

DOUBLY STOCHASTIC HILBERTIAN PROCESSES

SERGE GUILLAS,* *Université Pierre et Marie Curie—Paris VI
and École des Mines de Douai*

Abstract

In this paper, we consider a Hilbert-space-valued autoregressive stochastic sequence (X_n) with several regimes. We suppose that the underlying process (I_n) which drives the evolution of (X_n) is stationary. Under some dependence assumptions on (I_n) , we prove the existence of a unique stationary solution, and with a symmetric compact autocorrelation operator, we can state a law of large numbers with rates and the consistency of the covariance estimator. An overall hypothesis states that the regimes where the autocorrelation operator's norm is greater than 1 should be rarely visited.

Keywords: Hilbert; nonlinear; autoregressive; stability

AMS 2000 Subject Classification: Primary 60G10

Secondary 62M10

1. Introduction

Autoregressive Hilbertian processes of order one (ARH(1)) have been extensively investigated by Bosq [4], [5], mainly because of the interest in prediction of continuous-time stochastic processes (e.g. electricity consumption [9], traffic [2], or climatic variation [3]). The inclusion of exogenous variables in an additive way may improve the model; this has recently been considered theoretically [13] and applied to pollution forecasting [10]. However, in the case of real-valued time-series, several authors also examined nonadditive inclusion. For instance, we can mention regime models (according to Tjøsteim's terminology [21]) of the form

$$X_t = \theta_t X_t + e_t, \quad (1)$$

where (e_t) is a noise and (θ_t) is a stochastic process. When θ_t is regulated by X_{t-1} , (1) results in '(self-exciting) threshold autoregressive processes' (SETAR or just TAR). On the other hand, if (θ_t) is a Markov chain with a finite state space, we say that the model is doubly stochastic [20]. Random coefficient autoregressive models (RCA) [16] appear in the particular case where (θ_t) is i.i.d.

Tong [22], [23] studied TAR processes and an application to the prediction of ozone can be found in [15]. These nonlinear processes should fit better to phenomena with sudden changes such as river flows.

Petrucci and Woolford [17] considered the real-valued model

$$Z_t = \phi_1 Z_{t-1}^+ + \phi_2 Z_{t-1}^- + \varepsilon_t, \quad t = 1, 2, \dots,$$

where $x^+ = \max(x, 0)$, $x^- = \min(x, 0)$ and (ε_t) is a white noise. They showed that (Z_t) is ergodic if and only if

$$\phi_1 < 1, \quad \phi_2 < 1, \quad \phi_1 \phi_2 < 1,$$

Received 5 February 2002; revision received 20 June 2002.

* Current address: The University of Chicago, Center for Integrating Statistical and Environmental Science, Chicago, IL 60637, USA. Email address: guillas@galton.uchicago.edu

which is less restrictive than

$$|\phi_1| < 1, \quad |\phi_2| < 1.$$

Similarly, Brandt [8] and Bougerol and Picard [7] made use of the negativity of the top Lyapunov exponent in order to ensure the existence of a stationary solution, but Yao [24] pointed out that this criterion is difficult to apply, proposing to look at the spectral radius of the autocorrelation matrix in a multivariate framework. See also [18] in this context, [12] for extensions to the multivariate ARMA case, [25] for an investigation of the nonlinear case, and [14] for nonstationary (θ_t) and (e_t) .

From our ARH(1) point of view, the model imposes conditions of the form (see [5])

$$\|\rho^j\|_{\mathcal{L}} < 1 \quad \text{for some } j$$

for the existence of a stationary solution to the equation

$$X_n = \rho(X_{n-1}) + \varepsilon_n,$$

where H is a real and separable Hilbert space—usually an infinite-dimensional one—with norm $\|\cdot\|$, ρ is a bounded linear operator on H , $\|\cdot\|_{\mathcal{L}}$ is the linear norm of linear operators on H , and (ε_n) is a strong Hilbertian white noise (SHWN), that is, a sequence of i.i.d. random variables with values in H satisfying

$$E\varepsilon_n = 0, \quad 0 < E\|\varepsilon_n\|^2 = \sigma^2 < \infty \quad \text{for all } n \in \mathbb{Z}.$$

