Gremars Working Document (26 March 2005)

THE L^2-STRUCTURES OF STANDARD AND SWITCHING-REGIME GARCH MODELS

CHRISTIAN FRANCQ ,* *GREMARS Université Lille 3*

JEAN-MICHEL ZAKOIAN,** *GREMARS Université Lille 3 and CREST*

Abstract

This paper analyzes the probabilistic structure of Markov-switching GARCH(p, q) models, in which the volatility process is driven by a finite state-space Markov chain. We give necessary and sufficient conditions for the existence of moments of any order. We find that the squares and higher order powers of the process have the L^2 structures of ARMA processes, and hence admit ARMA representations. These results are applicable to standard GARCH models and have statistical implications in terms of order identification and parameter estimation.

Keywords: ARMA representation, GARCH, HMM, Markov-switching models.
2000 Mathematics Subject Classification: Primary 60J05
Secondary 60G10; 60G12; 62M10

1. Introduction

In the statistical analysis of returns on financial assets, one is confronted with the task of modeling time series with a very complicated structure. This structure is characterized by a set of empirical features that are common to many financial data, such as stock prices or indices, and exchange rates. These "stylized facts" are lack of serial correlation, heavy-tailed marginal distributions, volatility clustering (large and small observations tend to appear in clusters), and slow decrease of the sample autocorrelations of the squares (or absolute values) of the data. These properties make standard time series models, such as ARMA processes, inappropriate for such series.

* Postal address: GREMARS, UFR MSES, Université Lille 3, Domaine du Pont de bois, BP 149, 59653 Villeneuve d'Ascq Cedex, France
** Postal address: CREST, 3 Avenue P. Larousse, 92245 Malakoff Cedex,France

The most popular class of volatility models is the autoregressive conditional heteroskedasticity (ARCH) specification introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986). There is a still growing literature on GARCH and related models (see *e.g.* Berkes and Horváth (2003), Giraitis, Kokoszka and Leipus (2000), Hilebrand (2005), Klüppelberg, Lindner and Maller (2004), Bai, Russel and Tiao (2003), for recent references; see Mikosch (2001) for a comprehensive review on GARCH and stochastic volatility processes.) These models have the form $\epsilon_t = \sqrt{h_t}\eta_t$, where (h_t) is a positive stochastic process, called volatility, and (η_t) is an iid centered sequence, with h_t independent of η_t. Under moment conditions (ϵ_t) is an uncorrelated sequence. In the standard model, h_t is specified as a linear function of past values of ϵ_t^2 and h_t. GARCH models gained popularity among practitioners, and in the econometric literature, because they often give a sensible fit to many types of financial data of moderate sample size. However, empirical work has shown that these models have failed to fit the data very well over a long period of time: parameters change, requiring the practitioners to reestimate them.

This observation leads to Markov-Switching (MS) GARCH models: GARCH models in which the volatility parameters are assumed to depend on a Markov chain. The MS GARCH(p,q) model we consider is defined by the equations

$$\begin{cases} \epsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= \omega(\Delta_t) + \sum_{i=1}^q \alpha_i(\Delta_t)\epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j(\Delta_t)h_{t-j}, \end{cases} \tag{1.1}$$

where the parameters are functions of a finite state ergodic Markov chain, denoted by (Δ_t), which is independent of the centered iid sequence (η_t). Precise definitions and assumptions are given in Section 2. MS GARCH models generate series with a much more flexible dependence structure than in standard GARCH specifications. Since the seminal paper by Hamilton (1989), the use of Markov-Switching Models has become increasingly popular in econometrics. GARCH models of the type (1.1), with Markov-switching coefficients, have already been discussed by, for example, Hamilton and Susmel (1994), Cai (1994), Gray (1996), Dueker (1997), Francq, Roussignol and Zakoïan (2001). A particular case of interest is when (Δ_t) is an independent process (see Wong and Li (2001)).

The purpose of this paper is to investigate the L^2 structure of the MS GARCH processes and their even powers. This is a substantial mathematical problem due to the presence of a latent process in the volatility equation. We find that, under existence conditions, the L^2 structure of (ϵ_t^{2m}), for any nonnegative integer m, is that of a finite ARMA model. This is in particular the case for standard GARCH models. To our knowledge this result was known only for standard GARCH and $m = 1$. We are able to characterize the orders of this ARMA in terms of the MS GARCH parameters. The method for deriving these ARMA representations is novel and may find applications in other contexts.

The L^2-structure results presented here relate to a literature dealing with MS ARMA models. Francq and Zakoïan (2001) establish necessary and sufficient conditions for the existence of strict and second-order stationary solutions to MS ARMA models. Karlsen (1990), Poskitt and Chung (1996), and Zhang and Stine (2001), obtain ARMA representations for autoregressive processes governed by a Markov chain. However, these results do not apply to our framework for the following reason. Although a squared standard GARCH process follows an ARMA model, a squared MS GARCH process does not in general follow a MS ARMA model.

The results in the present paper can be applied to identify the orders and to estimate the parameters of model (1.1). Given the untractability of the likelihood in this framework, it is extremely important to develop alternative methods. In Francq and Zakoïan (2004), motivated by the structural results of the present paper, we investigate an estimation method. The idea is to select the MS model that better matches the second-order behaviour of the observed series and its powers.

The article is organized as follows. Section 2 introduces the main assumptions on the model and derives necessary and sufficient conditions for the existence of moments. Section 3 deals with the autocovariance structure of (ϵ_t^2). The existence of an ARMA representation is proved and its orders are explicitly given. Section 4 extends these results to the powers of (ϵ_t^2). We establish the existence of an ARMA representation for (ϵ_t^{2m}) and characterize the orders and parameters of this representation. An algorithm allowing to explicitly compute the ARMA coefficients is given in an Appendix. To facilitate reading, a list of the main notations is provided after the Appendix.

2. Assumptions and moment conditions

Throughout this paper we make the following standard assumption.

ASSUMPTION A: *the processes* (η_t) *and* (Δ_t) *are independent;* (η_t) *is a sequence of iid random variables with zero mean and unit variance;* (Δ_t) *is a homogeneous, stationary, irreducible, and aperiodic Markov chain with finite state-space* $\mathcal{E} = \{1, 2, ..., d\}$*; the coefficients satisfy the constraints* $\omega(k) > 0$, $\alpha_i(k) \geq 0$, *and* $\beta_j(k) \geq 0$ *for* $k \in \mathcal{E}$, $i \in \{1, \dots, q\}$, $j \in \{1, \dots, p\}$.

The stationary probabilities of (Δ_t) are denoted by $\pi(k) := P(\Delta_1 = k)$ and are positive under Assumption A, the transition probabilities are denoted by $p(k, \ell) := P(\Delta_t = \ell | \Delta_{t-1} = k)$, and the i-step transition probabilities by $p^{(i)}(k, \ell) = P(\Delta_t = \ell | \Delta_{t-i} = k)$ for $k, \ell \in \mathcal{E}$ and $i \geq 1$.

Consider the matrices

$$\mathbb{P}_f^{(i)} = \begin{pmatrix} p^{(i)}(1,1)f(1) & \cdots & p^{(i)}(d,1)f(1) \\ \vdots & & \vdots \\ p^{(i)}(1,d)f(d) & \cdots & p^{(i)}(d,d)f(d) \end{pmatrix}, \quad \text{and} \quad \Pi_f = \begin{pmatrix} \pi(1)f(1) \\ \vdots \\ \pi(d)f(d) \end{pmatrix}, \tag{2.1}$$

for any function $f : \mathcal{E} \to \mathcal{M}_{n \times n'}(\mathbb{R})$, where $\mathcal{M}_{n \times n'}(\mathbb{R})$ denotes the space of real $n \times n'$ matrices, and any positive integers i, n and n'. When $i = 1$ we simply use $\mathbb{P}_f$ to signify $\mathbb{P}_f^{(1)}$, and when $f \equiv 1$ we use $\mathbb{P} := \mathbb{P}_1 = (p(\ell, k))$ to denote the transpose of the transition matrix. Note that when $n = n'$, $\Pi_f = \mathbb{P}_f \Pi_{I_n}$, where I_n is the $n \times n$ identity matrix. Note also that

$$\text{when} \quad n = n' = 1, \qquad \det\{\mathbb{P}_f^{(i)}\} = \{\det \mathbb{P}\}^i \prod_{k=1}^{d} f(k), \tag{2.2}$$

and that, still when $n = n' = 1$

$$\text{when} \quad \prod_{k=1}^{d} f(k) \neq 0, \qquad \operatorname{rank}\{\mathbb{P}_f^{(i)}\} = \operatorname{rank}(\mathbb{P}^i) \quad \text{for any } i \geq 1. \tag{2.3}$$

We now derive necessary and sufficient conditions for the existence of a strictly stationary solution to model (1.1) with a finite $2m$-th order moment, for $m = 1, 2, \dots$. Along the lines of the work done by Bougerol and Picard (1992) for standard GARCH models, we consider the model

$$\underline{z}_t = \underline{b}_t + C_t \underline{z}_{t-1}, \tag{2.4}$$

where $\underline{b}_t = \left(\omega(\Delta_t)\eta_t^2, 0, \ldots, 0, \omega(\Delta_t), 0, \ldots, 0\right)' \in \mathbb{R}^{p+q}$, $\quad \underline{z}_t \in \mathbb{R}^{p+q}$, and

$$C_t = \begin{pmatrix} \boldsymbol{\alpha}_{1:q-1}(\Delta_t)\eta_t^2 & \alpha_q(\Delta_t)\eta_t^2 & \boldsymbol{\beta}_{1:p-1}(\Delta_t)\eta_t^2 & \beta_p(\Delta_t)\eta_t^2 \\ I_{q-1} & 0 & 0 & 0 \\ \boldsymbol{\alpha}_{1:q-1}(\Delta_t) & \alpha_q(\Delta_t) & \boldsymbol{\beta}_{1:p-1}(\Delta_t) & \beta_p(\Delta_t) \\ 0 & 0 & I_{p-1} & 0 \end{pmatrix}$$

where $\boldsymbol{\alpha}_{1:q-1}(\Delta_t) = (\alpha_1(\Delta_t), \ldots, \alpha_{q-1}(\Delta_t))$, $\boldsymbol{\beta}_{1:p-1}(\Delta_t) = (\beta_1(\Delta_t), \ldots, \beta_{p-1}(\Delta_t))$. When $p = 0$ or $q = 0$ obvious modifications have to be done in the definitions of $\underline{b}_t$ and C_t. There is a one-to-one correspondence between the solutions (ϵ_t) of (1.1) and the positive solutions $(\underline{z}_t)$ of (2.4). If (ϵ_t) is solution to (1.1), take $\underline{z}_t = \left(\epsilon_t^2, \ldots, \epsilon_{t-q+1}^2, h_t, \ldots, h_{t-p+1}\right)'$. Conversely, if $(\underline{z}_t)$ is a solution of (2.4) with positive components, take ϵ_t equal to the square root of the first component multiplied by the sign of η_t.

Before stating the main result of this section we need some further notation. Denote by $\rho(A)$ the spectral radius of any square matrix A. Let $\otimes$ denote the Kronecker product and, for any matrix A, let $A^{\otimes m} = A \otimes \cdots \otimes A$. A square positive matrix A is called irreducible if for all i, j there exists a positive integer n such that $A^{(n)}(i,j) > 0$, where $A^{(n)}(i,j)$ denotes the generic element of A^n. A process (X_t) is called nonanticipative if, for all t, X_t is measurable with respect to the σ-field generated by $\{\Delta_u, \eta_u;\ u \leq t\}$. In the statement of the results we will index the expectations by the processes involved, when necessary to avoid ambiguity.

We begin by stating a preliminary result which can be easily proved.

Lemma 2.1. *Let $f : \mathcal{E} \to \mathcal{M}_{n\times n}(\mathbb{R})$ and $g : \mathcal{E} \to \mathcal{M}_{n\times n'}(\mathbb{R})$. Then, for $k > 0$, and $h > k$,*

$$\begin{aligned} E\{f(\Delta_t)f(\Delta_{t-1})\ldots f(\Delta_{t-k+1})g(\Delta_{t-k}) \mid \Delta_{t-h}\} &= \mathbb{I}\mathbb{P}_f^k \mathbb{P}_g^{(h-k)}(e_{\Delta_{t-h}} \otimes I_{n'}), \\ E\{f(\Delta_t)f(\Delta_{t-1})\ldots f(\Delta_{t-k+1})g(\Delta_{t-k})\} &= \mathbb{I}\mathbb{P}_f^k \Pi_g \end{aligned}$$

where $\mathbb{I} = (I_n, \ldots, I_n)$ is a $n \times nd$ matrix and (e_i) denotes the canonical basis of $\mathbb{R}^d$.

We are now ready to state the following theorem.

Theorem 2.1. *Suppose that $E(\eta_t^{2m}) < \infty$ and*

$$\rho\{\mathbb{P}_{C^{(m)}}\} < 1, \tag{2.5}$$

where $C^{(m)}(\cdot) = E_\eta(C_t^{\otimes m}|\Delta_t = \cdot)$ *and* m *is a strictly positive integer. Then (1.1) has a unique nonanticipative and strictly stationary solution* (ϵ_t)*. This solution is also ergodic and satisfies* $E(\epsilon_t^{2m}) < \infty$.

Conversely, if $\mathbb{P}_{C^{(m)}}$ *is irreducible and* $\rho\{\mathbb{P}_{C^{(m)}}\} \geq 1$*, then there is no strictly stationary solution* (ϵ_t) *to model (1.1) such that* $E(\epsilon_t^{2m}) < \infty$.

