

AUTOCOVARANCE STRUCTURE OF POWERS OF SWITCHING-REGIME ARMA PROCESSES

CHRISTIAN FRANQC¹ AND JEAN-MICHEL ZAKOÏAN²

Abstract. In Francq and Zakoïan [4], we derived stationarity conditions for ARMA(p, q) models subject to Markov switching. In this paper, we show that, under appropriate moment conditions, the powers of the stationary solutions admit weak ARMA representations, which we are able to characterize in terms of p, q , the coefficients of the model in each regime, and the transition probabilities of the Markov chain. These representations are potentially useful for statistical applications.

Mathematics Subject Classification. 62M10.

INTRODUCTION

Consider the following multivariate ARMA model with random coefficients

$$X_t = c(\Delta_t) + \sum_{i=1}^p a_i(\Delta_t)X_{t-i} + \epsilon_t + \sum_{j=1}^q b_j(\Delta_t)\epsilon_{t-j}, \quad (1)$$

where X_t is a random vector with values in \mathbb{R}^K , (Δ_t) is an irreducible, aperiodic, Markov chain with finite state-space $\mathcal{E} = \{1, 2, \dots, d\}$ and stationary transition probabilities denoted $p(i, j) := \text{pr}(\Delta_t = j | \Delta_{t-1} = i)$, the $a_i(\Delta_t)$ and $b_j(\Delta_t)$ are $K \times K$ real random matrices, $c(\Delta_t)$ is a $K \times 1$ vector. It is assumed that (Δ_t) is stationary, with stationary probabilities denoted $\pi(i) := \text{pr}(\Delta_1 = i)$, $1 \leq i \leq d$. Write M^T for the transpose of any matrix M . We assume that

$$\epsilon_t = \epsilon_t(\Delta_t) = \sigma(\Delta_t)\eta_t$$

where $\sigma(\Delta_t)$ is a $K \times K$ random matrix and (η_t) is an independent and identically distributed sequence of K -dimensional centered variables, with $E(\eta_t \eta_t^T) = \Omega$, the covariance matrix Ω being nonsingular. In addition, assume that (η_t) is independent of (Δ_t) . Hence (ϵ_t) is a white noise.

The model (1) is called Markov-switching ARMA(p, q) model. Markov switching models (MSM) can be viewed as extensions of the hidden Markov models (HMM) introduced by Baum and Petrie [1]. By contrast with the HMM's, the observations of a MSM are not independent random variables conditional on the Markov chain. These models have found a variety of applications in econometrics since the paper by Hamilton [5]. In a recent paper, Francq and Zakoïan [4] have established necessary and sufficient conditions for the existence

Keywords and phrases: ARMA representation, hidden Markov models, Markov-switching models, identification.

¹ Université du Littoral-Côte d'Opale, LMPA J. Liouville, Centre Universitaire de la Mi-Voix, 50 rue F. Buisson, BP. 699, 62228 Calais Cedex, France; e-mail: Christian.Francq@lmpa.univ-littoral.fr

² Université de Lille 3 and CREST, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France; e-mail: zakoian@ensae.fr

of strict and second-order stationary solutions to model (1). Extending the results of Karlsen [7], Poskitt and Chung [9], and Zhang and Stine [12], obtained in the pure autoregressive case ($q = 0$), we showed that the covariance structure of any second-order stationary solution (X_t) is that of a vector ARMA model.

The purpose of this paper is to extend these previous results to the powers (in the Kronecker sense) of the (X_t) process. Such an extension is potentially useful for the statistical inference. In particular, identifying the number d of regimes in a HMM or a MSM can be a formidable task. In the recent literature, papers dealing with this problem use penalized-likelihood based information criteria (Leroux and Puterman [8], Rydèn [11]), Lagrange multiplier tests (Hamilton [6]), Bayesian procedures (Robert *et al.* [10]). The ARMA representations derived from model (1) should be useful for identifying d because, roughly speaking, their orders are increasing with d . Such a strategy has been followed by Zhang and Stine [12] in the case of HMM (which corresponds to $p = q = 0$ in (1)). However, for the general model, the orders of the ARMA representations also depend on p and q . Therefore, several ARMA representations, obtained for different powers of X_t , are likely to be needed to achieve complete identification of the orders of model (1).

In Section 2, we derive the autocovariance structure of (X_t) and its powers, under appropriate stationarity conditions. In Section 3 we consider three generic examples aimed to illustrate the general results. Section 3 concludes.

1. AUTOCOVARANCE STRUCTURES OF (X_t) AND ITS POWERS

Before dealing with the solutions of model (1), we recall some basic results on the characterization of the orders of a vector ARMA.

1.1. Characterization of an ARMA in terms of its vector autocovariance function

The autocovariance function of a K -multivariate second-order stationary process $X = (X_t)_{t \in \mathbb{Z}}$ is defined by

$$\Gamma_X(\ell) = E \left\{ (X_t - EX_t)(X_{t-\ell} - EX_{t-\ell})^T \right\}, \quad \ell \in \mathbb{Z}.$$

It will be convenient to work with the vector autocovariance function defined by

$$G_X(\ell) = E\{X_t \otimes X_{t-\ell}\} - EX_t \otimes EX_{t-\ell}, \quad \ell \in \mathbb{Z},$$

where \otimes denotes the matrix tensor product. The $K \times K$ identity matrix is denoted by I_K . The following result characterizes the existence of an ARMA representation through the vector autocovariance function of a stationary process.

Lemma 1. *Let (X_t) be a K -multivariate second-order stationary process with vector autocovariance function $G_X(\cdot)$. There exist $K \times K$ matrices $\Lambda_1, \dots, \Lambda_{p_0}$ such that*

$$G_X(\ell) = \sum_{i=1}^{p_0} (\Lambda_i \otimes I_K) G_X(\ell - i), \quad \forall \ell > q_0$$

if and only if (X_t) satisfies a (weak) ARMA(p_0, q_0) representation, that is

$$X_t = c + \sum_{i=1}^{p_0} \Lambda_i X_{t-i} + u_t + \sum_{i=1}^{q_0} \Theta_i u_{t-i}$$

where (u_t) is a K -multivariate white noise process, which is not independent in general, p_0, q_0 are integers, c is a $K \times 1$ vector and $\Lambda_1, \dots, \Lambda_{p_0}, \Theta_1, \dots, \Theta_{q_0}$ are $K \times K$ matrices.