Let $(I_n)_{n \in \mathbb{Z}}$ be a strictly stationary sequence of random variables taking their values in $\{0, 1\}$, and (ε_n) be a SHWN. Let ρ_0 and ρ_1 be two bounded linear operators on H . We consider the following model:

$$X_n = \begin{cases} \rho_0(X_{n-1}) + \varepsilon_n & \text{if } I_n = 0, \\ \rho_1(X_{n-1}) + \varepsilon_n & \text{if } I_n = 1, \end{cases} \tag{2}$$

which can be also written as

$$X_n = \rho_{I_n}(X_{n-1}) + \varepsilon_n.$$

We will also assume the following hypothesis:

$$(I_n) \text{ and } (\varepsilon_n) \text{ are independent.}$$

This assumption tells us that the two influences (the index of the operator and the additive noise) are not linked. We may talk of a double noise when (I_n) is also a noise.

In this paper, we extend doubly stochastic models to Hilbert-space-valued processes. When one of the two choices leads to an explosive case, for instance if $\|\rho_j^j\| \geq 1$ for all j , the alternative choice may stabilize the behaviour as we will see hereafter. Note that with this kind of sample paths, it seemed reasonable to study extremes for such processes (see [11] in the particular case of RCA models). In Section 2, we establish the existence and uniqueness of a stationary solution under the assumption that (I_n) is $(m - 1)$ -dependent for a positive integer m . The use of m -dependence, instead of more classical notions relying on the decay rate of mixing coefficients for instance, is motivated by purely technical reasons. In Section 3, we state a law of large numbers with rates, under the assumption that ρ_{I_n} is symmetric and compact and (I_n) is i.i.d. (RCA models). Finally, in Section 4, we show the consistency of an estimator of the covariance operator. Note that, for simplicity, our paper is devoted to the two regimes case, but it is easy to generalize to the multiple regimes case.

2. Stationary solution

We start with two technical lemmas. First, let us recall the following result of Bass [1] (see also [19]).

Lemma 1. (Bass [1].) *Let Y_1, \dots, Y_n be nonnegative random variables. Then*

$$E(Y_1 \cdots Y_n) \leq \int_0^1 Q_{Y_1}(u) \cdots Q_{Y_n}(u) \, du,$$

where Q_X denotes the quantile function of the random variable $|X|$ defined by

$$Q_X(u) = \inf\{t, P(|X| > t) \leq u\}.$$

We then infer the following upper bound from Lemma 1.

Lemma 2. *Let a and b be two positive real numbers. Let Y_1, \dots, Y_n be identically distributed random variables such that*

$$P(Y_1 = a) = 1 - q, \quad P(Y_1 = b) = q.$$

Then

$$E(Y_1 \cdots Y_n) \leq (1 - q)a^n + qb^n.$$

Proof. Apply Lemma 1:

$$E[Y_1 \cdots Y_n] \leq \int_0^1 (Q_Y(u))^n \, du,$$

with $Q_Y(u) := Q_{Y_1}(u)$. If $b \leq a$, then

$$P(Y_1 > t) = \begin{cases} 0 & \text{if } t \geq a, \\ 1 - q & \text{if } t \in [b, a[, \\ 1 & \text{otherwise,} \end{cases}$$

and, if $a \leq b$, then

$$P(Y_1 > t) = \begin{cases} 0 & \text{if } t \geq b, \\ q & \text{if } t \in [a, b[, \\ 1 & \text{otherwise.} \end{cases}$$

Hence, in the case where $b \leq a$,

$$Q_Y(u) = \begin{cases} a & \text{if } 0 \leq u < 1 - q, \\ b & \text{if } 1 - q \leq u < 1, \end{cases}$$

and, in the case where $a \leq b$,

$$Q_Y(u) = \begin{cases} b & \text{if } 0 \leq u < q, \\ a & \text{if } q \leq u < 1. \end{cases}$$

Thus, in general,

$$\int_0^1 (Q_Y(u))^n \, du = (1 - q)a^n + qb^n.$$

We will make use of the following quantities which evaluate the probability of visiting the possibly explosive regime. Let us set

$$q = P(I_n = 1),$$

and

$$c_k = (1 - q)\|\rho_0\|^k + q\|\rho_1\|^k.$$

Remark 1. The sequence (c_k) is not monotone in general. For instance, take $q = 0.5$, $\|\rho_0\| = 0.5$, and $\|\rho_1\| = 1.1$. We can see that $c_1 = 0.8$, $c_2 = 0.73$, and (c_k) tends to infinity as k goes to infinity. However, if $\|\rho_0\| \leq 1$ and $\|\rho_1\| \leq 1$, then $c_k \leq c_{k'}$ for $0 < k' < k$.