Proof. First note that there exists a solution (ϵ_t) of (1.1) belonging to L^{2m} if and only if there exists a solution $\underline{z}_t$ of (2.4) belonging to L^m. The proof is divided in three steps.

i) We show that (2.5) is a sufficient condition for the existence of a solution of (2.4) belonging to L^m. The vector

$$\underline{z}_t^* = \underline{b}_t + \sum_{k=1}^{\infty} C_t C_{t-1} \cdots C_{t-k+1} \underline{b}_{t-k} \tag{2.6}$$

is well-defined (its components belong to $\mathbb{R}^+ \cup \{+\infty\}$ with probability one, because the components of b_t and $C_t C_{t-1} \cdots C_{t-k+1} \underline{b}_{t-k}$ are nonnegative). For the series in (2.6) to belong to L^m, it suffices that

$$\underline{z}_{t,k} := C_t C_{t-1} \cdots C_{t-k+1} \underline{b}_{t-k}$$

(with $\underline{z}_{t,0} = \underline{b}_t$) converges to zero in L^m (endowed with any matrix norm) at an exponential rate, as k goes to infinity. For any vector X such that AX is well defined, we have $(AX)^{\otimes m} = A^{\otimes m} X^{\otimes m}$. It readily follows that

$$\underline{z}_{t,k}^{\otimes m} = C_t^{\otimes m} C_{t-1}^{\otimes m} \cdots C_{t-k+1}^{\otimes m} \underline{b}_{t-k}^{\otimes m}. \tag{2.7}$$

Thus we may write, in view of the independence of the matrices and vector $C_t, \cdots C_{t-k+1}$, $\underline{b}_{t-k}$ conditional on Δ_t,

$$\begin{aligned} E\left(\underline{z}_{t,k}^{\otimes m}\right) &= E\{E(C_t^{\otimes m} C_{t-1}^{\otimes m} \cdots C_{t-k+1}^{\otimes m} \underline{b}_{t-k}^{\otimes m} | \Delta_t, \dots, \Delta_{t-k})\} \\ &= E\{C^{(m)}(\Delta_t) \dots C^{(m)}(\Delta_{t-k+1}) \underline{b}^{(m)}(\Delta_{t-k})\} \end{aligned}$$

where $C^{(m)}(\Delta_t) = E(C_t^{\otimes m}|\Delta_t)$ and $\underline{b}^{(m)}(\Delta_t) = E(\underline{b}_t^{\otimes m}|\Delta_t)$. Applying Lemma 2.1 with $n = (p+q)^m$, we obtain

$$E\left(\underline{z}_{t,k}^{\otimes m}\right) = \mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^k \Pi_{\underline{b}^{(m)}}. \tag{2.8}$$

$$E\left(\underline{z}_{t,k}^{\otimes m}\right) = \mathbb{I}\, \mathbb{P}_{C^{(m)}}^k \Pi_{\underline{b}^{(m)}}. \tag{2.9}$$

Consider the matrix norm defined by $\|A\| = \sum_{i,j} |A(i,j)|$, where $A(i,j)$ denotes the generic element of some matrix A. Using the elementary equalities $\|A\|\|B\| = \|A \otimes B\| = \|B \otimes A\|$, it holds that $\|\underline{z}_{t,k}\|^m = \|\underline{z}_{t,k}^{\otimes m}\|$. Hence $E\|\underline{z}_{t,k}\|^m = E\|\underline{z}_{t,k}^{\otimes m}\| = \|E(\underline{z}_{t,k}^{\otimes m})\|$ because $E\|X\| = \|E(X)\|$ for any vector with positive components. In view of (2.9), we thus have

$$E\left\|\underline{z}_{t,k}\right\|^m = \left\|\mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^k \Pi_{\underline{b}^{(m)}}\right\|. \tag{2.10}$$

The matrix norm being multiplicative, we have

$$\|\underline{z}_{t,k}\|_{L^m} = \left\|\mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^k \Pi_{\underline{b}^{(m)}}\right\|^{1/m} \leq \|\mathbb{I}\|^{1/m} \|\{\mathbb{P}_{C^{(m)}}\}^k\|^{1/m} \left\|\Pi_{\underline{b}^{(m)}}\right\|^{1/m}.$$

If (2.5) holds, it follows from the Jordan decomposition that $\|\mathbb{P}_{C^{(m)}}{}^k\|$ converges to zero at an exponential rate as k goes to infinity. Hence, $\sum_{k=0}^K \underline{z}_{t,k}$ converges to $\underline{z}_t^*$ defined by (2.6), both in L^m and almost surely, for fixed t, as $K \to \infty$. For each K, $(\sum_{k=0}^K \underline{z}_{t,k})$ is a measurable function of $(\Delta_t, \eta_t, \ldots, \Delta_{t-K}, \eta_{t-K})$. Hence the limit $(\underline{z}_t^*)$ is a measurable function of the strictly stationary process (Δ_t, η_t). From Billingsley (1995, Theorem 36.4.) it follows that $(\underline{z}_t^*)$ is strictly stationary. It is clear that $\underline{z}_t^* = \underline{b}_t + C_t \underline{z}_{t-1}^*$. Therefore (2.4) admits a nonanticipative strictly stationary solution belonging to L^m. Moreover this solution is ergodic because (Δ_t, η_t) is an ergodic stationary process.

ii) Now assume that another strictly stationary solution of (1.1) exists, or equivalently that there is another strictly stationary solution $(\underline{\tilde{z}}_t)$ of (2.4). Then for any $n \geq 0$,

$$\|\underline{\tilde{z}}_t - \underline{z}_t^*\| = \|C_t(\underline{\tilde{z}}_{t-1} - \underline{z}_{t-1}^*)\| = \ldots \leq \|C_t \ldots C_{t-n}\| \ \ \|\underline{\tilde{z}}_{t-n-1} - \underline{z}_{t-n-1}^*\|.$$

By assumption $\mathbb{P}(\|\underline{\tilde{z}}_t - \underline{z}_t^*\| \neq 0) > 0$. But we know that $\|C_t \ldots C_{t-n}\| \to 0$ a.s. when $n \to \infty$ since the sum in (2.6) converges. It follows that $\mathbb{P}(\lim_{n\to\infty} \|\underline{\tilde{z}}_{t-n-1} - \underline{z}_{t-n-1}^*\| = \infty) > 0$, which entails that either $\limsup_{n\to\infty} \|\underline{\tilde{z}}_{t-n-1}\| = \infty$ or $\limsup_{n\to\infty} \|\underline{z}_{t-n-1}^*\| = \infty$ with a positive probability. This is absurd since the sequences $(\underline{\tilde{z}}_t)_t$ and $(\underline{z}_t^*)_t$ are stationary. The uniqueness of $(\underline{z}_t^*)_t$ is thus established.

iii) We now prove the necessary part. For two matrices A and B of same size, we use the notation $A \succeq B$ (resp. $A \succ B$) to indicate that the (real) components of $A - B$ are nonnegative (resp. strictly positive). In view of (2.4), for any k

$$\underline{z}_t = \underline{z}_{t,0} + \cdots + \underline{z}_{t,k} + C_t \ldots C_{t-k} \underline{z}_{t-k-1} \succeq \sum_{\ell=0}^{k} \underline{z}_{t,\ell}$$

because the matrices C_t and the vectors $\underline{z}_t$ have nonnegative components. Hence $\underline{z}_t \succeq \sum_{\ell=0}^{\infty} \underline{z}_{t,\ell}$ and

$$E(\underline{z}_t^{\otimes m}) \quad \succeq \quad E\left(\sum_{\ell=0}^{\infty} \underline{z}_{t,\ell}\right)^{\otimes m} \succeq \sum_{\ell=0}^{\infty} E(\underline{z}_{t,\ell}^{\otimes m}) = \sum_{\ell=0}^{\infty} \mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^{\ell} \Pi_{\underline{b}^{(m)}}, \tag{2.11}$$

using (2.9). For any irreducible $n \times n$ matrix $A \succeq 0$, we have $(I + A)^{n-1} \succ 0$ (see Lancaster and Tismenetsky (1985, p. 533)). In addition, the elementary relations $\|\mathbb{I}A\| = \|A\|$, $\|A\| \geq \rho(A)$, $\rho(A^n) = \{\rho(A)\}^n$, $n!(I + \cdots + A^n) \succeq (I + A)^n$ and $\|A + B\| = \|A\| + \|B\|$ hold, for any matrices $A \succeq 0$ and $B \succeq 0$ of appropriate dimensions. Moreover, $\|\sum_{\ell=0}^{\infty} A_\ell\| = \sum_{\ell=0}^{\infty} \|A_\ell\|$ for any sequence of matrices $A_\ell \succeq 0$ of the same dimension, by continuity of the norm. Therefore, by (2.11)

$$\begin{aligned}
\|E(\underline{z}_t^{\otimes m})\| &\geq \|\sum_{\ell=0}^{\infty} \mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^{\ell} \Pi_{\underline{b}^{(m)}}\| \\
&= \sum_{\ell=0}^{\infty} \|\mathbb{I}\{\mathbb{P}_{C^{(m)}}\}^{\ell} \Pi_{\underline{b}^{(m)}}\| \\
&= \sum_{\ell=0}^{\infty} \|\{\mathbb{P}_{C^{(m)}}\}^{\ell d(p+q)^m} \{I + \mathbb{P}_{C^{(m)}} + \cdots + \mathbb{P}_{C^{(m)}}^{d(p+q)^m - 1}\} \Pi_{\underline{b}^{(m)}}\| \\
&\geq \frac{1}{\{d(p+q)^m - 1\}!} \sum_{\ell=0}^{\infty} \|\{\mathbb{P}_{C^{(m)}}\}^{\ell d(p+q)^m} \{I + \mathbb{P}_{C^{(m)}}\}^{d(p+q)^m - 1} \Pi_{\underline{b}^{(m)}}\| \\
&\geq \frac{c}{\{d(p+q)^m - 1\}!} \sum_{\ell=0}^{\infty} \{\rho\{\mathbb{P}_{C^{(m)}}\}\}^{\ell d(p+q)^m},
\end{aligned} \tag{2.12}$$

where $c > 0$ is the smallest element of $\{I + \mathbb{P}_{C^{(m)}}\}^{d(p+q)^m - 1} \Pi_{\underline{b}^{(m)}}$. Hence, because (ϵ_t) belongs to L^{2m}, the left-hand side of (2.12) is finite. Therefore (2.5) must hold.

Remark 2.1. Instead of using (2.4), one could consider a recursive system with smaller dimension, $\underline{h}_t = \underline{\omega}_t + A_t \underline{h}_{t-1}$, involving only $\underline{h}_t = (h_t, \ldots, h_{t-r+1})$ with $r = \max\{p, q\}$. The latter system was used by Francq, Roussignol and Zakoïan (2001) to obtain the strict stationarity condition. However, in this representation the matrices A_t are dependent conditional on Δ_t, unless if $p = q = 1$ (A_t is in fact a function of $(\Delta_t, \eta_{t-1}, \ldots, \eta_{t-r})$). For the moment existence, the fact that the matrices C_t are independent conditional on Δ_t is crucial (see the argument after (2.7)). For this reason we used the representation on $\underline{z}_t$ (see however the proof of Corollary 2.2 below for the case $p = q = 1$).

In some special cases the condition for $E(\epsilon_t^{2m}) < \infty$ can be simplified.

Corollary 2.1. *When either i) the functions $\alpha_i(\cdot)$ and $\beta_j(\cdot)$ are constant for all i and j (but not necessarily the intercept $\omega(\cdot)$), or ii) (Δ_t) is an independent process, (2.5) can be replaced by*

$$\rho\{E_{\eta,\Delta}(C_t^{\otimes m})\} < 1. \tag{2.13}$$

Remark 2.2. In case *i)*, considered by Hamilton and Susmel (1994) and Cai (1994), $C_t^{\otimes m}$ depends on η_t only, whereas it depends on η_t and Δ_t in the independent-switching model. Hence, though this condition is apparently the same for the two cases, it does not imply the same restriction on the parameters. Considering the sub-case where the intercept is constant, this corollary applies to the standard GARCH(p,q). In this case, Ling and McAleer (2002) have shown that (2.13) is also necessary.

Proof. Case *i)* is easily proved, noting that C_t does not depend on Δ_t. Consider case *ii)*. Using (2.7) and the independence of the matrices $C_t^{\otimes m}$, we obtain

$$E\|\underline{z}_{t,k}\|^m \quad = \quad \|E\underline{z}_{t,k}^{\otimes m}\| = \left\|\{E(C_t^{\otimes m})\}^k E\{\underline{b}^{(m)}(\Delta_t)\}\right\|.$$

Along the lines of proof of i) in Theorem 2.1, we conclude that (2.13) is a sufficient condition for the existence of a strictly stationary solution with a finite $2m$-th order moment.

Corollary 2.2. *When $p = q = 1$, a necessary and sufficient condition for the existence of $E\epsilon_t^{2m}$ is*

$$\rho\left[\mathbb{P}_{E_\eta\{\alpha_1(\cdot)\eta_t^2+\beta_1(\cdot)\}^m}\right] < 1. \tag{2.14}$$

Moreover, if (Δ_t) is independent this condition takes the form

$$E_{\eta,\Delta}\left\{\alpha_1(\Delta_t)\eta_t^2 + \beta_1(\Delta_t)\right\}^m < 1. \tag{2.15}$$

Remark 2.3. When $\beta_1(\cdot) = 0$, (2.15) reduces to $E\left\{\alpha_1(\Delta_t)\right\}^m E\left(\eta_t^{2m}\right) < 1$. This condition was obtained by Wong and Li (Theorem 3, 2001) for $m = 2$.

Proof. Instead of applying Theorem 2.1, we give a self-contained proof of Corollary 2.2. When $p = q = 1$, the vector equation (2.4) is equivalent to the one-dimensional equation: $h_t = \underline{b}_t + c_t h_{t-1}$, with $\underline{b}_t = \omega(\Delta_t)$ and $c_t = \alpha_1(\Delta_t)\eta_{t-1}^2 + \beta_1(\Delta_t)$. Therefore, the theorem applies with $\mathbb{P}_{C^{(m)}}$ replaced by $\mathbb{P}_{c^{(m)}} := \mathbb{P}_{E\{\alpha_1(\cdot)\eta_t^2+\beta_1(\cdot)\}^m}$. In the proof of the necessary part, the analogous of (2.11) is given by

$$Eh_t^m \geq \sum_{\ell=0}^{\infty} \mathbb{I}\{\mathbb{P}_{c^{(m)}}\}^{\ell}\Pi_{\underline{b}^{(m)}}.$$

Because $\Pi_{\underline{b}^{(m)}} = \Pi_{\{\omega(\cdot)\}^m} \succ 0$, the arguments given to show (2.12) lead to the conclusion that, in the MS GARCH(1,1) case, (2.5) is also necessary.

When (Δ_t) is independent, for any non null function $f : \mathcal{E} \to \mathbb{R}$, the rank of $\mathbb{P}_f$ is 1. Thus, $\rho\{\mathbb{P}_f\} = \mathrm{Tr}\{\mathbb{P}_f\} = Ef(\Delta_t)$, where $\mathrm{Tr}(\cdot)$ denotes the trace of a square matrix. The last part of the proof follows.

Remark 2.4. It can be shown that, in the case $p = q = 1$, (2.14) is equivalent to (2.5). Indeed the (i, j)-th block of the matrix $\mathbb{P}_{C^{(m)}}$ is $p(j, i)E\binom{\eta_t^2}{1}^{\otimes m}(\alpha_1(i),\ \beta_1(i))^{\otimes m}$. It follows that the blocks of $\mathbb{P}_{C^{(m)}}$ are diagonalized by the same eigenvectors matrix, so that we can write $\mathbb{P}_{C^{(m)}} = M\mathbb{P}_{D^{(m)}}M^{-1}$, for some nonsingular matrix M, and where for any k, $D^{(m)}(k)$ is diagonal of rank 1, with non-zero term equal to $E\left\{\alpha_1(k)\eta_t^2 + \beta_1(k)\right\}^m$. The equivalence between the conditions (2.14) and (2.5) easily follows. This gives an alternative proof for the sufficient part of Corollary 2.2.

We now turn to the autocovariance structure of (ϵ_t^{2m}). We will first establish the existence of an ARMA representation for (ϵ_t^2), which can be done in a more straightforward way than for higher powers. The representation will also be more explicit in that particular case.

3. Autocovariance structure of (ϵ_t^2)

That (ϵ_t^2) has the autocorrelation structure of an ARMA$(\max\{p, q\}, p)$ is a well-known feature of standard GARCH(p, q) processes. This was first proved by Milhøj (1985) in the ARCH case, and extended to the GARCH by Bollerslev (1986, 1988). In this section, we show that this feature still holds when the GARCH parameters are allowed to switch in regime.

An approach for deriving ARMA representations of MS models is to find a state-space representation in which the state equation is a Vector AR(1) (see Krolzig (1997)). This method was used for models of the form $X_t = \mu(\Delta_t) + \eta_t$ by Poskitt and Chung (1996). Unfortunately, this method does not seem to apply to our general model (1.1).

We start by describing an alternative approach for characterizing the existence of ARMA models.

3.1. Preliminary result

A useful result for deriving ARMA representations is the following. Recall that a second-order stationary process (X_t) is called purely nondeterministic if $\bigcap_{-\infty}^{\infty} \mathcal{H}_n(X) = \{0\}$, where $\mathcal{H}_n(X)$ is the Hilbert space spanned by the X_t, $t \leq n$ (see Hannan 1970, p. 137).