Proof. It is well known that a K -multivariate stationary process $X = (X_t)$ follows an ARMA(p_0, q_0) equation if and only if its autocovariance function satisfies a linear difference equation of order p_0 from the rank $q_0 - p_0 + 1$:

$$\Gamma_X(\ell) = \sum_{i=1}^{p_0} \Lambda_i \Gamma_X(\ell - i), \quad \forall \ell > q_0.$$

In the univariate case ($K = 1$), this result is proved by a straightforward adaptation of Brockwell and Davis ([3], Prop. 3.2.1 and remark, p. 90). In the general case ($K \geq 1$), a detailed proof has been given by Berlinec ([2], Ths. 2 and 3). Using the elementary relation $\text{vec}(BC) = (C^T \otimes I)\text{vec}B$, provided the matrix product BC is well defined, the conclusion follows because

$$G_X(\ell) = \text{vec} \{ \Gamma_X(\ell)^T \}$$

and

$$\text{vec} \left\{ \left(\Gamma_X(\ell) - \sum_{i=1}^{p_0} \Lambda_i \Gamma_X(\ell - i) \right)^T \right\} = G_X(\ell) - \sum_{i=1}^{p_0} (\Lambda_i \otimes I_K) G_X(\ell - i). \quad \square$$

1.2. Existence of a m -th order stationary solution to model (1)

We first introduce a few notations. Let $M^{\otimes m} = M \otimes \dots \otimes M$ be the m -th tensor power of any matrix M , where $m \in \{0, 1, \dots\}$ (by convention $M^{\otimes 0} = I$ and $M^{\otimes 1} = M$). If $f(k)$ is a non random matrix, for $k \in \mathcal{E}$, set

$$\mathbb{P}(f) = \begin{pmatrix} p(1,1)f(1) & \cdots & p(d,1)f(1) \\ \vdots & & \vdots \\ p(1,d)f(d) & \cdots & p(d,d)f(d) \end{pmatrix}.$$

In particular we use $\mathbb{P} := \mathbb{P}(1) = (p(j, i))$ to denote the transpose of the transition matrix. Note that

$$\mathbb{P}(f) = \mathbb{P} \otimes f \tag{2}$$

when f is constant (*i.e.* $f(k) \equiv f, \forall k$).

Now let $A(k)$ be the $Kp \times Kp$ matrix

$$A(k) = \begin{pmatrix} a_1(k) & \cdots & a_p(k) \\ I_K & 0 & \cdots & 0 \\ 0 & I_K & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I_K & 0 \end{pmatrix} \tag{3}$$

and let the function $A^{\otimes m} : k \in \mathcal{E} \rightarrow A(k)^{\otimes m}$.

The following result gives a sufficient condition for the existence of a strictly stationary process (X_t) belonging to L^m and satisfying (1). We call nonanticipative any solution such that X_t is measurable with respect to the σ -field generated by $\{\eta_u, \Delta_u : u \leq t\}$.

Theorem 1. *Suppose that*

$$\rho\{\mathbb{P}(A^{\otimes m})\} < 1 \tag{4}$$

and $E\|\eta_t\|^m < \infty$. Then there exists a unique nonanticipative strictly stationary solution of (1) and this solution belongs to L^m .

Because the proof is quite similar to the one given in Francq and Zakoïan ([4], Th. 2) for the 2nd order stationarity, we omit it here. It is worth noting that the condition of this theorem is only constraining the AR part of the model. It is not always necessary, however, and it is possible to show, on particular cases, that the MA part can have importance for the stationarity issues (see Ex. 6 in Francq and Zakoïan [4]).

1.3. ARMA representation for $(X_t^{\otimes m})$

Assuming that the condition for the existence of a $2m$ -th order stationary and nonanticipative solution (X_t) is satisfied, we wish to show that the vector $(X_t^{\otimes m})$ admits an ARMA representation.

From (1), it is clear that the following first-order autoregressive-type representation holds

$$z_t = \omega_t + \Phi_t z_{t-1} \tag{5}$$

where $\omega_t = \omega(\Delta_t, \eta_t) = \underline{c}_t + \underline{\epsilon}_t = \underline{c}_t + \Sigma_t \eta_t$,

$$\underline{c}_t = \begin{pmatrix} c(\Delta_t) \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{K(p+q)}, \quad z_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \\ \epsilon_t \\ \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-q+1} \end{pmatrix} \in \mathbb{R}^{K(p+q)}, \quad \Sigma_t = \begin{pmatrix} \sigma(\Delta_t) \\ 0 \\ \vdots \\ 0 \\ \sigma(\Delta_t) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

is a $K(p+q) \times K$ matrix and

$$\Phi_t = \Phi(\Delta_t) = \begin{pmatrix} a_1(\Delta_t) & \dots & a_p(\Delta_t) & b_1(\Delta_t) & \dots & b_q(\Delta_t) \\ I_K & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & I_K & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & I_K & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & I_K & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & I_K & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & I_K & 0 \end{pmatrix}$$

is a $K(p+q) \times K(p+q)$ matrix. In this vector representation, we have implicitly assumed that $p \geq 1, q \geq 1$, without loss of generality because $a_p(\cdot)$ and $b_q(\cdot)$ can be equal to zero in (1).

We first establish the following basic result:

Lemma 2. *The following representation holds*

$$z_t^{\otimes m} = \sum_{i=0}^m \Psi_i^{(m)}(\Delta_t, \eta_t) z_{t-1}^{\otimes i} \tag{6}$$

where the $\Psi_i^{(m)}(\Delta_t, \eta_t)$'s are defined recursively by

$$\begin{aligned} \Psi_0^{(1)}(\Delta_t, \eta_t) &= \underline{c}_t + \Sigma_t \eta_t, & \Psi_1^{(1)}(\Delta_t, \eta_t) &= \Phi_t, & \text{and for } m > 0, \\ \Psi_i^{(m+1)}(\Delta_t, \eta_t) &= \underline{\omega}_t \otimes \Psi_i^{(m)}(\Delta_t, \eta_t) + \Phi_t \otimes \Psi_{i-1}^{(m)}(\Delta_t, \eta_t) \end{aligned} \tag{7}$$

with, by convention, $\Psi_i^{(m)}(\Delta_t, \eta_t) = 0$ when $i > m$ or $i < 0$, and $z_t^{\otimes 0} = \Psi_0^{(0)}(\Delta_t, \eta_t) = 1$.

