{l-s}\}=\{1,...,l\}}}$, i.e. it selects all the groups of indexes of dimension s (with $s = 1, .., l-1$) drawn from a number l of them and thus for any s and any permutation $j_{i_1} + ... + j_{i_{l-s}} + j_{r_1} + ... + j_{r_s} = j_1 + ... + j_l$. Note that $\widetilde{\sum}_{l,s} 1 = \binom{l}{s}$.

By Dudley (1989, Theorem 8.3.5), with probability one there exists a constant $K < \infty$ such that for all $k > K$

$$\prod_{j=1}^{k} z_{t-j}^2 = O(e^{\lambda k}), \quad a.s. \tag{46}$$

Imposing $E \mid ln\, z_t^2 \mid^{2+\epsilon} < \infty$ for some arbitrary $\epsilon > 0$, by the law of iterated logarithm for i.i.d. variates (see e.g. Stout (1974, Corollary 5.2.1)) one can refine case $\gamma = 0$ in (46) setting $\delta = \delta(k) = [9/2\mu(ln\, ln\, \mu\, k)/k]^{1/2}$ (cf. definition of λ in (12)) with $\mu := E(ln\, z_t^2)^2$.

For (43), for some $0 < \zeta < 1$, writing $\sum_{l=0}^{[k/M]} = \sum_{l=0}^{[\zeta\, k/M]-1} + \sum_{[\zeta\, k/M]}^{[k/M]}$, yields, for ζ suitably small,

$$\begin{aligned}(43) \qquad &= O\left(\psi_k \sum_{l=0}^{\infty} (e^{\lambda} \sum_{j=1}^{M} \psi_j)^l + \sum_{l=[\zeta\, k/M]}^{[k/M]} (e^{\lambda} \sum_{j=1}^{M} \psi_j)^l\right) \\ &= O\left(\psi_k + (e^{\lambda} \sum_{j=1}^{M} \psi_j)^{\zeta\, k/M}\right), \quad a.s.\end{aligned}$$

Concerning (44), along the same lines,

$$\begin{aligned}(44) \qquad &= O\left(\psi_k \sum_{l=0}^{\infty} (e^{\lambda} \sum_{j=M+1}^{\infty} \psi_j)^l + \sum_{l=[\zeta\, k/M]}^{[k/M]} (e^{\lambda} \sum_{j=M+1}^{\infty} \psi_j)^l\right) \\ &= O\left(\psi_k + (e^{\lambda} \sum_{j=M+1}^{\infty} \psi_j)^{\zeta\, k/M}\right), \quad a.s.\end{aligned}$$

For the 'mixed' term,

$$\begin{aligned}(45) \qquad &= O\left(\psi_k + \sum_{l=[\zeta\, k/M]}^{[k/M]} e^{\lambda\, l} \sum_{s=1}^{l} \binom{l}{s} (\sum_{j=M+1}^{\infty} \psi_j)^s (\sum_{j=1}^{M} \psi_j)^{l-s}\right) \\ &= O\left(\psi_k + (e^{\lambda} \sum_{j=1}^{\infty} \psi_j)^{\zeta\, k/M}\right), \quad a.s.\end{aligned}$$

Finally, for the second term on the RHS of (40),

$$\sum_{l=[k/M]+1}^{k} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t) \leq \sum_{l=[k/M]+1}^{k} (e^{\lambda} \sum_{j=1}^{\infty} \psi_j)^l$$
$$= O\left((e^{\lambda} \sum_{j=1}^{\infty} \psi_j)^{k/M} \right), \quad a.s.$$

Strict stationarity, ergodicity and non-degenerateness of the distribution of the σ_t^2 follow along the lines of Nelson (1990, proof of Theorem 2). □