For simplicity, we introduce the following symbols for lengthily repeated products:

$$\pi_{n,j} = \rho_{I_n} \cdots \rho_{I_{n-j+1}} = \prod_{p=0}^{j-1} \rho_{I_{n-p}},$$

with the convention that

$$\pi_{n,0} = \text{Id}.$$

Theorem 1. Suppose that (I_n) is $(m - 1)$ -dependent. If

$$c_{4m} < 1,$$

then (2) has a stationary solution given by

$$X_n = \sum_{j=0}^{\infty} \pi_{n,j}(\varepsilon_{n-j}), \tag{3}$$

and the series converges in $L^2_H(\mathbb{P})$ (that is, in the mean square sense). Under the conditions

$$E\|X_0\|^4 < \infty \tag{4}$$

and

$$c_{8m} < 1, \tag{5}$$

this solution is unique.

Proof. For the existence:

$$\Delta_{l'} := \left\| \sum_{j=l}^{l'} \pi_{n,j}(\varepsilon_{n-j}) \right\|_{L^2_H(\mathbb{P})}^2 = \sum_{j=l}^{l'} \|\pi_{n,j}(\varepsilon_{n-j})\|_{L^2_H(\mathbb{P})}^2$$

because, for $j \neq j'$,

$$E\langle \pi_{n,j}(\varepsilon_{n-j}), \pi_{n,j'}(\varepsilon_{n-j'}) \rangle = E\langle \varepsilon_{n-j}, \pi_{n,j}^* \pi_{n,j'}(\varepsilon_{n-j'}) \rangle = 0$$

since ε_{n-j} and $\pi_{n,j}^* \pi_{n,j'}(\varepsilon_{n-j'})$ are independent.

Therefore,

$$\begin{aligned} \Delta_l^{l'} &= \left| \sum_{j=l}^{l'} \mathbb{E} \langle \pi_{n,j}(\varepsilon_{n-j}), \pi_{n,j}(\varepsilon_{n-j}) \rangle \right| \\ &\leq \sum_{j=l}^{l'} \mathbb{E} [\|\pi_{n,j}\|_{\mathcal{L}}^2 \|\varepsilon_{n-j}\|_H^2] \\ &\leq \sum_{j=l}^{l'} \mathbb{E} \left[\prod_{p=0}^{j-1} \|\rho_{I_{n-p}}\|_{\mathcal{L}}^2 \right] \mathbb{E} [\|\varepsilon_{n-j}\|_H^2]. \end{aligned} \tag{6}$$

It remains to find an upper bound for $\mathbb{E}[\prod_{p=0}^{j-1} \|\rho_{I_{n-p}}\|_{\mathcal{L}}^2]$. For that purpose, note that the random variables $Y_p = \|\rho_{I_{n-p}}\|_{\mathcal{L}}^2$ are identically distributed, and use the $(m - 1)$ -dependence of the sequence (I_n) and then of the sequence (Y_n) . We may now write

$$Y_1 \cdots Y_j = Y_1 \cdots Y_m Y_{m+1} \cdots Y_{2m} Y_{2m+1} \cdots Y_{km} Y_{km+1} \cdots Y_{km+r},$$

where $j = km + r$ is the expression of the Euclidean division of j by m ($0 \leq r < m$). Let us denote

$$\begin{aligned} U_1 &= Y_1 \cdots Y_m, \\ V_1 &= Y_{m+1} \cdots Y_{2m}, \\ U_2 &= Y_{2m+1} \cdots Y_{3m}, \\ &\vdots \end{aligned}$$

so that

$$\begin{aligned} Y_1 \cdots Y_m Y_{m+1} \cdots Y_{2m} Y_{2m+1} \cdots Y_{km} \cdots Y_{km+r} \\ = \begin{cases} U_1 V_1 U_2 \cdots U_{\lfloor k/2 \rfloor + 1} Y_{km+1} \cdots Y_{km+r} & \text{if } k \text{ is odd,} \\ U_1 V_1 U_2 \cdots V_{k/2} Y_{km+1} \cdots Y_{km+r} & \text{if } k \text{ is even.} \end{cases} \end{aligned}$$