Lemma 3.1. *Let (X_t) be a real-valued second-order stationary process, with autocovariance function $\gamma(\cdot)$. Suppose that for some $\ell_0 \in \{0, 1, \ldots\}$,*

$$\left\{ \begin{array}{lcll} \gamma(\ell) & = & \delta' W(\ell), & \forall \ell, \\ W(\ell) & = & A W(\ell - 1), & \forall \ell > \ell_0, \end{array} \right. \tag{3.1}$$

where the $W(\ell)$ and δ are column-vectors of size n and A is a $n \times n$ matrix. Let $\mathcal{P}$ denote any real polynomial of degree n such that $\delta' \mathcal{P}(A) = 0$. Write

$$\mathcal{P}(x) = x^{n_0} \mathcal{P}_1(x) \mathcal{P}_2(x), \qquad n_0 \geq 0,$$

where $\mathcal{P}_1$ is a polynomial of degree n_1 whose roots are of modulus 1, and $\mathcal{P}_2$ has no root of modulus 1 or equal to 0. Then (X_t) admits an ARMA$(n - n_0, n + \ell_0 - 1)$ representation, with AR polynomial $x^n \mathcal{P}(x^{-1})$.

Moreover if (X_t) is purely nondeterministic, it admits an ARMA$(n - n_0 - n_1, n + \ell_0 - 1 - n_1)$ representation, with AR polynomial $x^{n - n_0 - n_1} \mathcal{P}_2(x^{-1})$.

Proof. First note that the existence of $\mathcal{P}$ is guaranteed by the fact that the characteristic polynomial of A cancels when applied to A, by the Cayley-Hamilton theorem. However we do not assume that $\mathcal{P}$ is the characteristic polynomial of A. It is not restrictive to assume that $\mathcal{P}(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_{n_0} x^{n_0}$ with $a_n = 1$. Since $W(\ell) = A^{\ell - \ell_0} W(\ell_0), \quad \ell \geq \ell_0$, we find, for all $\ell > n + \ell_0 - 1$,

$$\sum_{i=0}^{n-n_0} a_{n-i} \gamma(\ell - i) \quad = \quad \delta' \sum_{i=0}^{n-n_0} a_{n-i} A^{\ell - \ell_0 - i} W(\ell_0) = \delta' \mathcal{P}(A) A^{\ell - \ell_0 - n} W(\ell_0) = 0.$$

In view of the standard result characterizing the existence of an ARMA representation through the autocovariance function (see e.g. Brockwell and Davis (1991), Proposition

3.2.1 and Remark p.90.), we conclude that (X_t) admits an ARMA$(n-n_0, n+\ell_0-1)$ representation, with AR polynomial

$$\sum_{i=0}^{n-n_0} a_{n-i}x^i = x^n\mathcal{P}(x^{-1}) := x^{n-n_0-n_1}P_1(x)\mathcal{P}_2(x^{-1}),$$

where $P_1(x) = x^{n_1}\mathcal{P}_1(x^{-1})$ is a polynomial with all its roots on the unit circle. By Azencott and Dacunha-Castelle (1984, Theorem 2.3), when (X_t) is purely nondeterministic we have that

(i) the MA polynomial Q can be factorized as $Q(x) = Q_1(x)P_1(x)$, and

(ii) the common roots of modulus 1 in the AR and MA polynomials can be cancelled.

It follows that (X_t) admits an ARMA representation with AR polynomial $x^{n-n_0-n_1}\mathcal{P}_2(x^{-1})$, and MA polynomial $Q_1(x)$. The conclusion follows.

Remark 3.1. It is important to note that the ARMA representation of the theorem is not necessarily the *canonical* ARMA representation, *i.e.* the AR and MA polynomials may have common roots, or roots inside the unit disk. The canonical representation can be straightforwardly deduced by cancelling the common roots and by inverting the roots of modulus smaller than 1 (see e.g. Brockwell and Davis (1991), Remark 1 (a) p. 86 and Proposition 3.5.1).

Remark 3.2. The reciprocal of this result is trivial: if (X_t) is a second-order stationary ARMA(p,q) process, then a representation of the form (3.1) exists, with $W(\ell) = (\gamma(\ell), \dots, \gamma(\ell-p+1))'$, $\delta = (1, 0, \dots, 0)$, and $\ell_0 = q$. To apply the lemma, the difficulty is to find $W(\ell), \delta$ and A for which a relation of the form (3.1) can be exhibited. This will be illustrated below.

3.2. ARMA representation for (ϵ_t^2)

In this section we will show that the square of any second-order stationary solution of model (1.1) admits an ARMA representation. Let $r = \max(p,q)$, for all $i \leq r$

$$\xi_i(\cdot) = \alpha_i(\cdot) + \beta_i(\cdot), \tag{3.2}$$

with by convention $\alpha_i(\cdot) = 0$ (resp. $\beta_i(\cdot) = 0$) when $i > q$ (resp. $i > p$). Recall that $\omega(\cdot)$ plays the role of an intercept in model (1.1). Let the $d(r+1) \times d(r+1)$ matrix

$$A = \begin{pmatrix} \mathbb{P}^{(1)}_{\xi_1} & \cdots & \cdots & \mathbb{P}^{(r)}_{\xi_r} & \mathbb{P}_\omega \\ I_d & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & I_d & 0 & 0 \\ 0 & \cdots & 0 & 0 & \mathbb{P} \end{pmatrix}. \tag{3.3}$$

Denote by L the lag operator ($L^i X_t = X_{t-i}$ and $L^i f(t) = f(t-i)$ for any integers i and t, any process (X_t) and any function f defined on the integers set). Denote by $\gamma_X(.)$ the autocovariance function of any stationary process $X = (X_t)$. Let $n_0(M)$ denote the algebraic multiplicity of the zero eigenvalue of any square matrix M. Let $n_1(M)$ be the sum of the algebraic multiplicities of the eigenvalues of M of modulus 1.

Theorem 3.1. *Suppose that (2.5) holds with $m = 1, 2$, and let (ϵ_t) be the nonanticipative fourth-order stationary solution of model (1.1).*

Then (ϵ_t^2) is an ARMA process, with AR order $N := d(r+1) - n_0(A) - n_1(A)$, where $r = \max(p, q) \geq 1$, and MA order N when $p \geq q$, MA order $N - 1$ when $p < q$.

Remark 3.3. As will be seen in the proof, the AR polynomial can be deduced from the characteristic polynomial of A. First note that $n_1(A) \geq 1$, because 1 is an eigenvalue of $\mathbb{P}$, and that

$$\det(A) = (-1)^{d(r-1)} \det(\mathbb{P}) \det(\mathbb{P}^{(r)}_{\xi_r}) = (-1)^{d(r-1)} \{\det(\mathbb{P})\}^{r+1} \prod_{k=1}^{d} \xi_r(k),$$

in view of (2.2). Moreover it can be seen, using (2.3), that $\operatorname{rank}(A) = (r-1)d + \operatorname{rank}\{\mathbb{P}^{(r)}_{\xi_r}\} + \operatorname{rank}(\mathbb{P})$, *i.e.*

$$n_0(A) = n_0\{\mathbb{P}^{(r)}_{\xi_r}\} + n_0(\mathbb{P}). \tag{3.4}$$

The characteristic polynomial of A, defined by $\det(xI - A)$, is therefore of the form

$$\mathcal{P}_A(x) = x^{n_0(A)} \prod_{i=1}^{n_1(A)} (x - \lambda_i) \mathcal{P}_2(x), \qquad \mathcal{P}_2(0) \neq 0 \tag{3.5}$$

where the λ_i are the eigenvalues of modulus 1 of A, and $\mathcal{P}_2$ is a polynomial. We will see in the proof below that the roots of $\mathcal{P}_2$ are inside the unit circle. It will be also shown that the AR polynomial of the ARMA model is $L^{d(r+1) - n_0\{\mathbb{P}^{(r)}_{\xi_r}\} - n_0(\mathbb{P}) - n_1(A)} \mathcal{P}_2(L^{-1})$.

Remark 3.4. The orders given by the theorem are only maximal ones. However, the only possibility to reduce the AR order occurs when there is a common root of modulus greater than 1 in the AR and MA parts. In any case, the order of the AR part is therefore, after reduction, greater than (or equal to) that of the MA part. This deserves being stressed since it generalizes a standard GARCH property. In other words, if the AR and MA parts have no common root, then the AR order given by the theorem is minimal.

To establish the theorem, the following basic Lemma will be used. We give the proof for the sake of completeness.

Lemma 3.2. *Let, for $i \geq 1$, X_{t-i} be an integrable variable belonging to the $\sigma-$field generated by $\{\epsilon_{t-s}; s \geq i\}$ where (ϵ_t) is a nonanticipative solution of model (1.1). Then*

$$\pi(k)E[X_{t-i}|\Delta_t = k] = \sum_{j=1}^{d} E[X_{t-i}|\Delta_{t-1} = j]p(j,k)\pi(j) = \sum_{j=1}^{d} E[X_{t-i}|\Delta_{t-i} = j]p^{(i)}(j,k)\pi(j).$$

Proof. First, note that for any measurable partition (Ω_j) of Ω, the expectation of any integrable variable can be decomposed as $E(X) = \sum_j E(X|\Omega_j)P(\Omega_j)$ (see e.g. Billingsley, 3.4.1 and 3.4.2). Hence we deduce that

$$E[X_{t-i}|\Delta_t = k] = \sum_{j=1}^{d} E[X_{t-i}|\Delta_t = k, \Delta_{t-1} = j]P[\Delta_{t-1} = j|\Delta_t = k].$$

Now observe that for all j's such that $P[\Delta_t = k, \Delta_{t-1} = j] > 0$, and for all x

$$\begin{aligned} P[X_{t-i} < x|\Delta_t = k, \Delta_{t-1} = j] &= \frac{P[\Delta_t = k|\Delta_{t-1} = j, X_{t-i} < x]P[X_{t-i} < x, \Delta_{t-1} = j]}{P[\Delta_t = k, \Delta_{t-1} = j]} \\ &= \frac{P[\Delta_t = k|\Delta_{t-1} = j]P[X_{t-i} < x|\Delta_{t-1} = j]P[\Delta_{t-1} = j]}{P[\Delta_t = k, \Delta_{t-1} = j]} \\ &= P[X_{t-i} < x|\Delta_{t-1} = j]. \end{aligned}$$

Hence $E[X_{t-i}|\Delta_t = k, \Delta_{t-1} = j] = E[X_{t-i}|\Delta_{t-1} = j]$. In view of $P[\Delta_{t-1} = j|\Delta_t = k]\pi(k) = P[\Delta_t = k|\Delta_{t-1} = j]\pi(j)$, the first equality follows. The second equality can be shown similarly.

Proof of Theorem 3.1. The proof is split into three steps.

i) We proceed by evaluating the expectation of $\epsilon_t^2\epsilon_{t-\ell}^2$ conditional to the regime of

Δ_t. For $\ell > 0$ and $k \in \mathcal{E}$ we have

$$\begin{aligned} & E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = k] = E[h_t \epsilon_{t-\ell}^2 | \Delta_t = k] \\ = \quad & \omega(k) E[\epsilon_{t-\ell}^2 | \Delta_t = k] + \sum_{i=1}^{q} \alpha_i(k) E[\epsilon_{t-i}^2 \epsilon_{t-\ell}^2 | \Delta_t = k] + \sum_{j=1}^{p} \beta_j(k) E[h_{t-j} \epsilon_{t-\ell}^2 | \Delta_t = k]. \end{aligned}$$

Moreover, for $\ell > j$ we have $E[\epsilon_{t-j}^2 \epsilon_{t-\ell}^2 | \Delta_t = k] = E[h_{t-j} \epsilon_{t-\ell}^2 | \Delta_t = k]$, so that

$$E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = k] = \omega(k) E[\epsilon_{t-\ell}^2 | \Delta_t = k] + \sum_{i=1}^{r} \xi_i(k) E[\epsilon_{t-i}^2 \epsilon_{t-\ell}^2 | \Delta_t = k], \qquad \ell > p.$$

In view Lemma 3.2, we have, for $\ell > p$ and $\ell \geq q$,

$$\pi(k) E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = k] = \pi(k) \omega(k) E[\epsilon_{t-\ell}^2 | \Delta_t = k] + \sum_{i=1}^{r} \sum_{j=1}^{d} E[\epsilon_{t-i}^2 \epsilon_{t-\ell}^2 | \Delta_{t-i} = j] \xi_i(k) p^{(i)}(j,k) \pi(j)$$

which, by the notations in (2.1), can be written in vector form as

$$\begin{aligned} \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} \quad = \quad & \Pi_{\omega(\cdot) E[\epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} \\ & + \sum_{i=1}^{r} \mathbb{P}_{\xi_i}^{(i)} \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell+i}^2 | \Delta_t = \cdot\,]}, \qquad \ell \geq \max\{p+1, q\}. \end{aligned} \tag{3.6}$$

Further we have, using Lemma 3.2

$$\pi(k) E[\epsilon_{t-\ell}^2 | \Delta_t = k] = \sum_{j=1}^{d} E[\epsilon_{t-\ell}^2 | \Delta_{t-1} = j] p(j,k) \pi(j), \qquad \ell > 0,$$

or equivalently

$$\Pi_{E[\epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} = \mathbb{P} \Pi_{E[\epsilon_{t-\ell+1}^2 | \Delta_t = \cdot\,]}, \qquad \ell > 0. \tag{3.7}$$

Similarly

$$\Pi_{\omega(\cdot) E[\epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} = \mathbb{P}_\omega \Pi_{E[\epsilon_{t-\ell+1}^2 | \Delta_t = \cdot\,]}, \qquad \ell > 0. \tag{3.8}$$

This leads us to introduce the $d(r+1) \times 1$ vector

$$W(\ell) := \begin{pmatrix} \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} \\ \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell+1}^2 | \Delta_t = \cdot\,]} \\ \vdots \\ \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell+r-1}^2 | \Delta_t = \cdot\,]} \\ \Pi_{E[\epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]} \end{pmatrix}, \qquad \ell = 0, \pm 1, \ldots.$$

From (3.6), (3.7) and (3.8) we have,

$$W(\ell) = AW(\ell - 1), \quad \forall \ell > \max\{p+1, q\} - 1. \tag{3.9}$$

It should be noted that, because the components of $W(\ell)$ are positive, and bounded when ℓ increases by stationarity of (ϵ_t^2), (3.9) implies that $\rho(A) \leq 1$. The autocovariance function of (ϵ_t^2) is given by

$$\gamma_{\epsilon^2}(\ell) = \sum_{k=1}^{d} E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = k]\pi(k) - \gamma_\epsilon^2(0)$$

where $\gamma_\epsilon(0) = E(\epsilon_t^2)$. In this equality, the right-hand terms can be expressed as

$$\gamma_\epsilon(0) = \underline{1}_d' \Pi_{E[\epsilon_{t-\ell-r+1}^2 | \Delta_t = \cdot\,]} \quad \text{and} \quad \sum_{k=1}^{d} E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = k]\pi(k) = \underline{1}_d' \Pi_{E[\epsilon_t^2 \epsilon_{t-\ell}^2 | \Delta_t = \cdot\,]},$$

where $\underline{1}_d = (1, \dots, 1)'$ denotes the d element sum vector. We therefore find that, in view of (3.9)

$$\begin{cases} \gamma_{\epsilon^2}(\ell) & = & \delta' W(\ell + r - 1), \quad \forall \ell, \\ W(\ell + r - 1) & = & AW(\ell + r - 2), \quad \forall \ell > \max\{p+1, q\} - r, \end{cases} \tag{3.10}$$

with $\delta' = (0, \dots, 0, \underline{1}_d', -\underline{1}_d' \gamma_\epsilon(0))$. By the first part of Lemma 3.1 with $n_0 = n_0(A)$, $n = d(r+1)$ and $\mathcal{P} = \mathcal{P}_A$, we find that (ϵ_t^2) admits an ARMA representation, with orders $(d(r+1) - n_0(A), d(r+1))$ when $p \geq q$, and $(d(r+1) - n_0(A), d(r+1) - 1)$ when $p < q$.