PROOF OF THEOREM 3: Following Theorem 2, take an arbitrary constant integer $M > 0$ with $M < k$ and split the RHS of (11) as

$$N_k(t) \leq \sum_{l=0}^{[k/M]} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t) + \sum_{l=[k/M]+1}^{k} \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} B_l(t), \quad (47)$$

with $B_l(t)$ defined in (41). Given $l \leq [k/M]$, we have seen that the first sum on the RHS of (47) can be written as the sum of three terms (43)-(45). The term involving the sums $\sum_{j_i=1}^{M}$ $(i = 1, .., l)$ only is bounded by

$$O\left(\sum_{l=0}^{[k/M]} (\rho)^{k-l} (\, M\, A\, e^{\lambda}\,)^l \right) = O\left((\rho)^k [1 + ((\rho)^{-1}\, M\, A\, e^{\lambda})^{[k/M]}] \right)$$
$$= O\left((\rho')^k \right), \quad a.s.,$$

using (46) and choosing M large enough such that

$$\rho' := \rho((\rho)^{-1} M\, A\, e^{\lambda})^{1/M} < 1.$$

Second, concerning the term involving the sums $\sum_{j_i=M+1}^{k-l+1-j_1-\ldots-j_{i-1}+(i-1)}$ $(i = 1, .., l)$ only, one obtains

$$O\left(A \sum_{l=0}^{[k/M]} (\rho)^{k-l} \left\{ \sum_{j_1=M+1}^{k-l+1} \cdots \sum_{j_l=M+1}^{k-j_1-\ldots-j_{l-1}} (\alpha')^l z_{t-j_1}^2 \ldots z_{t-j_1-\ldots-j_l}^2 \right\} \right)$$
$$= O\left(\sum_{l=0}^{k} (\rho)^{k-l} (\alpha')^l \sum_{j_1=1}^{k-l+1} \cdots \sum_{j_l=1}^{k-j_1-\ldots-j_{l-1}} z_{t-j_1}^2 \ldots z_{t-j_1-\ldots-j_l}^2 \right)$$
$$= O\left(\prod_{i=1}^{k} (\rho + \alpha' z_{t-i}^2) \right) = O\left(exp(k/2E[ln(\rho + \alpha' z_t^2)]) \right), \quad a.s.,$$

and $\alpha' := \alpha + \delta_M$ for some $\delta_M > 0$ with $\delta_M = o(1)$, arbitrarily small by choosing M large enough.

For the 'mixed' term, involving the sums $\sum_{j_i=1}^{M}$, $\sum_{j_h=M+1}^{k-l+1-j_1-...-j_{h-1}+(h-1)}$ $(i,h=1,..,l)$, using (46),

$$\begin{aligned} & O\left(\sum_{l=0}^{[k/M]}(\rho)^{k-l}\sum_{s=1}^{l-1}(M\,A\,e^{\lambda})^s(\alpha' e^{\lambda})^{l-s}\binom{l}{s}\binom{k-s}{l-s}\right) \\ &= O\left(\sum_{l=0}^{[k/M]}(\rho)^{k-l}(\alpha' e^{\lambda})^l\binom{k}{l}{}_2F_1(-l,-l;-k;-M\,A/\alpha')\right) \\ &= o\left((\rho)^k[(1+M\,A\,/\alpha')\,(\rho)^{-1}\alpha'\,e^{\lambda}]^{[k/M]}\binom{k+1}{[k/M]}\right) \\ &= o\left(k^{3/2}(\rho'')^k\right),\ a.s., \end{aligned}$$

where ${}_2F_1(\cdot,\cdot;\cdot,\cdot)$ is the $(2,1)$ generalized hypergeometric series (Gradshteyn and Ryzhik 1994, Section 9.1). The result follows using Gradshteyn and Ryzhik (1994, # 9.132) yielding

$${}_2F_1(-l,-l;-k;-z) = 2(1+z)^l\frac{\Gamma(-k)}{\Gamma(-l)\Gamma(-k+l)}{}_2F_1(-l,-k+l;1;1/(1+z)),$$

where $\Gamma(\cdot)$ is the Gamma function, simplifying terms, noting that as $z\to\infty$ ${}_2F_1(-l,-k+l;1;1/(1+z)) = o(\binom{k}{l})$, using Gradshteyn and Ryzhik (1994, # 0.156) and Stirling's formula (Brockwell and Davis 1987, pg. 522) and choosing M large enough such that

$$\rho'' := \rho\left((1+M\,A\,/\alpha')\,(\rho)^{-1}\alpha'\,e^{\lambda}\right)^{1/M} < 1.$$