Hence, for odd k , and denoting $V_{\lfloor k/2 \rfloor + 1} = Y_{km+1} \cdots Y_{km+r}$,

$$\begin{aligned} \mathbb{E} \left[\prod_{p=0}^{j-1} Y_p \right] &\leq \left(\mathbb{E} \left[\prod_{i=0}^{\lfloor k/2 \rfloor + 1} U_i^2 \right] \right)^{1/2} \left(\mathbb{E} \left[\prod_{i=0}^{\lfloor k/2 \rfloor + 1} V_i^2 \right] \right)^{1/2} \\ &\leq [((1 - q)\|\rho_0\|^{4m} + q\|\rho_1\|^{4m})^{\lfloor k/2 \rfloor + 1}]^{1/2} \\ &\quad \times [((1 - q)\|\rho_0\|^{4m} + q\|\rho_1\|^{4m})^{\lfloor k/2 \rfloor}]^{1/2} \\ &\quad \times [(1 - q)\|\rho_0\|^{4r} + q\|\rho_1\|^{4r}]^{1/2} \end{aligned}$$

by Lemma 2. Accordingly, for odd k , and since $c_{4m} < 1$,

$$\mathbb{E} \left[\prod_{p=0}^{j-1} Y_p \right] \leq [(1 - q)\|\rho_0\|^{4r} + q\|\rho_1\|^{4r}]^{1/2} [(1 - q)\|\rho_0\|^{4m} + q\|\rho_1\|^{4m}]^{\lfloor k/2 \rfloor},$$

an upper bound which holds also for even k by a similar argument. If we set

$$M_m = \sup_{r=0, \dots, m-1} [(1 - q)\|\rho_0\|^{4r} + q\|\rho_1\|^{4r}]^{1/2},$$

then we obtain

$$E \left[\prod_{p=0}^{j-1} Y_p \right] \leq M_m c_{4m}^{\lfloor k/2 \rfloor},$$

and

$$E \left[\prod_{p=0}^{j-1} \|\rho_{I_{n-p}}\|_{\mathcal{L}}^2 \right] \leq M_m c_{4m}^{\lfloor k/2 \rfloor}.$$

Since

$$j = km + r, \quad k = \frac{j - r}{m} > \frac{j - m}{m} = \frac{j}{m} - 1, \quad c_{4m} < 1,$$

we get

$$\left(E \left[\prod_{p=0}^{j-1} \|\rho_{I_{n-p}}\|_{\mathcal{L}}^2 \right] \right)^{1/2} \leq KM_m [(c_{4m})^{1/2m}]^j, \tag{7}$$

with a constant $K > 0$. Consequently, the upper bound is the general term of a convergent series. Thus, $\Delta_l^{l'}$ tends to 0 when l, l' tend to infinity and, by the Cauchy criterion, the series in (3) converges in $L^2_H(\mathbb{P})$.

Consider the stationary process

$$W_n = \sum_{j=0}^{\infty} \pi_{n,j}(\varepsilon_{n-j}).$$

Since ρ_0 and ρ_1 are bounded,

$$\begin{aligned} W_n - \rho_{I_n}(W_{n-1}) &= \sum_{j=0}^{\infty} \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) - \sum_{j=0}^{\infty} \rho_{I_n} \left(\prod_{p=0}^{j-1} \rho_{I_{n-1-p}} \right) (\varepsilon_{n-1-j}) \\ &= \sum_{j=0}^{\infty} \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) - \sum_{j=0}^{\infty} \left(\prod_{p=0}^j \rho_{I_{n-p}} \right) (\varepsilon_{n-1-j}) \\ &= \sum_{j=0}^{\infty} \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) - \sum_{j'=1}^{\infty} \left(\prod_{p=0}^{j'-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j'}) \\ &= \varepsilon_n. \end{aligned}$$