ii) To reduce the ARMA orders, in view of the second part of Lemma 3.1, we now show that $X_t := \epsilon_t^2 - E(\epsilon_t^2)$ is purely nondeterministic. It is sufficient to show that

$$E|X_t - P_{\mathcal{H}_{t-h}(X)} X_t|^2 \to EX_t^2, \quad \text{when } h \to \infty,$$

where $P_{\mathcal{H}_n(X)}$ denotes the projection mapping onto $\mathcal{H}_n(X)$. This amounts to showing that $\operatorname{Var}\left(P_{\mathcal{H}_{t-h}(X)} X_t\right)$ tends to zero as $h \to \infty$. From

$$\begin{aligned} \operatorname{Var}(X_t) &= \operatorname{Var}\left(P_{\mathcal{H}_{t-h}(X)} X_t\right) + \operatorname{Var}\left(X_t - P_{\mathcal{H}_{t-h}(X)} X_t\right) \\ &= \operatorname{Var}\left\{E\left(X_t | X_{t-s}, s \geq h\right)\right\} + \operatorname{Var}\left\{X_t - E(X_t | X_{t-s}, s \geq h)\right\} \end{aligned}$$

we deduce that $\operatorname{Var}\{E(X_t | X_{t-s}, s \geq h)\} \geq \operatorname{Var}\left(P_{\mathcal{H}_{t-h}(X)} X_t\right)$. By the same argument we have $\operatorname{Var}\left\{E\left(X_t | \Delta_{t-s}, \eta_{t-s}, s \geq h\right)\right\} \geq \operatorname{Var}\left\{E\left(X_t | X_{t-s}, s \geq h\right)\right\}$. It is therefore sufficient to show that $\|$ Var $\{E(\underline{z}_t^*|\Delta_{t-s},\eta_{t-s},s\geq h)\}\|$ tends to zero as $h\to\infty$, where $\underline{z}_t^*$ is defined in (2.6). Using Lemma 2.1, we obtain

$$\begin{aligned} E(\underline{z}_t^*|\Delta_{t-s},\eta_{t-s},s\geq h) &= \mathbb{I}\,\mathbb{P}^{(h)}_{\underline{b}^{(1)}}e_{\Delta_{t-h}}+\sum_{k=1}^{h-1}\mathbb{I}\,\mathbb{P}^k_{C^{(1)}}\mathbb{P}^{(h-k)}_{\underline{b}^{(1)}}e_{\Delta_{t-h}} \\ &\quad +\sum_{k=h}^{\infty}\mathbb{I}\,\mathbb{P}^h_{C^{(1)}}\left\{e_{\Delta_{t-h}}\otimes I_{p+q}\right\}C_{t-h}\cdots C_{t-k+1}\underline{b}_{t-k}, \end{aligned}$$

where $\mathbb{I}=\underline{1}_d\otimes I_{p+q}$. Using the Minkovski inequality, $\left\|\mathbb{P}^h_{C^{(1)}}\right\|\to 0$ and

$$\mathbb{P}^{(h)}_{\underline{b}^{(1)}}e_i\to\Pi_{\underline{b}^{(1)}}\text{ as }h\to\infty,\qquad\text{for all }i=1,\dots,d,$$

it is easy to show that $\|E(\underline{z}_t^*|\Delta_{t-s},\eta_{t-s},s\geq h)-E(\underline{z}_t^*)\|_2\to 0$, as $h\to\infty$. We conclude that (X_t) is purely nondeterministic. Applying the second part of Lemma 3.1, we find that ϵ_t^2 has an ARMA representation, with orders $(N,d(r+1)-n_1(A))$ when $p\geq q$, and $(N,d(r+1)-n_1(A)-1)$ when $p<q$. The AR order is the announced result. When $n_0(A)=0$ we get the desired MA order.

iii) The MA order can be further reduced when $n_0(A)\neq 0$. We will show that (3.9) and (3.10) hold, where $W(\cdot)$, δ and A are replaced by matrices of smaller sizes, $\tilde{W}(\cdot)$, $\tilde{\delta}$ and $\tilde{A}$, with $\tilde{A}$ non-singular. There exist a full rank matrix $\mathbb{P}^*$, obtained by cancelling $n_0(\mathbb{P})$ redundant columns of $\mathbb{P}$, and a full rank $\{d-n_0(\mathbb{P})\}\times d$ matrix $\Lambda_{\mathbb{P}}$ such that $\mathbb{P}=\mathbb{P}^*\Lambda_{\mathbb{P}}$. Because $\omega(\cdot)>0$ we also have $\mathbb{P}_\omega=\mathbb{P}^*_\omega\Lambda_{\mathbb{P}}$ for some full rank matrix $\mathbb{P}^*_\omega$. We similarly introduce a full rank matrix $\mathbb{P}^{*(r)}$ and a full rank $\{d-n_0(\mathbb{P}^{(r)}_{\xi_r})\}\times d$ matrix $\Lambda_{\mathbb{P}^{(r)}}$ such that $\mathbb{P}^{(r)}_{\xi_r}=\mathbb{P}^{*(r)}\Lambda_{\mathbb{P}^{(r)}}$. Now define the block-diagonal $\{d(r+1)-n_0(A)\}\times d(r+1)$ matrix $\Lambda_A=\operatorname{diag}\left(I_{d(r-1)},\Lambda_{\mathbb{P}^{(r)}},\Lambda_{\mathbb{P}}\right)$, and let the $d(r+1)\times\{d(r+1)-n_0(A)\}$ matrix A^* such that $A=A^*\Lambda_A$. Let $\tilde{A}=\Lambda_A A^*$. Remark that A^*,Λ_A and $\tilde{A}$ are full-rank matrices, and that the square matrices A and $\tilde{A}$ have the same rank (when $n_0(A)>0$, A is not full-rank of course). We have by (3.10)

$$\tilde{W}(\ell+r-1):=\Lambda_A W(\ell+r-1)=\tilde{A}\tilde{W}(\ell+r-2),\quad\forall\ell>\max\{p+1,q\}-r.$$

Note that $\underline{1}_d'=\underline{1}_d'\mathbb{P}=\underline{1}_d'\mathbb{P}^*\Lambda_{\mathbb{P}}=\underline{1}'_{d-n_0(\mathbb{P})}\Lambda_{\mathbb{P}}=\underline{1}'_{d-n_0(\mathbb{P}^{(r)}_{\xi_r})}\Lambda_{\mathbb{P}^{(r)}}$. Thus $\delta'=\tilde{\delta}'\Lambda_A$ with $\tilde{\delta}'=(0,\dots,0,\underline{1}'_{d-n_0(\mathbb{P}^{(r)}_{\xi_r})},-\underline{1}'_{d-n_0(\mathbb{P})}\gamma_\epsilon(0))$. Hence $\gamma_{\epsilon^2}(\ell)=\tilde{\delta}'\tilde{W}(\ell+r-1)$. Applying Lemma 3.1, with $n_0=0$ (because $\tilde{A}$ has full rank), $n=d(r+1)-n_0(A)$ and $\mathcal{P}=\mathcal{P}_{\tilde{A}}$ (the characteristic polynomial of $\tilde{A}$), we find that (ϵ_t^2) admits an ARMA representation, with orders $(d(r+1)-n_0(A),d(r+1)-n_0(A))$ when $p\geq q$, and $(d(r+1)-n_0(A),d(r+$

$1) - n_0(A) - 1)$ when $p < q$. It is easy to show that A and $\tilde{A}$ have the same non-zero eigenvalues. By cancelling the roots of modulus 1, as in step *ii)*, we obtain the result of the theorem and the conclusion of Remark 3.3.

EXAMPLE 3.1: To illustrate the theorem consider the case $d = 2$ and $p = q = 1$. The matrix A, in which the subscripts have been omitted, is given in this case by

$$A = \begin{pmatrix} p(1,1)\{\alpha(1)+\beta(1)\} & p(2,1)\{\alpha(1)+\beta(1)\} & p(1,1)\omega(1) & p(2,1)\omega(1) \\ p(1,2)\{\alpha(2)+\beta(2)\} & p(2,2)\{\alpha(2)+\beta(2)\} & p(1,2)\omega(2) & p(2,2)\omega(2) \\ 0 & 0 & p(1,1) & p(2,1) \\ 0 & 0 & p(1,2) & p(2,2) \end{pmatrix}$$

Now take for instance $p(1,1) = p(2,1) = 0.2$, $\alpha(1) = \beta(1) = 0.05$, $\alpha(2) = \beta(2) = 0.1$. Then we have $\mathcal{P}_A(x) = x^2(x-1)(x-0.18)$, $n_0(A) = 2$, $n_1(A) = 1$. Applying Theorem 3.1 we may conclude that (ϵ_t^2) is solution of an ARMA(1,1) equation. By Remark 3.3 we know that the AR polynomial is of the form $1 - 0.18L$. By numerical computation of $E\epsilon_t^2$, Varϵ_t^2 and Cov$(\epsilon_t^2, \epsilon_{t-1}^2)$ (as described in the Appendix), we obtain the ARMA(1,1) representation:

$$\epsilon_t^2 - 2.195 = 0.18(\epsilon_{t-1}^2 - 2.195) + u_t - 0.088u_{t-1}, \quad \mathrm{var}(u_t) = 10.556$$

where u_t is a white noise.

4. Autocovariance structure of (ϵ_t^{2m})

In this section we extend the results of the previous section to powers of (ϵ_t^2). For statistical purposes, such as identification, considering only the ARMA representation for ϵ_t^2 will not suffice. Indeed, two different models of the form (1.1) may lead to the same ARMA representation for the squares, but, in general, higher powers will have different representations.

We will establish that for any positive integer m, the process (ϵ_t^{2m}) is, under mild assumptions, an ARMA. This result is much harder to prove than for (ϵ_t^2). Before dealing with the general case, we proceed with the case $p = q = 1$, which can be handled in a simpler way and provides sharper results as far as the ARMA orders are concerned. The $2m-$th order moment of the iid process is denoted by $\mu_{2m} = E(\eta_t^{2m})$.

4.1. The case $p = q = 1$

Let

$$\xi_{m_1,m_2}(\cdot) = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \{\omega(\cdot)\}^{m_1-m_2} E_\eta[\alpha(\cdot)\eta_{t-1}^2 + \beta(\cdot)]^{m_2} \frac{\mu_{2m_1}}{\mu_{2m_2}}. \tag{4.1}$$

Note that $\xi_{1,1}(\cdot) = \xi(\cdot)$, as defined in (3.2). This notation is introduced to establish formulas of the form $E[\epsilon_t^{2m_1}|\Delta_t] = \sum_{m_2=0}^{m_1} \xi_{m_1,m_2}(\Delta_t) E[\epsilon_{t-1}^{2m_2}|\Delta_t]$, for $m_1 \leq 2m$. Let the $d(m+1) \times d(m+1)$ matrix

$$A_m = \begin{pmatrix} \mathbb{P}_{\xi_{m,m}} & \mathbb{P}_{\xi_{m,m-1}} & \cdots & \cdots & \mathbb{P}_{\xi_{m,0}} \\ 0 & \mathbb{P}_{\xi_{m-1,m-1}} & \cdots & \cdots & \mathbb{P}_{\xi_{m-1,0}} \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & \mathbb{P}_{\xi_{1,1}} & \mathbb{P}_{\xi_{1,0}} \\ 0 & \cdots & 0 & 0 & \mathbb{P} \end{pmatrix}.$$

Theorem 4.1. *Let $p = q = 1$ and suppose $\alpha(\cdot) + \beta(\cdot) > 0$. Suppose that $\rho\left[\mathbb{P}_{\xi_{2m,2m}}\right] < 1$ and let (ϵ_t) be the nonanticipative $4m$th-order stationary solution of model (1.1).*

Then (ϵ_t^{2m}) is an ARMA(N_1, N_1) process, with $N_1 := (m+1)\{d - n_0(\mathbb{P})\} - n_1(A_m)$.

Remark 4.1. Recall that by Corollary 2.2, $\rho\left[\mathbb{P}_{\xi_{2m,2m}}\right] < 1$ is a necessary and sufficient condition for the existence of $E\epsilon_t^{4m}$. In view of (2.2) we have

$$\det(A_m) = \{\det(\mathbb{P})\}^{m+1} \prod_{m_1=1}^{m} \prod_{k=1}^{d} \xi_{m_1,m_1}(k), \quad \text{with} \quad \xi_{m_1,m_1}(\cdot) = E_\eta\{\alpha(\cdot)\eta_{t-1}^2 + \beta(\cdot)\}^{m_1}. \tag{4.2}$$

Notice that, since $\xi_{m_1,m_1}(\cdot) > 0$, we have $n_0(A_m) = (m+1)n_0(\mathbb{P})$. The characteristic polynomial of A_m is therefore of the form

$$\mathcal{P}_{A_m}(x) = x^{(m+1)n_0(\mathbb{P})} \prod_{i=1}^{n_1(A_m)} (x - \lambda_i)\mathcal{P}_2(x), \qquad \mathcal{P}_2(0) \neq 0 \tag{4.3}$$

where the λ_i are the eigenvalues of modulus 1 of A_m, and $\mathcal{P}_2$ is a polynomial whose roots are inside the unit circle. It will be shown below that the AR polynomial of the ARMA model is $L^{(m+1)\{d-n_0(\mathbb{P})\}-n_1(A_m)}\mathcal{P}_2(L^{-1})$.

Remark 4.2. In the special case when $d = 1$ we have $\alpha(\cdot) = \alpha$, $\beta(\cdot) = \beta$, $\mathbb{P} = 1$ and $n_1(A_m) = 1$. Therefore we find that for the standard GARCH$(1,1)$ model, the process (ϵ_t^{2m}) is an ARMA(m, m). This result is well-known when $m = 1$ and, to our knowledge, had not been proved for higher exponents m. There is an important difference, however, between the case $m = 1$ and the general case. When $m = 1$, the ARMA representation is a strong one in the sense that the noise is a martingale difference. This will generally not be the case for higher exponents m.

Remark 4.3. The AR polynomial of the ARMA representation for (ϵ_t^{2m}) is the product of the AR polynomial of the ARMA representation for $(\epsilon_t^{2(m-1)})$, and a polynomial obtained from the matrix $\mathbb{P}_{\xi_{m,m}}$. In particular, the AR part of the representation does not depend on $\omega(\cdot)$, because $\xi_{m_1,m_1}(k) = E\{\alpha(k)\eta_{t-1}^2 + \beta(k)\}^{m_1}$, and hence the block-diagonal terms of A_m, are independent of $\omega(\cdot)$. Note also that Remark 3.4 remains valid.

Before proving the theorem, we will establish a basic result we will use to investigate the autocovariance structure of (ϵ_t^{2m}).