Proof. Formula (6) for $m = 1$ is given by (5). Assuming that (6) holds for some $m \geq 1$ we have

$$\begin{aligned} z_t^{\otimes(m+1)} &= \sum_{i=0}^m (\underline{\omega}_t + \Phi_t z_{t-1}) \otimes \left\{ \Psi_i^{(m)}(\Delta_t, \eta_t) z_{t-1}^{\otimes i} \right\} \\ &= \sum_{i=0}^m \underline{\omega}_t \otimes \left\{ \Psi_i^{(m)}(\Delta_t, \eta_t) z_{t-1}^{\otimes i} \right\} + (\Phi_t \otimes \Psi_i^{(m)}(\Delta_t, \eta_t)) (z_{t-1} \otimes z_{t-1}^{\otimes i}) \\ &= \sum_{i=0}^{m+1} \left\{ \underline{\omega}_t \otimes \Psi_i^{(m)}(\Delta_t, \eta_t) \right\} z_{t-1}^{\otimes i} + \left\{ \Phi_t \otimes \Psi_{i-1}^{(m)}(\Delta_t, \eta_t) \right\} z_{t-1}^{\otimes i} \end{aligned}$$

and we get the stated result. □

If $X_t \in L^{2m}$, then we have, for $m' \leq m$ and $\ell \geq 1$

$$\begin{aligned} E[z_t^{\otimes m'} \otimes z_{t-\ell}^{\otimes m} | \Delta_t] &= k] \pi(k) \\ &= \sum_{i=0}^{m'} E \left[\left\{ \Psi_i^{(m')}(\Delta_t, \eta_t) z_{t-1}^{\otimes i} \right\} \otimes z_{t-\ell}^{\otimes m} | \Delta_t = k \right] \pi(k) \\ &= \sum_{i=0}^{m'} E \left[\left\{ \Psi_i^{(m')}(\Delta_t, \eta_t) \otimes I_{\{K(p+q)\}^m} \right\} (z_{t-1}^{\otimes i} \otimes z_{t-\ell}^{\otimes m}) | \Delta_t = k \right] \pi(k) \\ &= \sum_{i=0}^{m'} \sum_{j=1}^d E \left[\left\{ \Psi_i^{(m')}(\Delta_t, \eta_t) \otimes I_{\{K(p+q)\}^m} \right\} (z_{t-1}^{\otimes i} \otimes z_{t-\ell}^{\otimes m}) | \Delta_t = k, \Delta_{t-1} = j \right] p(j, k) \pi(j) \\ &= \sum_{i=0}^{m'} E \left[\Psi_i^{(m')}(k, \eta_t) \otimes I_{\{K(p+q)\}^m} \right] \sum_{j=1}^d E \left[(z_{t-1}^{\otimes i} \otimes z_{t-\ell}^{\otimes m}) | \Delta_t = k, \Delta_{t-1} = j \right] p(j, k) \pi(j) \\ &= \sum_{i=0}^{m'} \Upsilon_i^{(m', m)}(k) \sum_{j=1}^d E \left[z_{t-1}^{\otimes i} \otimes z_{t-\ell}^{\otimes m} | \Delta_{t-1} = j \right] p(j, k) \pi(j) \end{aligned} \tag{8}$$

where $\Upsilon_i^{(m', m)}(k) = E \left[\Psi_i^{(m')}(k, \eta_t) \otimes I_{\{K(p+q)\}^m} \right]$.

A few comments can be made. The first equality in (8) comes straightforwardly from Lemma 2. The second equality follows from a standard property of the Kronecker product (namely $(E \otimes B)(C \otimes D) = EC \otimes BD$ provided that the sizes of the matrices are compatible). The third equality is obtained by a standard argument

of the conditional expectation theory. The fourth equality comes from the independence between η_t and the process (Δ_t) . The last equality is the consequence of the Markov property of (Δ_t) and the independence between (η_t) and (Δ_t) : hence if X is any real random variable, measurable with respect to the σ -field generated by $\{\eta_u, \Delta_u : u \leq t-1\}$, we have $\text{pr}(X < x | \Delta_t = k, \Delta_{t-1} = j) = \text{pr}(X < x | \Delta_{t-1} = j)$, $\forall x$.

To write (8) in vector form we now introduce, for $m' \leq m$, the vectors

$$\mathcal{U}_\ell(m', m) := \left(\pi(1)E[\underline{z}_t^{\otimes m'} \otimes \underline{z}_{t-\ell}^{\otimes m} | \Delta_t = 1]^T, \dots, \pi(d)E[\underline{z}_t^{\otimes m'} \otimes \underline{z}_{t-\ell}^{\otimes m} | \Delta_t = d]^T \right)^T$$

and

$$\mathcal{W}_m(\ell) = (\mathcal{U}_\ell(m, m)^T, \dots, \mathcal{U}_\ell(0, m)^T)^T.$$

In view of (8), we then have for $\ell \geq 1$,

$$\begin{aligned} \mathcal{W}_m(\ell) &= \begin{pmatrix} \mathbb{P}(\Upsilon_m^{(m,m)}) & \mathbb{P}(\Upsilon_{m-1}^{(m,m)}) & \cdots & \mathbb{P}(\Upsilon_0^{(m,m)}) \\ 0 & \mathbb{P}(\Upsilon_{m-1}^{(m-1,m)}) & \cdots & \mathbb{P}(\Upsilon_0^{(m-1,m)}) \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 0 & \mathbb{P}(I_{\{K(p+q)\}^m}) \end{pmatrix} \mathcal{W}_m(\ell-1). \\ &:= \mathcal{A}_m \mathcal{W}_m(\ell-1) \end{aligned} \quad (9)$$