Assume $1 \le \rho + Ae^{\lambda}$. The last term on the RHS of (47) is bounded by

$$\begin{aligned} & O\left(\sum_{l=[k/M]+1}^{k}(\rho)^{k-l}A^{l+1}\sum_{j_1=1}^{k-l+1}\cdots\sum_{j_l=1}^{k-j_1-...-j_{l-1}}z_{t-j_1}^2\cdots z_{t-j_1-...-j_l}^2\right) \\ &= O\left((\rho)^{k-[k/M]}\,(A\,e^{\lambda})^{[k/M]}\binom{k}{[\frac{k}{M}]}{}_2F_1(1,-k+[\frac{k}{M}];1+[\frac{k}{M}],-Ae^{\lambda}/\rho\right) \\ &= O(k^{3/2}c^k),\ a.s., \end{aligned}$$

for some constant $c = c(\rho, Ae^{\lambda}, M)$ with $0 < c < 1$, using Gradshteyn and Ryzhik (1994, # 3.196 (1)) such that

$$
\begin{aligned}
&{}_2F_1(1, -k + [\frac{k}{M}]; 1 + [\frac{k}{M}], -Ae^{\lambda}/\rho) \\
&= \rho^{-k+[k/M]}\,(A\,e^{\lambda})^{-[k/M]}[\frac{k}{M}]\int_0^{Ae^{\lambda}} (x+\rho)^{k-[k/M]}(Ae^{\lambda}-x)^{[k/M]-1}dx,
\end{aligned}
$$

where the integral on the RHS above is $O(c^k)$ by splitting $\int_0^{Ae^{\lambda}} = \int_0^{1-\rho-\epsilon} + \int_{1-\rho-\epsilon}^{Ae^{\lambda}}$ for some $\epsilon > 0$ with $Ae^{\lambda} - 1 < 1 - \rho - \epsilon < 1 - \rho$ and choosing M suitably. When $\rho + Ae^{\lambda} < 1$ the result follows trivially (cf. (14)). □

PROOF OF THEOREM 4: (i) Assume that $E(\sigma_t^2) < \infty$. Then necessity of (23)-(25) follows taking expectation on both sides of (4). On the other hand, assume that (23)-(25) hold. Then, premultiply both sides of (4) by $1_t := 1(\bigcap_{j=0}^{\infty}\{\sigma_{t-j}^2 < M\})$ for some constant $M < \infty$ where $1(A)$ equals one when the event A holds and zero otherwise. Note that when σ_t^2 is bounded from above, so are all the σ_{t-k}^2 for $k \geq 1$. Setting $P(A)$ equal to the probability of the event A, taking expectations,

$$E(\sigma_t^2 1_t) \leq \tau P(\sigma_t^2 < M)/(1 - \kappa \sum_{i=1}^{\infty} \psi_i),$$

as, using the law of iterated expectations,

$$E(\sigma_{t-k}^2 1_t) = E(P(\sigma_t^2 < M, ..., \sigma_{t-k+1}^2 < M \,|\, \sigma_{t-k}^2, ...)1_{t-k}\sigma_{t-k}^2) \leq E(\sigma_{t-k}^2 1_{t-k}) = E(\sigma_t^2 1_t),$$

the latter equality holding by strict stationarity. Finally, letting $M \to \infty$, $P(\sigma_t^2 < M) \to 1$ by Theorem 2.
(ii) Assume that $E(\sigma_t^4) < \infty$. Then, (26)-(28) follow squaring both terms in (4) and taking expectations,

$$E(\sigma_t^4) = (\phi/\kappa)^2 + \theta E(\sigma_t^4) \sum_{j_1, j_2 = 1}^{\infty} \psi_{j_1}\psi_{j_2} \left(\sum_{i=0}^{\infty} \delta_i \delta_{i+(j_1-j_2)} \right),$$

using

$$E[(\epsilon_t^2 - \phi)(\epsilon_{t+u}^2 - \phi)] = E(\nu_t^2) \sum_{j=0}^{\infty} \delta_j \delta_{j+u}, \quad u = 0, \pm 1, ...,$$

as the ν_t are square integrable martingale differences and $\sum_{i=0}^{\infty} \delta_i^2 < \infty$ by (21). Sufficiency is obtained following the corresponding part in the proof of (i). □