Lemma 4.1. *Under the assumptions of Theorem 4.1 and Lemma 3.2, if $j < i$, we have for any $m_1 \in \{0, 1, \dots m\}$*

$$E[\epsilon_{t-j}^{2m_1} X_{t-i} | \Delta_t = k] = \mu_{2m_1} E[h_{t-j}^{m_1} X_{t-i} | \Delta_t = k]$$

and, for nonnegative integers m_1 and m_2 such that $m_1 + m_2 \leq m$

$$E[\epsilon_{t-j}^{2m_1} h_{t-j}^{m_2} X_{t-i} | \Delta_t = k] = \frac{\mu_{2m_1}}{\mu_{2(m_1+m_2)}} E[\epsilon_{t-j}^{2(m_1+m_2)} X_{t-i} | \Delta_t = k].$$

Proof. The first statement is a trivial consequence of the equality $\epsilon_{t-j} = \sqrt{h_{t-j}}\eta_{t-j}$ and the independence between η_{t-j} and all other variables involved in the conditional expectation. To prove the second statement, it suffices to write $E[\epsilon_{t-j}^{2m_1} h_{t-j}^{m_2} X_{t-i} | \Delta_t = k] = E[\eta_{t-j}^{2m_1} h_{t-j}^{m_1+m_2} X_{t-i} | \Delta_t = k] = \mu_{2m_1} E[h_{t-j}^{m_1+m_2} X_{t-i} | \Delta_t = k]$ and to use the first part of the lemma.

Proof of Theorem 4.1. Observe that for $\ell > 0$, and for all $m_1 \in \{0, 1, \dots, m\}$, $k \in \mathcal{E}$,

$$\begin{aligned} & E[\epsilon_t^{2m_1} \epsilon_{t-\ell}^{2m} | \Delta_t = k] \\ = \;& \mu_{2m_1} \sum_{m_2=0}^{m_1} \binom{m_1}{m_2} E\left[\{\omega(\Delta_t)\}^{m_1-m_2} \{\alpha(\Delta_t)\epsilon_{t-1}^2 + \beta(\Delta_t)h_{t-1}\}^{m_2} \epsilon_{t-\ell}^{2m} | \Delta_t = k\right]. \end{aligned} \quad (4.4)$$

Moreover, for $\ell > 1$

$$
\begin{aligned}
& E[\{\alpha(\Delta_t)\epsilon_{t-1}^2 + \beta(\Delta_t)h_{t-1}\}^{m_2}\epsilon_{t-\ell}^{2m}|\Delta_t = k] \\
= \;& E[\{\alpha(\Delta_t)\eta_{t-1}^2 + \beta(\Delta_t)\}^{m_2}h_{t-1}^{m_2}\epsilon_{t-\ell}^{2m}|\Delta_t = k] \\
= \;& E\{\alpha(k)\eta_{t-1}^2 + \beta(k)\}^{m_2}E[h_{t-1}^{m_2}\epsilon_{t-\ell}^{2m}|\Delta_t = k] \\
= \;& E\{\alpha(k)\eta_{t-1}^2 + \beta(k)\}^{m_2}\frac{1}{\mu_{2m_2}}E[\epsilon_{t-1}^{2m_2}\epsilon_{t-\ell}^{2m}|\Delta_t = k]
\end{aligned}
$$

by Lemma 4.1. Hence, in view of (4.4) and Lemma 3.2, we have for $\ell > 1$

$$
\begin{aligned}
& \pi(k)E[\epsilon_t^{2m_1}\epsilon_{t-\ell}^{2m}|\Delta_t = k] = \pi(k)\sum_{m_2=0}^{m_1}\xi_{m_1,m_2}(k)E[\epsilon_{t-1}^{2m_2}\epsilon_{t-\ell}^{2m}|\Delta_t = k] \\
= \;& \sum_{m_2=0}^{m_1}\sum_{j=1}^{d}\xi_{m_1,m_2}(k)E[\epsilon_{t-1}^{2m_2}\epsilon_{t-\ell}^{2m}|\Delta_{t-1} = j]p(j,k)\pi(j) \qquad (4.5)
\end{aligned}
$$

which can be written in vector form as

$$
\Pi_{E[\epsilon_t^{2m_1}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} = \sum_{m_2=0}^{m_1}\mathbb{P}_{\xi_{m_1,m_2}}\Pi_{E[\epsilon_t^{2m_2}\epsilon_{t-\ell+1}^{2m}|\Delta_t=\cdot\,]}, \quad \text{for all} \quad \ell > 1. \qquad (4.6)
$$

Because (4.6) holds for all $m_1 \in \{0, 1, \dots m\}$, we have for $\ell > 1$,

$$
W_m(\ell) = A_m W_m(\ell - 1)
$$

where

$$
W_m(\ell) = \begin{pmatrix} \Pi_{E[\epsilon_t^{2m}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \\ \Pi_{E[\epsilon_t^{2(m-1)}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \\ \vdots \\ \vdots \\ \Pi_{E[\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \end{pmatrix}.
$$

Moreover,

$$
\gamma_{\epsilon^{2m}}(\ell) = \underline{1}_d'\Pi_{E[\epsilon_t^{2m}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} - E(\epsilon_t^{2m})\underline{1}_d'\Pi_{E[\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} = \nu' W_m(\ell), \qquad (4.8)
$$

with $\nu' = (\underline{1}_d', 0, \dots, 0, -E(\epsilon_t^{2m})\underline{1}_d')$.

Hence, by Lemma 3.1, with $n_0 = n_0(A_m)$, $\ell_0 = 1$, $n = d(m+1)$ and $\mathcal{P} = \mathcal{P}_{A_m}$, we find that (ϵ_t^{2m}) admits an ARMA representation with orders $(d(m+1) - n_0(A_m), d(m+1))$.

With the notations of the proof of Theorem 3.1, we introduce the full-rank block-diagonal $(m+1)\{d-n_0(\mathbb{P})\} \times d(m+1)$ matrix $\Lambda_{A_m} = \mathrm{diag}\,(\Lambda_{\mathbb{P}}, \dots, \Lambda_{\mathbb{P}})$, and the matrices A_m^* and $\tilde{A}_m$ such that $A_m = A_m^* \Lambda_{A_m}$, and $\tilde{A}_m = \Lambda_{A_m} A_m^*$ is a non singular $(m+1)\{d-n_0(\mathbb{P})\} \times (m+1)\{d-n_0(\mathbb{P})\}$ matrix . We have

$$\tilde{W}_m(\ell) := \Lambda_{A_m} W_m(\ell) = \tilde{A}_m \tilde{W}_m(\ell-1), \quad \forall \ell > 1.$$

Thus $\nu' = \tilde{\nu}' \Lambda_{A_m}$ with $\tilde{\nu}' = (\underline{1}'_{d-n_0(\mathbb{P})}, 0, \dots, 0, -E(\epsilon_t^{2m})\underline{1}'_{d-n_0(\mathbb{P})})$. Hence $\gamma_{\epsilon^{2m}}(\ell) = \tilde{\nu}' \tilde{W}_m(\ell)$. Applying the first part of Lemma 3.1, with $n = (m+1)\{d-n_0(\mathbb{P})\}$, $n_0(\tilde{A}_m) = 0$, $\ell_0 = 1$, we may conclude that ϵ_t^{2m} is an ARMA$((m+1)\{d-n_0(\mathbb{P})\}, (m+1)\{d-n_0(\mathbb{P})\})$ process.

An argument similar to that used in Step *ii)* of the proof of Theorem 3.1 shows that $\epsilon_t^{2m} - E(\epsilon_t^{2m})$ is purely nondeterministic. The proof is completed by applying the second part of Lemma 3.1.

Example 4.1 (Example 3.1 continued): By Remark 4.3, the AR polynomial of the representation for ϵ_t^4 is a multiple of the AR polynomial obtained for ϵ_t^2. It thus suffices to compute the eigenvalues of $\mathbb{P}_{\xi_{2,2}}$. Assuming that η_t is gaussian, we have $\xi_{2,2}(1) = 0.015$ and $\xi_{2,2}(2) = 0.06$. The non-zero eigenvalue of $\mathbb{P}_{\xi_{2,2}}$ is 0.051. We therefore obtain an ARMA(2,2) representation for ϵ_t^4 with AR polynomial $(1-0.18L)(1-0.051L)$.

Example 4.2: To numerically illustrate the remark at the beginning of this section, we consider the two following GARCH(1,1) specifications, with η_t standard gaussian:

(I) Standard GARCH with $\omega = 1$, $\alpha_1 = 0.1728$, $\beta_1 = 0.4272$.

(II) Independent-switching with $\omega(1) = \omega(2) = 1$, $\alpha_1(1) = 0.1415$, $\beta_1(1) = 1.0585$, $\alpha_1(2) = \beta_1(2) = 0$, $p(1,1) = p(1,2) = p(2,1) = p(2,2) = 0.5$.

The two models lead to the same ARMA(1,1) representation for ϵ_t^2 but the ARMA(2,2) representations for ϵ_t^4 are different:

$$\begin{array}{ll} (I) & (1-0.6L)\epsilon_t^2 = 1 + (1-0.427L)u_t^{(1)}, \\ & (1-0.6L)(1-0.4197L)\epsilon_t^4 = 4.8 + (1-0.537L)(1-0.3058L)u_t^{(2)}, \\ (II) & (1-0.6L)\epsilon_t^2 = 1 + (1-0.427L)v_t^{(1)}, \\ & (1-0.6L)(1-0.74L)\epsilon_t^4 = 4.8 + (1-0.687L)(1-0.5349L)v_t^{(2)}, \end{array}$$

where the $u_t^{(i)}$ and $v_t^{(i)}$ are white noises. In these ARMA models, the AR coefficients were obtained by computing the eigenvalues of matrices of the form A_2 and A_4. The intercepts and MA coefficients were derived by the method presented in the Appendix.

4.2. The general case

We start this section by briefly describing the scheme of proof and by introducing some new notation. For ease of presentation we first assume that $p > 1$ and $q > 1$ but the theorem of this section will be stated for $p \geq 1$, $q \geq 1$. Appropriate extensions of the notations will be given in Remark 4.6 below.

Proceeding as in Section 4.1, we will compute moments of the process (ϵ_t) conditional on the Markov chain. However, instead of moments of variables of the form $\epsilon_t^{2m_1}\epsilon_{t-\ell}^{2m}$ as in the case $p = q = 1$, it will be necessary to consider variables of the form $\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}$ where $\Upsilon_t(\underline{m}')$ is a product of for

$$\begin{aligned} \underline{m}' &= (m_0, \dots, m_{q-1}, n_1 \dots, n_{p-1}) \in \mathbb{N}^{p+q-1}, \\ \Upsilon_t(\underline{m}') &= \epsilon_t^{2m_0} \dots \epsilon_{t-q+1}^{2m_{q-1}} h_{t-1}^{n_1} \dots h_{t-p+1}^{n_{p-1}}. \end{aligned} \tag{4.9}$$

To compute the expectation of $\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}$ conditional on $\Delta_t = k$, we will replace ϵ_t^2 by $h_t\eta_t^2$ and apply the multinomial formula to $h_t^{m_0}$, with h_t expressed in function of the ϵ_{t-i}^2 and h_{t-j}. It will therefore be convenient to introduce the real numbers

$$\xi_{\underline{m}',\underline{l}}(k) = \frac{m_0!}{\Pi_{j=0}^{p+q} l_j!}\omega(k)^{l_0}\alpha_1(k)^{l_1}\dots\alpha_q(k)^{l_q}\beta_1(k)^{l_{q+1}}\dots\beta_p(k)^{l_{p+q}}\frac{\mu_{2m_0}\mu_{2(l_1+m_1)}}{\mu_{2(l_1+m_1+l_{q+1}+n_1)}},$$

defined for any $k \in \mathcal{E}$, $\underline{l} = (l_0, \dots, l_{p+q}) \in \mathbb{N}^{p+q+1}$ and $m_0, m_1 \in \mathbb{N}$. Thus we will obtain a formula relying the expectation of $\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}$ conditional on $\Delta_t = k$, and the expectations, conditional on $\Delta_{t-1} = j$, of terms of the form

$$\epsilon_{t-1}^{2(l_1+m_1+l_{q+1}+n_1)}\dots\epsilon_{t-q}^{2(l_q+m_q)}h_{t-2}^{l_{q+2}+n_2}\dots h_{t-p}^{l_{p+q}+n_p}\epsilon_{t-\ell}^{2m} := \Upsilon_{t-1}(\underline{m}'\star\underline{l})\epsilon_{t-\ell}^{2m}, \tag{4.10}$$

with for notational convenience $m_q = n_p = 0$. This formula involves an operator, $\star$, from $\mathbb{N}^{p+q-1} \times \mathbb{N}^{p+q+1}$ to $\mathbb{N}^{p+q-1}$ defined, for $p \geq 2$ and $q \geq 2$, by

$$\underline{m}' \star \underline{l} = (l_1 + m_1 + l_{q+1} + n_1, l_2 + m_2, \dots, l_q + m_q, l_{q+2} + n_2, \dots, l_{p+q} + n_p),$$

where $\underline{m}'$ is defined as above and $\underline{l} = (l_0, l_1, \dots, l_{p+q})$. For convenience, the operation carried out by $\star$ is displayed in Figure 4.1.

$$
\begin{array}{rcl}
\underline{m}' & \to & (m_0,\ m_1,\ \ldots,\ m_{q-1},\ 0,\ n_1,\ \ldots,\ n_{p-1},\ 0) \\
\underline{l} & = & (l_0,\ l_1,\ \ldots,\ l_{q-1},\ l_q,\ l_{q+1},\ \ldots,\ l_{p+q-1},\ l_{p+q}) \\
 & & \Big\downarrow \text{ sum} \\
 & & (l_0 + m_0, \boxed{l_1 + m_1}, \ldots, l_{q-1} + m_{q-1}, l_q, \boxed{l_{q+1} + n_1}, \ldots, l_{p+q-1} + n_{p-1}, l_{p+q}) \\
 & & \Big\downarrow \text{ cancellation of the 1-st element, summation of the boxed terms} \\
 & & (\boxed{l_1 + m_1 + l_{q+1} + n_1}, l_2 + m_2, \ldots, l_{q-1} + m_{q-1}, l_q, l_{q+2} + n_2, \ldots, l_{p+q-1} + n_{p-1}, l_{p+q}) \\
 & = & \underline{m}' \star \underline{l}
\end{array}
$$

FIGURE 4.1: Operator $\star$

Because we are interested in the autocovariance structure of ϵ_t^{2m}, we will focus on variables $\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}$ such that the sum of the exponents in the right-hand side of (4.9) is less than or equal to $2m$. Similarly, having applied the multinomial formula to $h_t^{m_0}$, the sum of the elements of $\underline{l}$ in (4.10) is equal to m_0. We will therefore introduce simplexes on $\mathbb{N}^{p+q-1}$ and $\mathbb{N}^{p+q+1}$ defined, for $m' \leq m$ by

$$\mathcal{R}_{m'} = \{\underline{m}' \in \mathbb{N}^{p+q-1} \mid m_0 + \cdots + m_{q-1} + n_1 + \cdots + n_{p-1} = m'\},$$

with by convention $n_1 + \cdots + n_{p-1} = 0$ if $p = 1$, and for any positive integer $m_0 \leq m$, by

$$\mathcal{S}_{m_0} = \{\underline{l} = (l_0, \ldots, l_{p+q}) \in \mathbb{N}^{p+q+1} \mid l_0 + \ldots + l_{p+q} = m_0\}.$$