Now let λ be the (unique) $K^{2m} \times d \{K(p+q)\}^m [1 + K(p+q) + \cdots + \{K(p+q)\}^m]$ matrix such that

$$G_{X^{\otimes m}}(\ell) := E[X_t^{\otimes m} \otimes X_{t-\ell}^{\otimes m}] - \{E[X_t^{\otimes m}]\}^{\otimes 2} = \lambda \mathcal{W}_m(\ell).$$

Examples of such a matrix λ will be given in the next section. To the matrix \mathcal{A}_m corresponds the set of pairs of positive integers

$$\begin{aligned} \mathcal{N}(\mathcal{A}_m) &= \{(n, n') | \exists K^m \times K^m \text{ matrices } C_1, \dots, C_n, \lambda \mathcal{A}_m^{n+n'} + (C_1 \otimes I_{K^m}) \lambda \mathcal{A}_m^{n+n'-1} \\ &\quad + \cdots + (C_{n-1} \otimes I_{K^m}) \lambda \mathcal{A}_m^{n'+1} + (C_n \otimes I_{K^m}) \lambda \mathcal{A}_m^{n'} = 0\}. \end{aligned} \quad (10)$$

Let $\mathcal{P}_{\mathcal{A}_m}(x) = a_n + a_{n-1}x + \cdots + a_1x^{n-1} + x^n$ be an annihilating polynomial of \mathcal{A}_m . Then we see that $(n, 0) \in \mathcal{N}(\mathcal{A}_m)$ with $C_i = a_i I_{K^m}$. We define a (not necessarily unique) smallest element of $\mathcal{N}(\mathcal{A}_m)$, $(n(\mathcal{A}_m), n'(\mathcal{A}_m))$, as follows:

- i) $(n(\mathcal{A}_m), n'(\mathcal{A}_m)) \in \mathcal{N}(\mathcal{A}_m)$ and
- ii) $\forall (n, n') \in \mathcal{N}(\mathcal{A}_m)$, $(n, n') \neq (n(\mathcal{A}_m), n'(\mathcal{A}_m))$, $n > n(\mathcal{A}_m)$ or $n' > n'(\mathcal{A}_m)$.

In view of (9), we have for any $\ell > i-1$

$$G_{X^{\otimes m}}(\ell) = \lambda \mathcal{A}_m \mathcal{W}_m(\ell-1) = \lambda \mathcal{A}_m^i \mathcal{W}_m(\ell-i).$$

Hence we have, for $\ell > n(\mathcal{A}_m) + n'(\mathcal{A}_m) - 1$ and suitable $K^m \times K^m$ matrices $C_1, \dots, C_{n(\mathcal{A}_m)}$,

$$\begin{aligned} G_{X^{\otimes m}}(\ell) &+ (C_1 \otimes I_{K^m}) G_{X^{\otimes m}}(\ell-1) + \cdots + (C_{n(\mathcal{A}_m)} \otimes I_{K^m}) G_{X^{\otimes m}}(\ell - n(\mathcal{A}_m)) \\ &= \{\lambda \mathcal{A}_m^{n(\mathcal{A}_m)+n'(\mathcal{A}_m)} + (C_1 \otimes I_{K^m}) \lambda \mathcal{A}_m^{n(\mathcal{A}_m)+n'(\mathcal{A}_m)-1} + \cdots \\ &\quad + (C_{n(\mathcal{A}_m)} \otimes I_{K^m}) \lambda \mathcal{A}_m^{n'(\mathcal{A}_m)}\} \mathcal{W}_m(\ell - n(\mathcal{A}_m) - n'(\mathcal{A}_m)) = 0. \end{aligned}$$

From Lemma 1 we therefore have proved that the process $(X_t^{\otimes m})$ admits an ARMA representation of orders $n(\mathcal{A}_m)$ (autoregressive part) and $n(\mathcal{A}_m) + n'(\mathcal{A}_m) - 1$ (moving-average part).

We now may state the main result of this paper.

Theorem 2. *Let $m \in \{1, 2, \dots\}$ and assume that $\rho\{\mathbb{P}(A^{\otimes 2m})\} < 1$ and $E\|\eta_t\|^{2m} < \infty$. Let (X_t) be the nonanticipative strictly stationary solution of (1). Let \mathcal{A}_m be the matrix defined in (9) and let $(n(\mathcal{A}_m), n'(\mathcal{A}_m))$ be a smallest element of the set $\mathcal{N}(\mathcal{A}_m)$ defined in (10). Then the process $(X_t^{\otimes m})$ admits an ARMA($n(\mathcal{A}_m)$, $n(\mathcal{A}_m) + n'(\mathcal{A}_m) - 1$) representation.*

We now illustrate this general result on several important particular cases.

2. GENERIC EXAMPLES

Three cases of particular interest are the HMM, the Markov switching pure AR and the Markov switching pure MA. The first case has been studied in the seminal paper by Poskitt and Chung [9]. They have shown that HMM's satisfy ARMA representations, which can be used to solve the important problem of identifying the number of regimes. The second case has been considered by Zhang and Stine [12]. Theorem 2 encompasses and extends these previous results. The last case enjoys a remarkable property: under the assumption of a symmetric distribution for η_t , the odd powers of Markov switching pure MA's are simple MA's, of the same order, whereas the even powers have more complicated ARMA representations.

Example 1. THE HIDDEN MARKOV MODEL. Poskitt and Chung ([9], Th. 2.1) have shown that in the particular case when $K = 1$ and $p = q = 0$ in (1), X_t satisfies an ARMA($d - 1, d - 1$) model. To derive (9), we have implicitly assumed that $p \geq 1$ and $q \geq 1$. It is however easy to see that in the situation when $K = 1, p = q = 0$ and $m = 1$, setting $\underline{z}_t = X_t$, (9) holds with

$$\mathcal{A}_1 = \begin{pmatrix} 0 & \begin{pmatrix} c(1)p(1, 1) & \cdots & c(1)p(d, 1) \\ \vdots & & \vdots \\ c(d)p(1, d) & \cdots & c(d)p(d, d) \end{pmatrix} \\ 0 & \mathbb{P} \end{pmatrix}. \tag{11}$$