Conversely, let (X_n) be a stationary solution of (2) satisfying

$$E\|X_0\|^4 < \infty,$$

then by a straightforward induction, it follows that

$$X_n = \left(\prod_{p=0}^k \rho_{I_{n-p}} \right) (X_{n-k-1}) + \sum_{j=0}^k \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) \tag{8}$$

and

$$\begin{aligned} \mathbb{E} \left\| X_n - \sum_{j=0}^k \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) \right\|^2 &\leq \mathbb{E} \left\| \left(\prod_{p=0}^k \rho_{I_{n-p}} \right) (X_{n-k-1}) \right\|^2 \\ &\leq \mathbb{E} \left\| \prod_{p=0}^k \rho_{I_{n-p}} \right\|_{\mathcal{L}}^2 \|X_{n-k-1}\|^2 \\ &\leq \left(\mathbb{E} \left\| \prod_{p=0}^k \rho_{I_{n-p}} \right\|_{\mathcal{L}}^4 \right)^{1/2} (\mathbb{E} \|X_{n-k-1}\|^4)^{1/2}. \end{aligned}$$

The term $(\mathbb{E} \|X_{n-k-1}\|^4)^{1/2}$ is constant because of the stationarity of (X_n) . We can obtain a similar bound as in (7) under the assumption (5). Those points lead to the conclusion:

$$\mathbb{E} \left\| X_n - \sum_{j=0}^k \left(\prod_{p=0}^{j-1} \rho_{I_{n-p}} \right) (\varepsilon_{n-j}) \right\|^2 \rightarrow 0$$

when $k \rightarrow \infty$.

Let us note that using the Cauchy–Schwarz inequality in (6) and not independence, the assumptions ensuring the existence and uniqueness of a stationary solution would be (4), (5) and

$$\mathbb{E} \|\varepsilon_0\|^4 < \infty.$$

For the strong stationarity case where no second order moment is required, we can mimic the proof of Theorem 1 of [8], replacing absolute values by norms. We obtain the following result.

Theorem 2. *If the sequence (I_n, ε_n) is stationary and ergodic, if*

$$-\infty \leq \mathbb{E} \log \|\rho_{I_0}\| < 0,$$

and if

$$\mathbb{E} \log(\|\varepsilon_0\|)^+ < \infty,$$

where $x^+ = \max(0, x)$, then (2) has a unique strictly stationary solution given by

$$X_n = \sum_{j=0}^{\infty} \pi_{n,j}(\varepsilon_{n-j}).$$

The series converges absolutely almost surely.

Note that if there is an m such that $c_m < 1$, then by the concavity of the log function we get

$$\mathbb{E} \log \|\rho_{I_0}\| = (1 - q) \log \|\rho_0\| + q \log \|\rho_1\| < 0.$$

3. Law of large numbers

In the sequel, we will assume that ρ_{I_n} is symmetric and compact of the form

$$\rho_{I_n} = \sum_{j=0}^{\infty} A_{j,I_n} e_j \otimes e_j,$$

where (e_j) is a fixed orthonormal basis of H and $(A_{j,0})_j, (A_{j,1})_j$ are two nonincreasing, nonnegative sequences tending to 0 such that, for all j and all n ,

$$\rho_{I_n}(e_j) = A_{j,I_n} e_j.$$

Actually, only the eigenvalues are random, the change of regime is just in terms of level of influence.

The following proposition will be a technical tool used in the proof of our results. It generalizes the property—holding in the real-valued context—that it is possible to take out of the conditional expectation random variables which are measurable with respect to the considered σ -algebra.

Proposition 1. *Let $(L_j)_{j \in \mathbb{N}}$ be a sequence of real random variables defined on a probability space (Ω, \mathcal{A}, P) such that*

$$L = \sum_{j=0}^{\infty} L_j e_j \otimes e_j$$

is a bounded linear operator on H , (e_j) being a fixed orthonormal basis of H . Let \mathcal{A}_0 be a sub- σ -algebra of \mathcal{A} . Let X be an integrable H -valued random variable. If the sequence $(L_j)_{j \in \mathbb{N}}$ is \mathcal{A}_0 -measurable, then

$$E^{\mathcal{A}_0}[L(X)] = L(E^{\mathcal{A}_0}[X]).$$

Proof. In order to show that $L(E^{\mathcal{A}_0}[X])$ is the conditional expectation (see [5, (1.34)] for the definition) of $L(X)$ relative to \mathcal{A}_0 , it is sufficient to show that, for all \mathcal{A}_0 -measurable A ,