Denote the first component of $\underline{m}'$ by $m_0(\underline{m}')$. We now order the elements of $\widetilde{\mathcal{R}}_m := \bigcup_{m'=0}^{m} \mathcal{R}_{m'}$ in the following way: if $\underline{m}' \in \mathcal{R}_{m'}$ and $\underline{m}'' \in \mathcal{R}_{m''}$ with $m' > m''$ then we say that $\underline{m}' > \underline{m}''$; for the elements of $\mathcal{R}_{m'}$ we take the so-called lexicographical order. Let $\tau(\underline{m}')$ denote the rank of any element $\underline{m}'$ in $\widetilde{\mathcal{R}}_m$. For instance $\tau(m, 0, \ldots, 0) = |\widetilde{\mathcal{R}}_m|$, where $|A|$ denotes the cardinal of any finite set A, and $\tau(0, \ldots, 0) = 1$. Let

$$\mathcal{A}_m = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1|\widetilde{\mathcal{R}}_m|} \\ A_{21} & A_{22} & \cdots & A_{2|\widetilde{\mathcal{R}}_m|} \\ \vdots & \ddots & & \vdots \\ A_{|\widetilde{\mathcal{R}}_m|-1,1} & \cdots & & A_{|\widetilde{\mathcal{R}}_m|-1,|\widetilde{\mathcal{R}}_m|} \\ 0 & \cdots & 0 & \mathbb{P} \end{pmatrix} \tag{4.11}$$

where

$$A_{ij} = \sum_{\underline{l} \in \mathcal{L}_{i,j}} \mathbb{P}_{\xi_{\tau^{-1}\{|\widetilde{\mathcal{R}}_m|-i+1\},\underline{l}}} = \mathbb{P}_{\sum_{\underline{l} \in \mathcal{L}_{i,j}} \xi_{\tau^{-1}\{|\widetilde{\mathcal{R}}_m|-i+1\},\underline{l}}}$$

and $\mathcal{L}_{i,j} = \{\underline{l} \in \mathcal{S}_{m_0\{\tau^{-1}(|\widetilde{\mathcal{R}}_m|-i+1)\}} \quad | \quad \tau\{\tau^{-1}(|\widetilde{\mathcal{R}}_m| - i + 1) \star \underline{l}\} = |\widetilde{\mathcal{R}}_m| - j + 1\}$.

Remark 4.4. It can be seen that the sum of the elements of $\underline{m}' \star \underline{l}$ is always less than that of the elements of $\underline{m}'$ when $\underline{l} \in \mathcal{S}_{m_0(\underline{m}')}$. So $A_{ij} = 0$ if $i > |\mathcal{R}_m| + \dots |\mathcal{R}_{m-k}|$ and $j \leq |\mathcal{R}_m| + \dots |\mathcal{R}_{m-k}|$, for any $k \geq 0$. The matrix $\mathcal{A}_m$ is therefore upper block-triangular, with blocks of size $d|\mathcal{R}_m|, \dots, d|\mathcal{R}_1|$ and $d|\mathcal{R}_0| = d$ on the diagonal.

Remark 4.5. From $n^{(k)} = n(n-1)\dots(n-k+1)$ and the formula $\sum_{n=k}^{N} n^{(k)} = \frac{(N+1)^{(k+1)}}{k+1}$ we have

$$\begin{aligned} |\widetilde{\mathcal{R}}_m| &= \sum_{m'=0}^{m} |\mathcal{R}_{m'}| = \sum_{m'=0}^{m} \frac{(m'+p+q-2)^{(p+q-2)}}{(p+q-2)!} \\ &= \frac{(m+p+q-1)^{(p+q-1)}}{(p+q-1)!} = \begin{pmatrix} m+p+q-1 \\ p+q-1 \end{pmatrix}. \end{aligned}$$

Remark 4.6. The previous notations can be extended as follows to the case $p = 1$ or $q = 1$. The definition of $\Upsilon_t(\underline{m}')$ in (4.9) remains valid for $q = 1$, and it can be extended to the case $p = 1$ with the following convention, which we will use throughout: $X_1^{a_1} \dots X_n^{a_n} = 1$ if $n < 1$. For instance if $p = q = 1$ we have $\mathcal{R}_{m'} = \{m'\}$ and $\Upsilon_t(\underline{m}') = \Upsilon_t(m') = \epsilon_t^{2m'}$. The definition of the operator $\star$ is extended as follows, for $p = 1$ and $q \geq 2$, $\underline{m}' \star \underline{l} := (l_1 + m_1 + l_{q+1}, l_2 + m_2, \dots, l_q + m_q) \in \mathbb{N}^q$; for $p \geq 2$ and $q = 1$, $\underline{m}' \star \underline{l} := (l_1 + l_2 + n_1, l_3 + n_2, \dots, l_{p+1} + n_p) \in \mathbb{N}^p$; for $p = 1$ and $q = 1$, $\underline{m}' \star \underline{l} = l_1 + l_2$.

To illustrate these notations, consider the following example.

Example 4.3: Let $p = 1, q = 2$ and $m = 2$. Then $\widetilde{\mathcal{R}}_2$ has 6 elements: $(2,0) > (1,1) > (0,2) > (1,0) > (0,1) > (0,0)$. Hence $\tau(0,0) = 1, \tau(0,1) = 2, \tau(1,0) = 3, \tau(0,2) = 4, \tau(1,1) = 5, \tau(2,0) = 6$. For instance let us compute A_{11}. We have

$$m_0\{\tau^{-1}(6)\} = m_0\{(2,0)\} = 2, \quad \mathcal{S}_2 = \{(2,0,0,0), (0,2,0,0), \dots, (0,0,1,1)\}.$$

Moreover, $(2,0) \star (l_0, l_1, l_2, l_3) = (l_1 + l_3, l_2)$. Hence

$$\mathcal{L}_{1,1} = \{\underline{l} \in \mathcal{S}_2 | \tau\{(l_1 + l_3, l_2)\} = 6\} = \{(0,2,0,0), (0,0,0,2), (0,1,0,1)\}.$$

Therefore we compute

$$\xi_{(2,0),(0,2,0,0)}(\cdot) = \mu_4\alpha_1^2(\cdot), \quad \xi_{(2,0),(0,0,0,2)}(\cdot) = \beta_1^2(\cdot), \quad \xi_{(2,0),(0,1,0,1)}(\cdot) = 2\alpha_1(\cdot)\beta_1(\cdot).$$

Thus $A_{11} = \mathbb{P}_{\alpha_1^2\mu_4+2\alpha_1\beta_1+\beta_1^2}$. Proceeding in the same way for the other matrices A_{ij} we find

$$\mathcal{A}_2 = \begin{pmatrix} \mathbb{P}_{\alpha_1^2\mu_4+2\alpha_1\beta_1+\beta_1^2} & \mathbb{P}_{2\alpha_2(\alpha_1+\beta_1)\mu_4} & \mathbb{P}_{\alpha_2^2\mu_4} & \mathbb{P}_{2\omega(\alpha_1+\beta_1)\mu_4} & \mathbb{P}_{2\omega\alpha_2\mu_4} & \mathbb{P}_{\omega^2\mu_4} \\ \mathbb{P}_{\alpha_1+\frac{\beta_1}{\mu_4}} & \mathbb{P}_{\alpha_2} & 0 & \mathbb{P}_{\omega} & 0 & 0 \\ \mathbb{P} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{P}_{\alpha_1+\beta_1} & \mathbb{P}_{\alpha_2} & \mathbb{P}_{\omega} \\ 0 & 0 & 0 & \mathbb{P} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbb{P} \end{pmatrix}. \tag{4.12}$$

Similarly, when $p = 2, q = 1$ and $m = 2$, we obtain

$$\mathcal{A}_2 = \begin{pmatrix} \mathbb{P}_{\alpha_1^2\mu_4+2\alpha_1\beta_1+\beta_1^2} & \mathbb{P}_{2\beta_2(\alpha_1+\beta_1)\mu_4} & \mathbb{P}_{\beta_2^2\mu_4} & \mathbb{P}_{2\omega(\alpha_1+\beta_1)\mu_4} & \mathbb{P}_{2\omega\beta_2\mu_4} & \mathbb{P}_{\omega^2\mu_4} \\ \mathbb{P}_{\frac{\alpha_1+\beta_1}{\mu_4}} & \mathbb{P}_{\beta_2} & 0 & \mathbb{P}_{\omega} & 0 & 0 \\ \mathbb{P}_{\frac{1}{\mu_4}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{P}_{\alpha_1+\beta_1} & \mathbb{P}_{\beta_2} & \mathbb{P}_{\omega} \\ 0 & 0 & 0 & \mathbb{P} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbb{P} \end{pmatrix}.$$

We are now in position to state the following result.

Theorem 4.2. *Let $p, q \geq 1$ and suppose $\rho\{\mathbb{P}_{C^{(m')}}\} < 1$ for all $m' \leq 2m$, where $C^{(m')}(\cdot)$ is defined in Theorem 2.1. Let (ϵ_t) be the nonanticipative $4m$th-order stationary solution of model (1.1).*

Then (ϵ_t^{2m}) is an ARMA(N_2, N_2) process, with

$$N_2 := \begin{pmatrix} m+p+q-1 \\ p+q-1 \end{pmatrix} \{d - n_0(\mathbb{P})\} - n_1(\mathcal{A}_m).$$

Remark 4.7. In the case $p = q = 1$ it can be seen, in view of Remark 4.4, that $\mathcal{A}_m = A_m$, where A_m is the $d(m+1) \times d(m+1)$ matrix of Section 4.1. Thus the results of Theorems 4.1 and 4.2 are the same in this special case.

Remark 4.8. Define f_{ij} such that $A_{ij} = \mathbb{P}_{f_{ij}}$ for $i,j = 1,\dots,|\widetilde{\mathcal{R}}_m|$ and let $D_{ij} = \text{diag}\,(f_{ij}(1),\dots,f_{ij}(d))$. Then we have

$$\mathcal{A}_m = \begin{pmatrix} D_{11} & D_{12} & \cdots & D_{1|\widetilde{\mathcal{R}}_m|} \\ D_{21} & D_{22} & \cdots & D_{2|\widetilde{\mathcal{R}}_m|} \\ \vdots & \ddots & & \vdots \\ D_{|\widetilde{\mathcal{R}}_m|-1,1} & \cdots & & D_{|\widetilde{\mathcal{R}}_m|-1,|\widetilde{\mathcal{R}}_m|} \\ 0 & \cdots & 0 & I_d \end{pmatrix} \begin{pmatrix} \mathbb{P} & 0 & \cdots & 0 \\ 0 & \mathbb{P} & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \mathbb{P} & 0 \\ 0 & \cdots & 0 & \mathbb{P} \end{pmatrix}. \tag{4.13}$$

It follows that

$$\det(\mathcal{A}_m) = \{\det(\mathbb{P})\}^{|\widetilde{\mathcal{R}}_m|} \prod_{k=1}^{d} \det f(k), \tag{4.14}$$

where $f(.)$ is the $(|\widetilde{\mathcal{R}}_m|-1)\times(|\widetilde{\mathcal{R}}_m|-1)$ matrix with (i,j)th element $f_{ij}(.)$, and that

$$n_0(\mathcal{A}_m) \geq |\widetilde{\mathcal{R}}_m| n_0(\mathbb{P}),$$

with equality if and only if $\det f(\cdot) \neq 0$.

Proof. Again, the proof is divided into three steps.

i) We first derive the autocovariance function $\gamma_{\epsilon^{2m}}(\ell)$ of ϵ^{2m}.

For $\ell > 1$, $p > 1$ and $q > 1$ we have, by Lemmas 3.2 and 4.1

$$\begin{aligned}
& \pi(k)E[\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}|\Delta_t = k] \\
= \; & \mu_{2m_0} E[h_t^{m_0}\epsilon_{t-1}^{2m_1}\dots\epsilon_{t-q+1}^{2m_{q-1}} h_{t-1}^{n_1}\dots h_{t-p+1}^{n_{p-1}}\epsilon_{t-\ell}^{2m}|\Delta_t = k]\pi(k) \\
= \; & \mu_{2m_0}\sum_{\underline{l}\in\mathcal{S}_{m_0}} \frac{m_0!}{\Pi_{j=0}^{p+q} l_j!} E[\omega(\Delta_t)^{l_0}\alpha_1(\Delta_t)^{l_1}\dots\alpha_q(\Delta_t)^{l_q}\beta_1(\Delta_t)^{l_{q+1}}\dots\beta_p(\Delta_t)^{l_{p+q}} \\
& \qquad \times \epsilon_{t-1}^{2(l_1+m_1)}\epsilon_{t-2}^{2(l_2+m_2)}\dots\epsilon_{t-q+1}^{2(l_{q-1}+m_{q-1})}\epsilon_{t-q}^{2l_q} \\
& \qquad \times h_{t-1}^{l_{q+1}+n_1}\dots h_{t-p+1}^{l_{p+q-1}+n_{p-1}} h_{t-p}^{l_{p+q}}\epsilon_{t-\ell}^{2m}|\Delta_t = k]\pi(k) \\
= \; & \sum_{\underline{l}\in\mathcal{S}_{m_0}} \xi_{\underline{m}',\underline{l}}(k)\sum_{j=1}^{d} E[\epsilon_{t-1}^{2(l_1+m_1)}\dots\epsilon_{t-q}^{2(l_q+m_q)}\epsilon_{t-1}^{2(l_{q+1}+n_1)} \\
& \qquad \times h_{t-2}^{l_{q+2}+n_2}\dots h_{t-p}^{l_{p+q}+n_p}\epsilon_{t-\ell}^{2m}|\Delta_{t-1} = j]p(j,k)\pi(j)
\end{aligned} \tag{4.15}$$

with $m_q = n_p = 0$. It can be seen that, in view of Remark 4.6, (4.15) remains valid for $p \geq 1$ and $q \geq 1$. Hence (4.15) generalizes (4.5), obtained for $p = q = 1$. Note that

$$\epsilon_{t-1}^{2(l_1+m_1)}\dots\epsilon_{t-q}^{2(l_q+m_q)}\epsilon_{t-1}^{2(l_{q+1}+n_1)} h_{t-2}^{l_{q+2}+n_2}\dots h_{t-p}^{l_{p+q}+n_p} = \Upsilon_{t-1}(\underline{m}'\star\underline{l})$$

and that for $\underline{l} \in \mathcal{S}_{m_0}$, $\underline{m}' \star \underline{l}$ belongs to $\mathcal{R}_{m''}$, with $m'' = m' - l_0 \leq m'$. Hence (4.15) can be written as

$$E[\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}|\Delta_t = k]\pi(k) = \sum_{\underline{l}\in\mathcal{S}_{m_0}} \xi_{\underline{m}',\underline{l}}(k) \sum_{j=1}^{d} E[\Upsilon_{t-1}(\underline{m}' \star \underline{l})\epsilon_{t-\ell}^{2m}|\Delta_{t-1} = j]p(j,k)\pi(j)$$

or, in vector form, as

$$\Pi_{E[\Upsilon_t(\underline{m}')\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} = \sum_{\underline{l}\in\mathcal{S}_{m_0}} \mathbb{P}_{\xi_{\underline{m}',\underline{l}}} \Pi_{E[\Upsilon_t(\underline{m}'\star\underline{l})\epsilon_{t-\ell+1}^{2m}|\Delta_t=\cdot\,]} \quad \text{for all} \quad \ell > 1. \tag{4.16}$$