Note that $G_X(h) = \Gamma_X(h) = \lambda\mathcal{W}(\ell)$, where $\lambda = (e^T, -E(X_t)e^T)$ and $e = (1, \dots, 1)^T \in \mathbb{R}^d$. Let $\mathcal{P}_{\mathbb{P}}(x)$ be the characteristic polynomial of \mathbb{P} . This polynomial is of degree d . Because

$$\mathcal{P}_{\mathbb{P}}(\mathcal{A}_1) = \begin{pmatrix} \mathcal{P}_{\mathbb{P}}(0)I_d & \times \\ 0 & \mathcal{P}_{\mathbb{P}}(\mathbb{P}) \end{pmatrix} = \begin{pmatrix} \mathcal{P}_{\mathbb{P}}(0)I_d & \times \\ 0 & 0 \end{pmatrix},$$

we have $\mathcal{P}_{\mathbb{P}}(\mathcal{A}_1)\mathcal{A}_1 = \mathcal{A}_1\mathcal{P}_{\mathbb{P}}(\mathcal{A}_1) = 0$. Theorem 2 entails that X_t follows an ARMA(d, d) model. Since \mathbb{P} has an eigenvalue equal to 1, the AR polynomial has a unit root. Because (X_t) is stationary, the MA polynomial admits also a zero equal to one (see Brockwell and Davis [3], p. 86). The AR and MA parts can then be divided by the lag polynomial $1 - L$. As in Poskitt and Chung [9], we conclude that X_t follows an ARMA($d - 1, d - 1$) representation.

This result can be extended to the multivariate case and to powers of X_t . Consider the general case $K \geq 1$. Let m be a positive integer. Assume the existence of a m -th order stationary and non anticipative solution (X_t) of (1). Now we set $\underline{z}_t = X_t \in \mathbb{R}^K$ and

$$\mathcal{A}_m = \begin{pmatrix} 0 & \cdots & 0 & \mathbb{P} \left\{ E(c(\cdot) + \sigma(\cdot)\eta_t)^{\otimes m} \otimes I_{K^m} \right\} \\ \vdots & & & \\ 0 & \cdots & 0 & \mathbb{P}(c \otimes I_{K^m}) \\ 0 & \cdots & 0 & \mathbb{P}(I_{K^m}) \end{pmatrix}. \tag{12}$$

We have $G_{X^{\otimes m}}(\ell) = \lambda\mathcal{W}_m(\ell)$, where $\lambda = (e^T \otimes I_{K^{2m}}, 0, \dots, 0, -e^T \otimes E(X_t^{\otimes m}) \otimes I_{K^m})$. By (2), $\mathcal{P}_{\mathbb{P}}(\mathbb{P}(I_{K^m})) = \mathcal{P}_{\mathbb{P}}(\mathbb{P} \otimes I_{K^m}) = \{\mathcal{P}_{\mathbb{P}}(\mathbb{P})\} \otimes I_{K^m} = 0$. Arguing as in the case $m = 1$, we conclude that $X_t^{\otimes m}$ follows an

ARMA($d - 1, d - 1$) representation, whatever the values of m and K . Moreover the coefficients of the AR part are diagonal matrices with constant values on the diagonal. This implies that any linear combination of the components of $X_t^{\otimes m}$ also follows an ARMA($d - 1, d - 1$) representation. The result extends that of Zhang and Stine [12] obtained in the case $m = 1$. Note that the orders of the ARMA are not always minimal. Poskitt and Chung [9] have however shown that, under some regularity conditions, these orders are minimal in the case $K = m = 1$. The following corollary summarizes our results concerning the HMM.

Corollary 1. *Suppose that X_t satisfies model (1) with $p = q = 0$, and that the assumptions of Theorem 2 are satisfied. Let N be an integer and let F be a $N \times K^m$ real matrix. Then the process $(FX_t^{\otimes m})_t$ has an ARMA($d - 1, d - 1$) representation.*

Example 2. VAR(p) SWITCHING MODEL. Consider a model of the form (1) with $q = 0$:

$$X_t = c(\Delta_t) + \sum_{i=1}^p a_i(\Delta_t)X_{t-i} + \sigma(\Delta_t)\eta_t. \tag{13}$$

Zhang and Stine [12] considered the model (13) without intercept: $c \equiv 0$. In this case, they found that X_t admits an ARMA($d(Kp)^2, d(Kp)^2 - 1$) representation. In order to compare this result to ours, consider the case $m = 1$. We have $\underline{z}_t = (X_t^T, \dots, X_{t-p+1}^T)^T \in \mathbb{R}^{Kp}$ and

$$\mathcal{A}_1 = \begin{pmatrix} \mathbb{P}(\Phi(\cdot) \otimes I_{Kp}) & \mathbb{P}(\underline{c}(\cdot) \otimes I_{Kp}) \\ 0 & \mathbb{P}(I_{Kp}) \end{pmatrix}, \tag{14}$$

where $\Phi(k) = a_1(k)$ when $p = 1$, $\Phi(k) = A(k)$ defined by (3) when $p > 1$, and $\underline{c}(k) = (c(k)^T, 0)^T$ is a $pK \times 1$ vector. With arguments and notations already employed,

$$\mathcal{P}_{\mathbb{P}(\Phi(\cdot) \otimes I_{Kp})}(\mathcal{A}_1)\mathcal{P}_{\mathbb{P}(\Phi(\cdot) \otimes I_{Kp})}(\mathcal{A}_1) = \begin{pmatrix} 0 & \times \\ 0 & \times \end{pmatrix} \begin{pmatrix} \times & \times \\ 0 & 0 \end{pmatrix} = 0.$$

Since the degree of the polynomial $\mathcal{P}_{\mathbb{P}(\Phi(\cdot) \otimes I_{Kp})}(\mathcal{A}_1)\mathcal{P}_{\mathbb{P}(\Phi(\cdot) \otimes I_{Kp})}(\mathcal{A}_1)$ is less than $d + d(Kp)^2$, and 1 is an eigenvalue of \mathbb{P} , Theorem 2 shows that X_t follows an ARMA($d + d(Kp)^2 - 1, d + d(Kp)^2 - 2$). In the particular situation considered by Zhang and Stine [12], the process (X_t) is centred, so $\lambda = (\lambda_1, 0)$ where λ_1 is a $K^2 \times d(Kp)^2$ matrix and 0 denotes the $K^2 \times d(Kp)$ null matrix, and

$$\mathcal{A}_1 = \begin{pmatrix} \mathbb{P}(\Phi(\cdot) \otimes I_{Kp}) & 0 \\ 0 & \mathbb{P}(I_{Kp}) \end{pmatrix}. \tag{15}$$