$$E[\mathbf{1}_A L(E^{\mathcal{A}_0}[X])] = E[\mathbf{1}_A L(X)],$$

that is,

$$E[\langle \mathbf{1}_A L(E^{\mathcal{A}_0}[X]), e_j \rangle] = E[\langle \mathbf{1}_A L(X), e_j \rangle], \quad j \in \mathbb{N},$$

or

$$E[\mathbf{1}_A L_j(E^{\mathcal{A}_0}[X], e_j)] = E[\mathbf{1}_A L_j(X, e_j)], \quad j \in \mathbb{N},$$

or

$$E[\mathbf{1}_A L_j E^{\mathcal{A}_0} \langle X, e_j \rangle] = E[\mathbf{1}_A L_j \langle X, e_j \rangle], \quad j \in \mathbb{N}.$$

But, since the random variables $\mathbf{1}_A L_j$ and $\langle X, e_j \rangle$ are \mathbb{R} -valued, we recognize here the well-known result about conditional expectation:

$$E^{\mathcal{A}_0}[\mathbf{1}_A L_j \langle X, e_j \rangle] = \mathbf{1}_A L_j E^{\mathcal{A}_0} \langle X, e_j \rangle.$$

The law of large numbers will be stated under the hypothesis:

$$(I_n) \text{ is i.i.d.} \tag{9}$$

Note that under this hypothesis, (I_n) is $(m - 1)$ -dependent with $m = 1$. Such an influence on the change of regime connects our model with RCA ones.

Set $S_n = X_1 + \dots + X_n$, where X_1 is centred.

Theorem 3. *Under the hypothesis (9) and all the assumptions of Theorem 1, and if $c_1 < 1$, then we have*

$$E \left\| \frac{S_n}{n} \right\|^2 = O\left(\frac{1}{n}\right) \tag{10}$$

and

$$\frac{n^{1/2}}{(\ln n)^\beta} \frac{S_n}{n} \rightarrow 0 \text{ as } n \rightarrow \infty \text{ a.s. } \beta > \frac{1}{2}. \tag{11}$$

Proof. Note that

$$E \|X_n + \dots + X_{n+p-1}\|^2 = \sum_{n \leq j, j' \leq n+p-1} E \langle X_j, X_{j'} \rangle.$$

by the stationarity of (X_n) . We use now (8) with $n = h$ and $k = h - 1$ to compute $\langle X_0, X_h \rangle$:

$$\langle X_0, X_h \rangle = \left\langle X_0, \sum_{j=0}^{h-1} \left(\prod_{p=0}^{j-1} \rho_{I_{h-p}} \right) (\varepsilon_{h-j}) \right\rangle + \langle X_0, \rho_{I_h} \dots \rho_{I_1}(X_0) \rangle. \tag{12}$$

Alternatively,

$$\begin{aligned} E \left\langle X_0, \sum_{j=0}^{h-1} \left(\prod_{p=0}^{j-1} \rho_{I_{h-p}} \right) (\varepsilon_{h-j}) \right\rangle &= \sum_{j=0}^{h-1} E \left[\left\langle X_0, \left(\prod_{p=0}^{j-1} \rho_{I_{h-p}} \right) (\varepsilon_{h-j}) \right\rangle \right] \\ &= \sum_{j=0}^{h-1} E \left[\left\langle \sum_i \langle X_0, e_i \rangle e_i, \sum_i A_{i, I_h} \dots A_{i, I_{h-j-1}} \langle \varepsilon_{h-j}, e_i \rangle e_i \right\rangle \right] \\ &= \sum_{j=0}^{h-1} E \left[\sum_i A_{i, I_h} \dots A_{i, I_{h-j-1}} \langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle \right]. \end{aligned}$$

But

$$\sum_i E [|A_{i, I_h} \dots A_{i, I_{h-j-1}} \langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle|] < \infty.$$