Note that this formula is analogous to (4.6), obtained for $p = q = 1$. In view of (4.16) we have, for $\ell > 1$,

$$\mathcal{W}_m(\ell) \quad := \quad \begin{pmatrix} \Pi_{E[\Upsilon_t\{\tau^{-1}(|\widetilde{\mathcal{R}}_m|)\}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \\ \Pi_{E[\Upsilon_t\{\tau^{-1}(|\widetilde{\mathcal{R}}_m|-1)\}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \\ \vdots \\ \Pi_{E[\Upsilon_t\{\tau^{-1}(1)\}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} \end{pmatrix} = \mathcal{A}_m\mathcal{W}_m(\ell-1). \tag{4.17}$$

Now remark that

$$\Pi_{E[\Upsilon_t\{\tau^{-1}(|\widetilde{\mathcal{R}}_m|)\}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} = \Pi_{E[\epsilon_t^{2m}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},$$

and

$$\Pi_{E[\Upsilon_t\{\tau^{-1}(1)\}\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]} = \Pi_{E[\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]}.$$

Hence,

$$\gamma_{\epsilon^{2m}}(\ell) = v'\mathcal{W}_m(\ell),$$

with $v' = (\underline{1}_d', 0, \ldots, 0, -E(\epsilon_t^{2m})\underline{1}_d')$.

ii) By Lemma 3.1, with $n_0 = n_0(\mathcal{A}_m), \ell_0 = 1$, $n = d|\widetilde{\mathcal{R}}_m|$ and $\mathcal{P} = \mathcal{P}_{\mathcal{A}_m}$, we find that (ϵ_t^{2m}) admits an ARMA representation, whose orders are $(d|\widetilde{\mathcal{R}}_m| - n_0(\mathcal{A}_m) - n_1(\mathcal{A}_m), d|\widetilde{\mathcal{R}}_m| - n_1(\mathcal{A}_m))$ after cancellation of the roots of modulus 1 (it can be shown as in Step *ii)* of the proof of Theorem 3.1 that $\epsilon_t^{2m} - E(\epsilon_t^{2m})$ is purely nondeterministic). When $n_0(\mathbb{P}) = 0$, we get the announced result (with an over-identification of the AR order, as given in the theorem, if $\det f(k) = 0$ for some integers k. See Remark 4.8).

iii) Now suppose $n_0(\mathbb{P}) \neq 0$. With the notations of the proof of Theorem 3.1, we introduce the block-diagonal $|\widetilde{\mathcal{R}}_m|\{d - n_0(\mathbb{P})\} \times d|\widetilde{\mathcal{R}}_m|$ matrix $\Lambda_{\mathcal{A}_m} = \mathrm{diag}\,(\Lambda_{\mathbb{P}}, \ldots, \Lambda_{\mathbb{P}})$, and the matrices $\mathcal{A}_m^*$ and $\tilde{\mathcal{A}}_m$ such that $\mathcal{A}_m = \mathcal{A}_m^*\Lambda_{\mathcal{A}_m}$, $\tilde{\mathcal{A}}_m = \Lambda_{\mathcal{A}_m}\mathcal{A}_m^*$. We have

$$\tilde{\mathcal{W}}_m(\ell) := \Lambda_{\mathcal{A}_m}\mathcal{W}_m(\ell) = \tilde{\mathcal{A}}_m\tilde{\mathcal{W}}_m(\ell-1), \quad \forall \ell > 1.$$

Thus $v' = \tilde{v}'\Lambda_{\mathcal{A}_m}$ with $\tilde{v}' = (\underline{1}'_{d-n_0(\mathbb{P})}, 0, \dots, 0, -E(\epsilon_t^{2m})\underline{1}'_{d-n_0(\mathbb{P})})$. Hence $\gamma_{\epsilon^{2m}}(\ell) = \tilde{v}'\tilde{\mathcal{W}}_m(\ell)$. The proof is completed as that of Theorem 3.1 by applying Lemma 3.1.

EXAMPLE 4.3 (CONTINUED): When $p = 1, q = 2$ and $m = 2$, the quantities introduced in the proof take the form

$$\begin{aligned}
\Upsilon_t(\underline{m}') &= \epsilon_t^{2m_0}\epsilon_{t-1}^{2m_1}, \qquad \underline{m}' = (m_0, m_1) \in \widetilde{\mathcal{R}}_2 = \{(2,0),(1,1),(0,2),(1,0),(0,1),(0,0)\},\\
\mathcal{W}_m(\ell) &= \Big(\Pi_{E[\epsilon_t^4\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},\ \Pi_{E[\epsilon_t^2\epsilon_{t-1}^2\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},\ \Pi_{E[\epsilon_{t-1}^4\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},\\
&\qquad \Pi_{E[\epsilon_t^2\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},\ \Pi_{E[\epsilon_{t-1}^2\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]},\ \Pi_{E[\epsilon_{t-\ell}^{2m}|\Delta_t=\cdot\,]}\Big)'.
\end{aligned}$$

Moreover, we have by (4.14), $\det\mathcal{A}_2 = (\det\mathbb{P})^6\prod_{k=1}^d \alpha_2^4(k)\mu_4$. Thus the process (ϵ_t^4) is solution of an ARMA$[6\{d-n_0(\mathbb{P})\}-n_1(\mathcal{A}_2), 6\{d-n_0(\mathbb{P})\}-n_1(\mathcal{A}_2)]$ equation. If $\alpha_2(k) = 0$ for some k, the AR order can be reduced.

To be more specific suppose that $d = 1$ and take $\alpha_1 = \beta_1 = 0$. Then the model under consideration is the standard ARCH(2) model of the form : $\epsilon_t = \sqrt{\omega + \alpha_2\epsilon_{t-2}^2}\eta_t$. We have, by (4.12),

$$\mathcal{A}_2 = \begin{pmatrix} 0 & 0 & \alpha_2^2\mu_4 & 0 & 2\omega\alpha_2\mu_4 & \omega^2\mu_4\\ 0 & \alpha_2 & 0 & \omega & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \alpha_2 & \omega\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

so that

$$\mathcal{P}_{\mathcal{A}_2}(x) = (x^2-\alpha_2)(x-\alpha_2)(x^2-\alpha_2^2\mu_4)(x-1).$$

Applying Theorem 4.2 we therefore obtain an ARMA(5,5) model for (ϵ_t^4), with AR polynomial: $(1-\alpha_2L^2)(1-\alpha_2L)(1-\alpha_2^2\mu_4L^2)$. Direct calculations confirm these results and allow to reduce the ARMA orders. We have $\epsilon_t^2 = \omega + \alpha_2\epsilon_{t-2}^2 + h_t(\eta_t^2-1)$. Hence

$$(1-\alpha_2L^2)\epsilon_t^2 = \omega + u_t \tag{4.18}$$

where $u_t = h_t(\eta_t^2-1)$. Moreover, from $\epsilon_t^4 = \{\omega^2 + \alpha_2^2\epsilon_{t-2}^4 + 2\omega\alpha_2\epsilon_{t-2}^2\}\mu_4 + h_t^2(\eta_t^4-\mu_4)$ we get $(1-\alpha_2^2\mu_4L^2)\epsilon_t^4 = \{\omega^2 + 2\omega\alpha_2\epsilon_{t-2}^2\}\mu_4 + v_t$, where $v_t = h_t^2(\eta_t^4-\mu_4)$, and thus by (4.18)

$$(1-\alpha_2^2\mu_4L^2)(1-\alpha_2L^2)\epsilon_t^4 = \omega^2\mu_4(1+\alpha_2) + (1-\alpha_2L^2)v_t + 2\omega\alpha_2\mu_4u_{t-2}.$$

It is straightforward to show that $(1-\alpha_2 L^2)v_t + 2\omega\alpha_2\mu_4 u_{t-2}$ is a MA(2) process. Finally, (ϵ_t^4) is the solution of an ARMA(4,2) equation. The MA polynomial can be obtained numerically from the first autocovariances of (ϵ_t^4), for given values of α_2 and μ_4.

To conclude, it should be noted that Theorem 4.2 is of invaluable help for more complex models or higher powers than those of this example. In such cases the direct method for determining the ARMA orders and AR polynomial cannot be applied.

Appendix A.

In this appendix we show how to compute the ARMA coefficients, as well as the autocovariance function of (ϵ_t^{2m}). For simplicity, we will only consider the GARCH(1,1) case. For convenience we recall the model

$$\left\{ \begin{array}{rcl} \epsilon_t & = & \sqrt{h_t}\eta_t, \\ h_t & = & \omega(\Delta_t) + \alpha(\Delta_t)\epsilon_{t-1}^2 + \beta(\Delta_t)h_{t-1}. \end{array} \right. \tag{A.1}$$

Recall that, by Corollary 2.2, $E\epsilon_t^{4m} < \infty$ if and only if $\rho\left[\mathbb{P}_{E\{\alpha(\cdot)\eta_t^2+\beta(\cdot)\}^{2m}}\right] < 1$. Under this assumption, we first compute the mean and autocovariances of ϵ_t^{2m}.

A.1. Deriving the autocovariances

For $0 \le m_2 \le m_1 \le m$, let

$$\xi^*_{m,m_1,m_2}(k) = \left(\begin{array}{c} m_1 \\ m_2 \end{array} \right) \{\omega(k)\}^{m_1-m_2} E\left[\{\alpha(k)\eta_{t-1}^2 + \beta(k)\}^{m_2}\eta_{t-1}^{2m}\right] \frac{1}{\mu_{2(m_2+m)}}. \tag{A.2}$$

The following result allows to compute the autocovariances of ϵ_t^{2m} conditional on Δ_t.

Theorem A.1. *Under the assumptions of Theorem 4.1,*

$$\begin{aligned} &\Pi_{E[\epsilon_t^{2m_1}|\Delta_t=\cdot\,]} \\ &= \left\{ \begin{array}{ll} \Pi_1 & \textit{for } m_1 = 0 \\ \left\{I - \mathbb{P}_{\xi_{m_1,m_1}}\right\}^{-1} \sum_{m_2=0}^{m_1-1} \mathbb{P}_{\xi_{m_1,m_2}} \Pi_{E[\epsilon_t^{2m_2}|\Delta_t=\cdot\,]} & \textit{for } m_1 = 1,\dots,2m. \end{array} \right. \end{aligned} \tag{A.3}$$

For $m_1 = 1,\dots,m$

$$\Pi_{E[\epsilon_t^{2m_1}\epsilon_{t-1}^{2m}|\Delta_t=\cdot\,]} = \sum_{m_2=0}^{m_1} \mathbb{P}_{\xi^*_{m,m_1,m_2}} \Pi_{E[\epsilon_t^{2(m_2+m)}|\Delta_t=\cdot\,]}. \tag{A.4}$$

For $\ell > 1$,

$$W_m(\ell) = A_m W_m(\ell - 1). \tag{A.5}$$

This theorem allows to compute $W_m(\ell)$ for all $\ell \geq 0$. The unconditional means follow from $E(\epsilon_t^{2m_1}) = \underline{1}_d' \Pi_{E[\epsilon_t^{2m_1}|\Delta_t = \cdot\,]}$. The unconditional autocovariances similarly follow by summing appropriate components of $W_m(\ell)$.

Example 3.1 (continued): Simple computations show that $\Pi_1 = (\pi(1), \pi(2))' = (0.2, 0.8)'$. We complete the model by setting $\omega(1) = 1$ and $\omega(2) = 2$. Applying (A.3) yields:

$$\Pi_{E[\epsilon_t^2|\Delta_t=\cdot\,]} = \begin{pmatrix} \pi(1)E[\epsilon_t^2|\Delta_t = 1] \\ \pi(2)E[\epsilon_t^2|\Delta_t = 2] \end{pmatrix} = \begin{pmatrix} \frac{0.2}{0.82} \\ \frac{1.6}{0.82} \end{pmatrix}, \quad \Pi_{E[\epsilon_t^4|\Delta_t=\cdot\,]} = \begin{pmatrix} 0.909 \\ 14.557 \end{pmatrix}.$$

Therefore: $E(\epsilon_t^2) = 2.195$ and $E(\epsilon_t^4) = 15.467$. Hence $\gamma_{\epsilon^2}(0) = 10.648$. Now, applying (A.4) with $m = m_1 = 1$, we get:

$$\Pi_{E[\epsilon_t^2\epsilon_{t-1}^2|\Delta_t=\cdot\,]} = \begin{pmatrix} \pi(1)E[\epsilon_t^2\epsilon_{t-1}^2|\Delta_t = 1] \\ \pi(2)E[\epsilon_t^2\epsilon_{t-1}^2|\Delta_t = 2] \end{pmatrix} = \begin{pmatrix} 0.645 \\ 5.162 \end{pmatrix},$$

hence $\gamma_{\epsilon^2}(1) = 0.989$. In view of the AR coefficient, and because ϵ_t^2 admits an ARMA(1,1) representation in this particular case, we have that $\gamma_{\epsilon^2}(\ell) = (0.18)^{\ell-1}0.989$, for $\ell \geq 1$. To conclude, let us derive the MA coefficient, b say, as a solution to the second-order equation $b/\{1+b^2\} = \mathrm{cov}(\epsilon_t^2 - 0.18\epsilon_{t-1}^2, \epsilon_{t-1}^2 - 0.18\epsilon_{t-2}^2)/\left\{\mathrm{var}(\epsilon_t^2 - 0.18\epsilon_{t-1}^2)\right\} = -0.087$. Choosing the solution of modulus less than 1, we obtain the ARMA(1,1) representation:

$$\epsilon_t^2 - 2.195 = 0.18(\epsilon_{t-1}^2 - 2.195) + u_t - 0.088u_{t-1}, \quad \mathrm{var}(u_t) = 10.556.$$

In general, when $d > 2$, the ARMA representations involve higher orders and such computations cannot be done by hand. An algorithm is proposed in the next section to obtain the ARMA coefficients.

A.2. Deriving the ARMA representation

In Theorem 4.1, we have shown that, provided (ϵ_t^{2m}) is second-order stationary, it admits an ARMA representation of the form

$$\left(1 + \sum_{i=1}^{md+d-1} a_i^{(m)} L^i\right)\left\{\epsilon_t^{2m} - E\epsilon_t^{2m}\right\} = \left(1 + \sum_{i=1}^{md+d-1} b_i^{(m)} L^i\right) u_t^{(m)}, \tag{A.6}$$

because $n_1(A_m) \geq n_1(\mathbb{P}) = 1$ and $n_0(\mathbb{P}) \geq 0$. It can be assumed, without generality loss, that (A.6) is the canonical ARMA representation, i.e. the AR and MA polynomials have their roots outside the unit circle and have no common root. In this section, we provide an algorithm for computing the coefficients $a_i^{(m)}$, $b_i^{(m)}$ and $\mathrm{var}(u_t^{(m)})$ in function of Model (A.1) parameters (i.e. $\omega(i), \alpha(i), \beta(i), p(i,j)$, for $i, j \in \mathcal{E}$, and $\mu_{2m_1} = E(\eta_t^{2m_1})$, for $m_1 = 1, \ldots, 2m$.) In the Gaussian case, for example, we have $\mu_{2m_1} = (2m_1)!/(2^{m_1} m_1!)$. A practical approach to deriving the ARMA representation (A.6) proceeds as follows.