Then we have $\lambda \mathcal{P}_{\mathbb{P}(\Phi(\cdot) \otimes I_{Kp})}(\mathcal{A}_1) = 0$. Therefore, in the particular case when $q = 0$ and $c \equiv 0$, Theorem 2 leads to the conclusion given by Zhang and Stine [12]. We have shown the following result:

Corollary 2. *Assume that $\rho\{\mathbb{P}(A^{\otimes 2})\} < 1$. Let (X_t) be the nonanticipative strictly stationary solution of (13). The process (X_t) has an ARMA($d + d(Kp)^2 - 1, d + d(Kp)^2 - 2$) representation. When $c \equiv 0$, (X_t) has an ARMA($d(Kp)^2, d(Kp)^2 - 1$) representation.*

Example 3. VMA(q) SWITCHING MODEL. Consider the pure MA Markov-switching model given by

$$X_t = \sum_{j=0}^q b_j(\Delta_t)\sigma(\Delta_{t-j})\eta_{t-j} \tag{16}$$

with $b_0(\Delta_t) = 1$. We will show the following result:

Corollary 3. *Let (X_t) be the Markov-switching moving average defined by (16), where η_t admits a symmetric law with moments up to the order $2m$, for some $m \in \{1, 2, \dots\}$. When m is odd, the process $(X_t^{\otimes m})$ admits a MA(q) representation. When m is even, the process $(X_t^{\otimes m})$ admits an ARMA($d-1, d-1+q$) representation.*

Proof. The first part of the corollary can be shown directly. Indeed, $X_t^{\otimes m}$ can be straightforwardly written under the form

$$X_t^{\otimes m} = \sum_{j_1, \dots, j_m=0}^q \psi_{j_1, \dots, j_m}(\Delta_t, \Delta_{t-j_1}, \dots, \Delta_{t-j_m}) \{\eta_{t-j_1} \otimes \dots \otimes \eta_{t-j_m}\}.$$

From the independence between $(\eta_t, \dots, \eta_{t-q})$ and $(\eta_{t-k}, \dots, \eta_{t-k-q})$, for $k > q$, and the independence between the processes (η_t) and (Δ_t) , conditioning with respect to $\Delta_t, \dots, \Delta_{t-k-q}$ yields $\text{Cov}(X_t^{\otimes m}, X_{t-k}^{\otimes m}) = E(X_t^{\otimes m} X_{t-k}^{\otimes m T}) = 0$ when m is odd. Hence $(X_t^{\otimes m})$ is a MA(q). The argument is no longer valid when m is even. In this case, a proof using Theorem 2 is given in the Appendix. □

Remark. It is worth noting that the ARMA representation, in the case when m is even, cannot (in general) be reduced to an MA(q). Consider for instance the simple example of a MA(1) with 2 regimes:

$$X_t = \eta_t + b1_{\{\Delta_t=2\}}\eta_{t-1},$$

with $\eta_t \sim \mathcal{N}(0, 1)$ and $b \neq 0$. For simplicity the first regime is chosen to be a white noise ($b(1) = 0$) and the second regime is a MA(1). The second-order structure of (X_t^2) is obtained as follows:

$$\begin{aligned} E(X_t^2) &= 1 + b^2\pi(2), \\ \text{Var}(X_t^2) &= 2 + b^4\pi(2)(3 - \pi(2)) + 4b^2\pi(2), \\ \text{Cov}(X_t^2, X_{t-1}^2) &= b^2\pi(2) \{2 + b^2p(2, 2) - b^2\pi(2)\}, \\ \text{Cov}(X_t^2, X_{t-\ell}^2) &= b^4\pi(2) \{p(2, 2) - \pi(2)\} \{p(2, 2) - p(1, 2)\}^{\ell-1}, \quad \text{for } \ell > 1. \end{aligned}$$

Recall that the assumptions made on (Δ_t) imply that $\pi(2) \neq 0$. Thus, unless when $p(2, 2) = p(1, 2)$ (or equivalently $p(2, 2) = \pi(2)$), the process (X_t^2) is not a pure MA. More precisely (X_t^2) admits an ARMA(1,2) representation (in accordance with Cor. 3) of the form

$$X_t^2 - 1 - b^2\pi(2) = \{p(2, 2) - p(1, 2)\} \{X_{t-1}^2 - 1 - b^2\pi(2)\} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$$

where θ_1, θ_2 are constants.

3. CONCLUSION

In this paper we have established the existence of ARMA representations for any power of Markov-switching ARMA processes. The orders of the ARMA representations are directly linked with the number of regimes of the unobserved Markov chain. The results highlight the interest of considering different powers of X_t for statistical purposes. For instance, in the particular case of a Markov switching MA(q), the identification of q can be achieved, whatever the number of regimes, by simply considering the autocorrelation functions of odd powers of X_t . On the other hand, the autocorrelation functions of even powers of X_t are informative concerning both the number of regimes d and the order q of the MA regimes. Of course, the statistical analysis based on these considerations is far beyond the scope of this paper, and is left for future research.

APPENDIX: PROOF OF COROLLARY 3

We set $\underline{z}_t = (X_t^T, \epsilon_t^T, \dots, \epsilon_{t-q+1}^T)^T \in \mathbb{R}^{K(q+1)}$. In view of (7) it is straightforward to show, by induction, that

$$E\Psi_i^{(m')}(\cdot, \eta_t) = 0 \quad \text{when } i + m' \text{ is odd.} \tag{17}$$

More precisely, $\Psi_i^{(m')}(\cdot, \eta_t)$ is the following sum of $\binom{m'}{i}$ terms:

$$\Psi_i^{(m')}(\Delta_t, \eta_t) = \underline{\omega}_t^{\otimes(m'-i)} \otimes \Phi(\Delta_t)^{\otimes i} + \underline{\omega}_t \otimes \Phi(\Delta_t) \otimes \underline{\omega}_t^{\otimes(m'-i-1)} \otimes \Phi(\Delta_t)^{\otimes(i-1)} + \dots + \Phi(\Delta_t)^{\otimes i} \otimes \underline{\omega}_t^{\otimes(m'-i)}. \tag{18}$$