Indeed, supposing without loss of generality that $\|\rho_1\| \geq \|\rho_0\|$,

$$\begin{aligned} \sum_i E [|A_{i, I_h} \dots A_{i, I_{h-j-1}} \langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle|] &\leq \|\rho_1\|^j \sum_i E [|\langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle|] \\ &\leq \|\rho_1\|^j E \left[\sum_i |\langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle| \right] \\ &\leq \|\rho_1\|^j \frac{1}{2} E \left[\sum_i |\langle X_0, e_i \rangle|^2 + |\langle \varepsilon_{h-j}, e_i \rangle|^2 \right] \\ &\leq \|\rho_1\|^j \frac{1}{2} (E \|X_0\|^2 + E \|\varepsilon_0\|^2) < \infty. \end{aligned}$$

Consequently,

$$\begin{aligned} \mathbb{E}\left\langle X_0, \sum_{j=0}^{h-1} \left(\prod_{p=0}^{j-1} \rho_{I_{h-p}} \right) (\varepsilon_{h-j}) \right\rangle &= \sum_{j=0}^{h-1} \sum_i \mathbb{E}[A_{i,I_h} \cdots A_{i,I_{h-j-1}} \langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle] \\ &= \sum_{j=0}^{h-1} \sum_i \mathbb{E}[\mathbb{E}[A_{i,I_h} \cdots A_{i,I_{h-j-1}} \langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle \mid \sigma(I_1, \dots, I_h)]] \\ &= \sum_{j=0}^{h-1} \sum_i \mathbb{E}[A_{i,I_h} \cdots A_{i,I_{h-j-1}} \mathbb{E}[\langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle \mid \sigma(I_1, \dots, I_h)]] . \end{aligned}$$

Now, owing to the independence of X_0 and ε_{h-j} from $\sigma(I_1, \dots, I_h)$,

$$\mathbb{E}[\langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle \mid \sigma(I_1, \dots, I_h)] = \mathbb{E}[\langle X_0, e_i \rangle \langle \varepsilon_{h-j}, e_i \rangle] = 0,$$

and therefore

$$\mathbb{E}\left\langle X_0, \sum_{j=0}^{h-1} \left(\prod_{p=0}^{j-1} \rho_{I_{h-p}} \right) (\varepsilon_{h-j}) \right\rangle = 0.$$

We manage the second term of (12) in a similar way:

$$\begin{aligned} \mathbb{E}\langle X_0, \rho_{I_h} \cdots \rho_{I_1}(X_0) \rangle &= \mathbb{E}\left[\left\langle \sum_i \langle X_0, e_i \rangle e_i, \sum_i A_{i,I_h} \cdots A_{i,I_1} \langle X_0, e_i \rangle e_i \right\rangle \right] \\ &= \mathbb{E}\left[\sum_i A_{i,I_h} \cdots A_{i,I_1} \langle X_0, e_i \rangle^2 \right] \\ &= \sum_i \mathbb{E}[A_{i,I_h} \cdots A_{i,I_1} \langle X_0, e_i \rangle^2] \end{aligned}$$

because all the terms are nonnegative, so

$$\begin{aligned} \mathbb{E}\langle X_0, \rho_{I_h} \cdots \rho_{I_1}(X_0) \rangle &= \sum_i \mathbb{E}[\mathbb{E}[A_{i,I_h} \cdots A_{i,I_1} \langle X_0, e_i \rangle^2 \mid \sigma(I_1, \dots, I_h)]] \\ &= \sum_i \mathbb{E}[A_{i,I_h} \cdots A_{i,I_1} \mathbb{E}[\langle X_0, e_i \rangle^2 \mid \sigma(I_1, \dots, I_h)]] . \end{aligned}$$

But, owing to the independence of X_0 from I_1, \dots, I_h ,

$$\mathbb{E}[\langle X_0, e_i \rangle^2 \mid \sigma(I_1, \dots, I_h)] = \mathbb{E}[\langle X_0, e_i \rangle^2].$$

Hence,

$$\mathbb{E}\langle X_0, \rho_{I_h} \cdots \rho_{I_1}(X_0) \rangle = \sum_i \mathbb{E}[A_{i,I_h} \cdots A_{i,I_1}] \mathbb{E}[\langle X_0, e_i \rangle^2].$$