Step 1 (definition of intermediate tools): Define the functions

$$\xi_{m_1,m_2}(k) = \binom{m_1}{m_2} \{\omega(k)\}^{m_1-m_2} \frac{\mu_{2m_1}}{\mu_{2m_2}} \sum_{m_3=0}^{m_2} \binom{m_2}{m_3} \alpha(k)^{m_3} \mu_{2m_3} \beta(k)^{m_2-m_3},$$

$$\xi^*_{m,m_1,m_2}(k) = \binom{m_1}{m_2} \frac{\{\omega(k)\}^{m_1-m_2}}{\mu_{2(m_2+m)}} \sum_{m_3=0}^{m_2} \binom{m_2}{m_3} \alpha(k)^{m_3} \mu_{2(m_3+m)} \beta(k)^{m_2-m_3},$$

and $\mathbb{P}_{\cdot}$, introduced in (4.1), (A.2), and (2.1), respectively.

Step 2 (moments existence): If $\rho\left(\mathbb{P}_{\xi_{2m,2m}}\right) \geq 1$, then terminate. Otherwise, go to Step 3.

Step 3 (computation of the autocovariances): Compute, in this order,

- $\Pi_{E[\epsilon_t^{2m_1} | \Delta_t = \cdot\,]}$ for $m_1 = 1, \ldots, 2m$, using (A.3).
- $\Pi_{E[\epsilon_t^{2m_1} \epsilon_{t-1}^{2m} | \Delta_t = \cdot\,]}$ for $m_1 = 0, \ldots, m$ using (A.4).
- $\Pi_{E[\epsilon_t^{2m_1} \epsilon_{t-\ell}^{2m} | \Delta_t = \cdot\,]}$ for $m_1 = 0, \ldots, m$ and $\ell = 2, \ldots, 2m(d+1)-2$ using (A.5).
- $E\epsilon_t^{2m} = e'\Pi_{E[\epsilon_t^{2m} | \Delta_t = \cdot\,]}$, $\quad \gamma_{\epsilon^{2m}}(0) = e'\Pi_{E[\epsilon_t^{4m} | \Delta_t = \cdot\,]} - \left(E\epsilon_t^{2m}\right)^2$,
- $\gamma_{\epsilon^{2m}}(\ell) = e'\Pi_{E[\epsilon_t^{2m} \epsilon_{t-\ell}^{2m} | \Delta_t = \cdot\,]} - \left(E\epsilon_t^{2m}\right)^2$ for $\ell = 1, \ldots, 2d(m+1)-2$.

Step 4 (computation of the AR coefficients): Compute the eigenvalues of $\mathbb{P}, \mathbb{P}_{\xi_{1,1}}, \ldots, \mathbb{P}_{\xi_{m,m}}$ and replace those of modulus 1 by 0. Denote by $\lambda_1, \ldots, \lambda_{md+d-1}$, 0 these numbers. Compute the $a_i^{(m)}$ from the recursion formulas

$$\begin{cases} a_{0,0} &= 1, \\ a_{0,j} &= 1, \quad a_{i,j} = a_{i,j-1} - \lambda_j a_{i-1,j-1} \quad \forall i \in \{1, \ldots, j-1\}, \quad a_{j,j} = -\lambda_j a_{j-1,j-1} \\ & \qquad\qquad \text{for } j = 1, \ldots, md+d-1, \\ a_i^{(m)} &= a_{i,md+d-1}, \quad \forall i \in \{1, \ldots, md+d-1\}. \end{cases}$$

Step 5 (computation of the autocovariances of the AR part): `Compute recursively`

$$\begin{cases} \gamma^{(0)}(h) & = & \gamma_{\epsilon^{2m}}(h) \quad \texttt{for} \ \ h = 0,\ldots,2(md+d-1), \\ \gamma^{(i)}(h) & = & \gamma^{(i-1)}(h)(1+\lambda_i^2) - \lambda_i \left\{ \gamma^{(i-1)}(|h-1|) + \gamma^{(i-1)}(h+1) \right\} \\ & & \qquad \texttt{for} \ \ i = 1,\ldots,md+d-1 \quad \texttt{and} \ \ h = 0,\ldots,2(md+d-1)-i, \\ \gamma_{AR}(h) & = & \gamma^{(md+d-1)}(h) \quad \texttt{for} \ \ h = 0,\ldots,md+d-1. \end{cases}$$

`and set` $\quad \gamma_{AR}(h) = 0, \quad \forall h > md+d-1.$

Step 6 (computation of the MA coefficients and noise variance):

1. `Set` $i = 1$.

2. `Compute` $b_{i,1},\ldots,b_{i,i}$ `and` v_i, `from the recursion relations`

$$\begin{cases} v_0 & = & \gamma_{AR}(0), \\ b_{i,i-k} & = & v_k^{-1} \left\{ \gamma_{AR}(i-k) - \sum_{j=0}^{k-1} b_{i,i-j} b_{k,k-j} v_j \right\}, \quad k = 0,\ldots,i-1, \\ v_i & = & \gamma_{AR}(0) - \sum_{j=0}^{i-1} b_{i,i-j}^2 v_j. \end{cases}$$

3. `If` $i > md+d-1$ `and` $\max_{k=1,\ldots,md+d-1}\{|b_{i,k} - b_{i-1,k}|\}$ `is less than a given positive tolerance then set` $b_k^{(m)} = b_{i,k}$ `for` $k = 1,\ldots,md+d-1$. `Otherwise, replace` i `by` $i+1$ `and go to 2.`

A few comments are in order. The computation of the $a_i^{(m)}$ is derived from $1 + \sum_{i=1}^{md+d-1} a_i^{(m)} L^i = \prod_{i=1}^{md+d-1}(1-\lambda_i L)$. In Step 5 we compute the autocovariances $\gamma_{AR}(h)$ of the MA process

$$AR_t := \left(1 + \sum_{i=1}^{md+d-1} a_i^{(m)} L^i \right) \left\{ \epsilon_t^{2m} - E\epsilon_t^{2m} \right\} = \prod_{i=1}^{md+d-1} (1-\lambda_i L) \left\{ \epsilon_t^{2m} - E\epsilon_t^{2m} \right\}.$$

Simple computations show that $\gamma_{AR}(h) = \prod_{i=1}^{md+d-1}(1-\lambda_i L)(1-\lambda_i L^{-1})\gamma_{\epsilon^{2m}}(h), \quad \forall h \in \mathbb{Z}$ with the convention that $L^k \gamma_{\epsilon^{2m}}(h) = \gamma_{\epsilon^{2m}}(h-k)$ for all $k \in \mathbb{Z}$. An alternative way to compute these autocovariances is to use $\gamma_{AR}(h) = \sum_{i,j=0}^{md+d-1} a_i^{(m)} a_j^{(m)} \gamma_{\epsilon^{2m}}, (|h+j-i|)$ for $h = 0,\ldots,md+d-1$. In Step 6, we use the innovations algorithm presented in Brockwell and Davis (p. 172 and 245, 1991). By their Proposition 5.2.2 and Remark 3 p.176, $\lim_{i\to\infty} v_i = \text{Var}\left(u_t^{(m)}\right), \quad \lim_{i\to\infty} b_{i,k} = b_k^{(m)}, \quad \forall k \geq 1$, with $b_k^{(m)} = 0$ for $k > md+d-1$.

Remark A.1. In the particular case where $d=2$, the AR coefficients of Step 4 can be analytically obtained, thus avoiding numerical computation. The eigenvalues of $\mathbb{P}_{\xi_{0,0}} = \mathbb{P}$ are 1 and $\lambda_1 := 1 - p(1,1) - p(2,2)$. Straightforward calculations show that, for $m \geq 1$, the sum of the eigenvalues of $\mathbb{P}_{\xi_{m,m}}$ is $s_m := p(1,1)\xi_{m,m}(1) + p(2,2)\xi_{m,m}(2)$, and their product is $p_m := \xi_{m,m}(1)\xi_{m,m}(2)(p(1,1)+p(2,2)-1)$. Hence, for $m > 1$, the AR polynomial of ϵ_t^{2m} is obtained by multiplying the AR polynomial of $\epsilon_t^{2(m-1)}$ by $1 - s_m L + p_m L^2$. For instance, the AR coefficients $a_i^{(1)}, i = 0, \ldots, 3$ for $m=1$, and $a_i^{(2)}, i = 0, \ldots, 5$ for $m = 2$ are

$$a_0^{(1)} = 1, \quad a_1^{(1)} = -s_1 - \lambda_1, \quad a_2^{(1)} = p_1 + s_1\lambda_1, \quad a_3^{(1)} = -p_1\lambda_1,$$

$$a_0^{(2)} = 1, \quad a_1^{(2)} = a_1^{(1)} - s_2, \quad a_2^{(2)} = a_2^{(1)} - s_2 a_1^{(1)} + p_2, \quad a_3^{(2)} = a_3^{(1)} - s_2 a_2^{(1)} + p_2 a_1^{(1)},$$

$$a_4^{(2)} = -s_2 a_3^{(1)} + p_2 a_2^{(1)}, \quad a_5^{(2)} = p_2 a_3^{(1)}$$

with λ_1 replaced by 0 if it is equal to 1 or -1.

Main notations

(ϵ_t) MS GARCH process

(h_t) volatility process

(p,q) order of the MS GARCH model

(η_t) iid (0, 1) process

$\alpha_i(k), \beta_j(k)$ GARCH coefficients in the state k

(Δ_t) Markov chain

$\mathcal{E}$ state space of the Markov chain

d number of states of the Markov chain

$\pi(k)$ stationary probability of state k

$p(k,\ell)$ one-step transition probability

$p^{(i)}(k,\ell)$ i-step transition probability

$\mathbb{P}$ transpose of the transition matrix

f generic matrix function defined on $\mathcal{E}$

$\mathbb{P}_f^{(i)}$ matrix with (k,ℓ)-block $p^{(i)}(k,\ell)f(\ell)$

Π_f vector of matrix blocks $\pi(k)f(k)$

$\underline{z}_t = \left(\epsilon_t^2, \dots, \epsilon_{t-q+1}^2, h_t, \dots, h_{t-p+1}\right)'$

$\underline{z}_t = \underline{b}_t + C_t \underline{z}_{t-1}$ vector GARCH representation

m positive integer

$C^{(m)}$ function : $k \mapsto E(C_t^{\otimes m}|\Delta_t = k)$

$\underline{b}^{(m)}$ function : $k \mapsto E(\underline{b}_t^{\otimes m}|\Delta_t = k)$

μ_{2m} moment of order $2m$ of η_t

ξ_i function : $k \mapsto \alpha_i(k) + \beta_i(k)$

$r = \max(p,q)$

$n_0(M)$ number of zero eigenvalues of M

$n_1(M)$ number of eigenvalues of M of modulus 1

$\underline{1}_d = (1,\dots,1)'$ the d element sum vector

ξ_{m_1,m_2} functions in the case $p = q = 1 : k \mapsto \binom{m_1}{m_2}\{\omega(k)\}^{m_1-m_2} E[\alpha(k)\eta_{t-1}^2 + \beta(k)]^{m_2} \frac{\mu_{2m_1}}{\mu_{2m_2}}$

$M = M^* \Lambda_M$, $\tilde{M} = \Lambda_M M^*$ (M^* full rank matrix obtained by cancelling the redundant columns of M)

$\mathcal{R}_{m'}$ subset of $\mathbb{N}^{p+q-1}$ (elements sum equal to m')

$\widetilde{\mathcal{R}}_m$ subset of $\mathbb{N}^{p+q-1}$ (elements sum less than m)

$\mathcal{S}_{m_0}$ subset of $\mathbb{N}^{p+q+1}$ (elements sum equal to m_0)

$\star$ operator from $\mathbb{N}^{p+q-1} \times \mathbb{N}^{p+q+1}$ to $\mathbb{N}^{p+q-1}$

τ lexicographical ranking over $\widetilde{\mathcal{R}}_m$

$\underline{m}'$ element of $\mathbb{N}^{p+q-1}$

$\Upsilon_t(\underline{m}') = \epsilon_t^{2m_0} \dots \epsilon_{t-q+1}^{2m_{q-1}} h_{t-1}^{n_1} \dots h_{t-p+1}^{n_{p-1}}$

$\mathcal{P}_M(x)$ characteristic polynomial of a matrix M

$\xi_{\underline{m}',\underline{l}}$ functions $\mathcal{E} \to \mathbb{R}$ generalizing the ξ_{m_1,m_2}

References

[1] Azencott, R., and D. Dacunha-Castelle (1984). *Séries d'observations irrégulières.* Masson, Paris.

[2] Bai, X., Russel, J.R., and G.C. Tiao (2003). Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics.* **114,** 349–360.

[3] Berkes, I., and L. Horváth (2003). Asymptotic results for long memory LARCH sequences, *The Annals of Applied Probability.* **13,** 641–668.

[4] Billingsley, P. (1995). *Probability and Measure* 3rd ed., New York: John Wiley & Sons.

[5] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics.* **31,** 307–327.

[6] Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis.* **9,** 121–131.

[7] Bougerol, P., and N. Picard (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics.* **52,** 115–127.

[8] Brockwell, P. J., and R. A. Davis (1991). *Time Series: Theory and Methods.* Springer-Verlag New York.

[9] Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics.* **12,** 309–316.

[10] Dueker, M. J. (1997). Markov switching in GARCH processes and mean-reverting stock-market volatility. *Journal of Business & Economic Statistics.* **15,** 26–34.

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

[12] Francq, C., Roussignol, M., and J-M. Zakoïan (2001). Conditional heteroskedasticity driven by hidden Markov chains. *Journal of Time Series Analysis.* **22,** 197–220.

[13] Francq, C., and J-M. Zakoïan (2001). Stationarity of multivariate Markov-switching ARMA models. *Journal of Econometrics.* **102,** 339–364.

[14] Francq, C., and J-M. Zakoïan (2004). L^2-structures of standard and switching-regime GARCH models and their implications for statistical inference. Unpublished document, *http://www.crest.fr.*

[15] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH Models: Dependence Structure and Central Limit Theorem. *Econometric Theory.* 16, 3–22.

[16] Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics.* **42,** 27–62.

[17] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica.* **57,** 357–384.

[18] Hamilton, J. D., and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics.* **64,** 307–333.

[19] Hannan, E. J. (1970). *Multiple Time Series.* John Wiley, New York.

[20] Hilebrand, E. (2005). Neglecting parameter changes in GARCH models. *Forthcoming in the Journal of Econometrics.*

[21] Karlsen, H. (1990). *A class of non-linear time series models.* Ph. D. Thesis, University of Bergen, Norway.

[22] 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.

[23] Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis.* Lecture Notes in Economics and Mathematical Systems, Volume 454, Berlin: Springer.

[24] Lancaster, P., and M. Tismenetsky (1985). *The theory of matrices.* New-York: Academic Press.

[25] Ling, S., and M. McAleer (2002). Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r,s) models. *Econometric Theory.* **18,** 722–729.

[26] Mikosch, T. (2001). *Modeling financial time series.* Lecture in "New Directions in Time Series Analysis", CIRM, Luminy.

[27] Milhøj, A. (1985). The moment structure of ARCH processes. *Scandinavian Journal of Statistics.* **12,** 281–292.

[28] Poskitt, D. S., and S. H. Chung (1996). Markov chain models, time series analysis and extreme value theory. *Advances in Applied Probability.* **28,** 405–425.

[29] Wong, C. S., and W. K. Li (2001). On a Mixture Autoregressive Conditional Heteroskedastic Model. *Journal of the American Statistical Association.* **96,** 982–995.

[30] Zhang, J. and R.A. Stine (2001). Autocovariance structure of Markov regime switching models and model selection. *Journal of Time Series Analysis.* **22,** 107–124.