In view of (17), (9) holds with

$$\mathcal{A}_m = \begin{pmatrix} \mathbb{P}(\Upsilon_m^{(m,m)}) & 0 & \mathbb{P}(\Upsilon_{m-2}^{(m,m)}) & \mathbb{P}(\Upsilon_0^{(m,m)}) \\ 0 & \mathbb{P}(\Upsilon_{m-1}^{(m-1,m)}) & 0 & \mathbb{P}(\Upsilon_0^{(m-1,m)}) \\ \vdots & & \ddots & \ddots \\ 0 & & \dots & \mathbb{P}(\Upsilon_1^{(1,m)}) \\ 0 & & \dots & 0 & \mathbb{P}(I_{\{K(q+1)\}}^m) \end{pmatrix}.$$

With obvious notations, the matrix \mathcal{A}_m can be partitioned into blocks $\mathcal{A}_m(i, j)$ defined by

$$\mathcal{A}_m(i, j) = \begin{cases} 0 & \text{when } i > j \text{ or } i + j \text{ odd} \\ \mathbb{P}(\Upsilon_{m-j+1}^{(m-i+1,m)}) & \text{otherwise,} \end{cases}$$

for $i, j = 1, \dots, m + 1$. Therefore the matrix \mathcal{A}_m^2 can be similarly partitioned into blocks defined by

$$\mathcal{A}_m^2(i, j) = \sum_{k=1}^{m+1} \mathcal{A}_m(i, k)\mathcal{A}_m(k, j) = \begin{cases} 0 & \text{when } i > j \text{ or } i + j \text{ odd} \\ \sum_{k \in \mathcal{S}_{i,j}} \mathbb{P}(\Upsilon_{m-k+1}^{(m-i+1,m)})\mathbb{P}(\Upsilon_{m-j+1}^{(m-k+1,m)}) & \text{otherwise,} \end{cases}$$

where $\mathcal{S}_{i,j}$ denotes the set of indices $\mathcal{S}_{i,j} = \{i, i + 2, \dots, j\}$. Similarly, the blocks of \mathcal{A}_m^{q+1} are given by

$$\mathcal{A}_m^{q+1}(i, j) = \sum_{k_1 \in \mathcal{S}_{i,j}} \sum_{k_2 \in \mathcal{S}_{k_1,j}} \dots \sum_{k_q \in \mathcal{S}_{k_{q-1},j}} \mathbb{P}(\Upsilon_{m-k_1+1}^{(m-i+1,m)})\mathbb{P}(\Upsilon_{m-k_2+1}^{(m-k_1+1,m)}) \dots \mathbb{P}(\Upsilon_{m-j+1}^{(m-k_{q-1}+1,m)}) \tag{19}$$

when $i \leq j$ and $i + j$ even. Now note that, because

$$\Phi(k) = \begin{pmatrix} 0 & b_1(k) & \dots & b_q(k) \\ 0 & 0 & \dots & 0 \\ 0 & I_K & \ddots & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & I_K & 0 \end{pmatrix},$$

we have

$$\prod_{\ell=1}^{q+1} \Phi(k_\ell) = 0, \quad \forall k_1, \dots, k_{q+1} \in \{1, \dots, d\}. \tag{20}$$

Each of the blocks defined by (19) can themselves be partitioned into blocks $[\mathcal{A}_m^{q+1}(i, j)](i', j')$ defined by

$$[\mathcal{A}_m^{q+1}(i, j)](i', j') = \sum_{k_1 \in \mathcal{S}_{i,j}} \sum_{k_2 \in \mathcal{S}_{k_1,j}} \cdots \sum_{k_q \in \mathcal{S}_{k_{q-1},j}} \sum_{k'_1=1}^d \cdots \sum_{k'_q=1}^d \tag{21}$$

$$p(k'_1, i') \Upsilon_{m-k_1+1}^{(m-i+1,m)}(i') p(k'_2, k'_1) \Upsilon_{m-k_2+1}^{(m-k_1+1,m)}(k'_1) \cdots p(j', k'_q) \Upsilon_{m-j+1}^{(m-k_{q-1}+1,m)}(k'_q)$$

for $i', j' = 1, \dots, d$. For a diagonal block such that $i = j < m + 1, k_1 = k_2 = \cdots = k_q = i$ and, using (20) and $\Psi_i^{(i)}(\cdot, \eta_i) = \Phi(\cdot)^{\otimes i}$,

$$\begin{aligned} & \Upsilon_{m-k_1+1}^{(m-i+1,m)}(i') \Upsilon_{m-k_2+1}^{(m-k_1+1,m)}(k'_1) \cdots \Upsilon_{m-j+1}^{(m-k_{q-1}+1,m)}(k'_q) \\ &= \Upsilon_{m-i+1}^{(m-i+1,m)}(i') \prod_{\ell=1}^q \Upsilon_{m-i+1}^{(m-i+1,m)}(k'_\ell) \\ &= \left\{ \Phi(i')^{\otimes(m-i+1)} \otimes I_{\{K(q+1)\}^m} \right\} \prod_{\ell=1}^q \left\{ \Phi(k'_\ell)^{\otimes(m-i+1)} \otimes I_{\{K(q+1)\}^m} \right\} \\ &= \left\{ \Phi(i') \prod_{\ell=1}^q \Phi(k'_\ell) \right\}^{\otimes(m-i+1)} \otimes I_{\{K(q+1)\}^m} = 0. \end{aligned}$$