PROOF OF THEOREM 5: From

$$(1 + \sum_{j=1}^{\infty} \delta_j L^j)(1 - \kappa \sum_{i=1}^{\infty} \psi_i L^i) = 1,$$

by the fundamental theorem for polynomials and simple yet tedious calculations, (35) follows. By a truncating argument, similar to the one used in the proof of Theorem 2, it straightforwardly follows that, as $j \to \infty$,

$$\delta_j = \kappa \psi_j + O\left((\kappa \sum_{i=1}^{\infty} \psi_i)^{\zeta j} \right), \tag{48}$$

for some $0 < \zeta \leq 1$, re-discovering that condition (25) is needed for summability of the δ_j. Note that, when $\psi_j \geq 0 \, (j \geq 1)$, then (35) implies $\delta_j \geq 0 \, (j \geq 0)$.

Hence, for any $u > 0$ and some $0 < c < \infty$,

$$\chi_\delta(u) = \sum_{j=0}^{u} \delta_j \delta_{j+u} + \sum_{j=u+1}^{\infty} \delta_j \delta_{j+u} \sim c\, \delta_u (1 + o(1)), \;\; u \to \infty,$$

given that, for u large enough,

$$\sum_{j=u+1}^{\infty} \delta_j \delta_{j+u} \leq \delta_{u+n} \sum_{j=u+1}^{\infty} \delta_j = o(\delta_u),$$

for some constant $1 \leq n < \infty$. Finally, comparing (48) with (35) and given $\kappa > 0, \psi_i \geq 0 \, (i \geq 1)$, yields

$$\delta_k \sim c \psi_k, \;\; k \to \infty.$$

□

References

BOLLERSLEV, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 302–327.

BOUGEROL, P., AND N. PICARD (1992): "Stationarity of GARCH processes and of some nonnegative time series," *Journal of Econometrics*, 52, 115–127.

BROCKWELL, P., AND R. DAVIS (1987): *Time series: theory and methods.* New York: Springer Verlag.

DAVIS, R., T. MIKOSCH, AND B. BASRAK (1999): "Sample ACF of multivariate stochastic recurrence equations with applications to GARCH," *Preprint.*

DUDLEY, R. (1989): *Real analysis and probability.* California: Pacific Grove.

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

GIRAITIS, L., P. KOKOSZKA, AND R. LEIPUS (1998): "Stationary ARCH models: dependence structure and central limit theorem," forthcoming *Econometric Theory.*

GIRAITIS, L., AND P. M. ROBINSON (1998): "Least squares (Whittle) estimate procedure for ARCH models," *Preprint.*

GRADSHTEYN, I., AND I. RYZHIK (1994): *Table of integrals, series and products.* San Diego: Academic Press, fifth edn.

KARANASOS, M. (1999): "The second moment and the autocovariance function of the squared errors of the GARCH model," *Journal of Econometrics*, 90, 63–76.

LEE, S., AND B. HANSEN (1994): "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator," *Econometric Theory*, 10, 29–52.

LOÈVE, M. (1978): *Probability Theory II.* New York: Springer Verlag, fourth edn.

LORETAN, M., AND P. PHILLIPS (1994): "Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial instruments," *Journal of Empirical Finance*, 1, 211–248.

LUMSDAINE, R. (1996): "Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models," *Econometrica*, 64/3, 575–596.

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

NELSON, D., AND C. CAO (1992): "Inequality constraints in the univariate GARCH model," *Journal of Business & Economic Statistics*, 10, 229–235.

ROBINSON, P. M. (1991): "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression," *Journal of Econometrics*, 47, 67–84.

ROBINSON, P. M., AND P. ZAFFARONI (1997): "Modelling nonlinearity and long memory in time series," in *Nonlinear Dynamics and Time Series, Fields Institute Communications, vol. 11*, ed. by C. Cutler,and D. Kaplan, pp. 161–170. Providence: American Mathematical Society.

STOUT, W. (1974): *Almost sure convergence*. New York; London: Academic Press Inc.

WEISS, A. (1986): "Asymptotic theory for ARCH models: estimation and testing," *Econometric Theory*, 2, 101–131.

PAOLO ZAFFARONI
Servizio Studi
Banca d'Italia
Via Nazionale, 91
00184 Roma Italy
tel. + 39 06 4792 4178
zaffaroni.paolo@insedia.interbusiness.it