We have shown that $[\mathcal{A}_m^{q+1}(i, i)] = 0$ for $i < m + 1$. Now, let us consider the term $[\mathcal{A}_m^{q+1}(1, 3)]$ for $m > 2$. In (21), consider for instance the term which corresponds to $k_1 = k_2 = 1$ and $k_3 = \cdots = k_q = 3$. In this situation, we have, using (18) and (20),

$$\begin{aligned} & \Upsilon_{m-k_1+1}^{(m-i+1,m)}(i') \Upsilon_{m-k_2+1}^{(m-k_1+1,m)}(k'_1) \cdots \Upsilon_{m-j+1}^{(m-k_{q-1}+1,m)}(k'_q) \\ &= \Upsilon_m^{(m,m)}(i') \Upsilon_m^{(m,m)}(k'_1) \Upsilon_{m-2}^{(m,m)}(k'_2) \prod_{\ell=3}^q \Upsilon_{m-2}^{(m-2,m)}(k'_\ell) \\ &= \left\{ \Phi(i')^{\otimes m} \otimes I_{\{K(q+1)\}^m} \right\} \left\{ \Phi(k'_1)^{\otimes m} \otimes I_{\{K(q+1)\}^m} \right\} \\ & \quad \times \left\{ E \left(\underline{\omega}_t^{\otimes 2} \otimes \Phi(k'_2)^{\otimes(m-2)} + \underline{\omega}_t \otimes \Phi(k'_2) \otimes \underline{\omega}_t \otimes \Phi(k'_2)^{\otimes(m-3)} + \cdots + \Phi(k'_2)^{\otimes(m-2)} \otimes \underline{\omega}_t^{\otimes 2} \right) \otimes I_{\{K(q+1)\}^m} \right\} \\ & \quad \times \prod_{\ell=3}^q \left\{ \Phi(k'_\ell)^{\otimes(m-2)} \otimes I_{\{K(q+1)\}^m} \right\} \\ &= E \left[\left\{ \Phi(i') \Phi(k'_1) \underline{\omega}_t \right\}^{\otimes 2} \otimes \left\{ \Phi(i') \prod_{\ell=1}^q \Phi(k'_\ell) \right\}^{\otimes(m-2)} \otimes I_{\{K(q+1)\}^m} \right. \\ & \quad \left. + \left\{ \Phi(i') \Phi(k'_1) \underline{\omega}_t \right\} \otimes \left\{ \Phi(i') \prod_{\ell=1}^q \Phi(k'_\ell) \right\} \otimes \left\{ \Phi(i')^{\otimes(m-2)} \Phi(k'_1)^{\otimes(m-2)} \left(\underline{\omega}_t \otimes \Phi(k'_2)^{\otimes(m-3)} \right) \prod_{\ell=3}^q \Phi(k'_\ell)^{\otimes(m-3)} \right\} \right. \\ & \quad \left. \otimes I_{\{K(q+1)\}^m} + \cdots + \left\{ \Phi(i') \prod_{\ell=1}^q \Phi(k'_\ell) \right\}^{\otimes(m-2)} \otimes \left\{ \Phi(i') \Phi(k'_1) \underline{\omega}_t \right\}^{\otimes 2} \otimes I_{\{K(q+1)\}^m} \right] \\ &= 0. \end{aligned}$$

Continuing this way, we show that $[\mathcal{A}_m^{q+1}(1, 3)] = 0$ when $m > 2$. By the same arguments, $[\mathcal{A}_m^{q+1}(i, j)] = 0$ for all $j < m + 1$.

First consider the case when m is odd. We already saw in Section 2 that $(X_t^{\otimes m})$ admits a weak MA(q) representation. This is also obtained by Theorem 2 because $\mathcal{A}_m^{q+1}(1, \cdot) = 0$ and $\lambda \mathcal{A}_m^{q+1} = 0$.

Now consider the case when m is even. The previous argument does not apply because $\mathcal{A}_m^{q+1}(1, m+1) \neq 0$. Let $\mathcal{P}_{\mathbb{P}}$ be any annihilating polynomial of \mathbb{P} . Using (2), we have $\mathcal{P}_{\mathbb{P}}\{\mathbb{P}(I_{\{K(q+1)\}^m})\} = 0$. Therefore we have

$$\mathcal{A}_m^{q+1} = \begin{pmatrix} 0 & 0 & & \times \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \{\mathbb{P}(I_{\{K(q+1)\}^m})\}^{q+1} \end{pmatrix} \text{ and } \mathcal{P}_{\mathbb{P}}(\mathcal{A}_m) = \begin{pmatrix} \times & 0 & \times & \times \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & \times & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

It follows that $\mathcal{P}_{\mathbb{P}}(\mathcal{A}_m)\mathcal{A}_m^{q+1} = \mathcal{A}_m^{q+1}\mathcal{P}_{\mathbb{P}}(\mathcal{A}_m) = 0$. Then, choosing the characteristic polynomial of \mathbb{P} for $\mathcal{P}_{\mathbb{P}}(\cdot)$, Theorem 2 shows that, when m is even, $(X_t^{\otimes m})$ is an ARMA($d, d+q$). Because one of the roots of $\mathcal{P}_{\mathbb{P}}(x) = 0$ is equal to one, an argument already given shows that $(X_t^{\otimes m})$ is an ARMA($d-1, d-1+q$).

REFERENCES

- [1] L.E. Baum and T. Petrie, Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statist.* **30** (1966) 1554-1563.
- [2] A. Berline, *Estimation des degrés d'un ARMA multivarié*, Pub. IRMA, Vol. 4. Lille (1982).
- [3] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*. Springer-Verlag, New York (1991).
- [4] C. Francq and J.-M. Zakoïan, Stationarity of Multivariate Markov-switching ARMA Models. *J. Econometrics* **102** (2001) 339-364.
- [5] J.D. Hamilton, A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57** (1989) 357-384.
- [6] J.D. Hamilton, Specification testing in Markov switching time series models. *J. Econometrics* **45** (1996) 39-70.
- [7] H. Karlsen, *A class of non-linear time series models*, Ph.D. Thesis. University of Bergen, Norway (1990).
- [8] B.G. Leroux and L.M. Puterman, Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48** (1992) 545-558.
- [9] D.S. Poskitt and S.H. Chung, Markov chain models, time series analysis and extreme value theory. *Adv. Appl. Probab.* **28** (1996) 405-425.
- [10] C.P. Robert, T. Rydén and D.M. Titterton, Bayesian inference in hidden Markov models through the reversible jump Markov Chain Monte-Carlo method. *J. Roy. Statist. Soc. B* **62** (2000) 57-75.
- [11] T. Rydén, Estimating the orders of hidden Markov models. *Statistics* **26** (1995) 345-354.
- [12] J. Zhang and R.A. Stine, Autocovariance structure of Markov regime switching models and model selection. *J. Time Ser. Anal.* **22** (2001) 107-124.