Since

$$\begin{aligned} \mathbb{E}[A_{i,I_h} \cdots A_{i,I_1}] &\leq \mathbb{E}[A_{1,I_h} \cdots A_{1,I_1}] \\ &\leq \sum_{j=0}^h C_h^j q^j (1-q)^{h-j} \|\rho_1\|^j \|\rho_0\|^{h-j} \\ &= ((1-q)\|\rho_0\| + q\|\rho_1\|)^h , \end{aligned}$$

we get

$$\begin{aligned} E\langle X_0, \rho_{I_h} \cdots \rho_{I_1}(X_0) \rangle &\leq [(1 - q)\|\rho_0\| + q\|\rho_1\|]^h \sum_i E[\langle X_0, e_i \rangle^2] \\ &= [(1 - q)\|\rho_0\| + q\|\rho_1\|]^h E\|X_0\|^2. \end{aligned}$$

Accordingly,

$$E\|X_n + \cdots + X_{n+p-1}\|^2 \leq 2pE\|X_0\|^2 \sum_{h=0}^{p-1} [(1 - q)\|\rho_0\| + q\|\rho_1\|]^h.$$

As $(1 - q)\|\rho_0\| + q\|\rho_1\| < 1$,

$$E\|X_n + \cdots + X_{n+p-1}\|^2 = O(p),$$

which implies (10), and applying Lemma 8.1 of [6], we obtain (11).

Let us note that the recent result obtained by Bosq [6, Lemma 8.1] seriously improved the rates of convergence. Indeed, the use of [5, Corollary 2.3] would have lead to rates with $n^{1/4}$ instead of $n^{1/2}$ in (11).

4. Estimation of the covariance estimator

At this point, we use the same method as in [4]. Let us recall that the covariance operator and the empirical covariance operator are defined respectively by

$$\begin{aligned} C(\cdot) &= E[\langle X_0, \cdot \rangle X_0], \\ C_n(\cdot) &= \frac{1}{n} \sum_{t=1}^n \langle X_t, \cdot \rangle X_t. \end{aligned}$$

Also, $(\mathcal{S}, \|\cdot\|_{\mathcal{S}})$ is the space of Hilbert–Schmidt operators on H and $(\mathcal{N}, \|\cdot\|_{\mathcal{N}})$ is the space of nuclear operators on H .

Theorem 4. *Under the hypothesis (9), all the assumptions of Theorem 1, and if $c_2 < 1$, we have*

$$E\|C_n - C\|_{\mathcal{S}}^2 = O\left(\frac{1}{n}\right), \tag{13}$$

$$\|C_n - C\|_{\mathcal{S}} \rightarrow 0 \text{ as } n \rightarrow \infty \text{ a.s.} \tag{14}$$

Proof. Define the following random variables with values in \mathcal{S} :

$$Z_t = \langle X_t, \cdot \rangle X_t - C, \quad t = 1, 2, \dots$$

Note that (Z_t) is a stationary process. Moreover,

$$\begin{aligned} E\|Z_t\|_{\mathcal{S}}^2 &\leq 2E\|\langle X_t, \cdot \rangle X_t\|_{\mathcal{S}}^2 + 2\|C\|_{\mathcal{S}}^2 \\ &\leq 2E\|X_0\|^4 + 2\|C\|_{\mathcal{N}}^2 \end{aligned}$$

by [5, (1.49), (1.55)]. Hence,

$$E\|Z_t\|_{\mathcal{H}}^2 \leq 2E\|X_0\|^4 + 2(E\|X_0\|^2)^2 =: 2M < \infty$$

by [5, (1.59)].

Because of the stationarity,

$$E\|Z_n + \dots + Z_{n+p-1}\|_{\mathcal{H}}^2 \leq 2pM + 2 \sum_{h=1}^{p-1} (p-h)E\langle Z_0, Z_h \rangle_{\mathcal{H}}.$$

Let us find an upper bound for $E\langle Z_0, Z_h \rangle_{\mathcal{H}}$. Use the relation

$$E\langle Z_0, Z_h \rangle_{\mathcal{H}} = \sum_{j,l} E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle X_h, e_j \rangle \langle X_h, e_l \rangle) - \sum_{j,l} \langle C(e_j), e_l \rangle^2 \quad (15)$$

(see [4, (25)]), then replace X_h by $\varepsilon_h + \rho_{I_h}(\varepsilon_{h-1}) + \dots + \rho_{I_h} \dots \rho_{I_2}(\varepsilon_1) + \rho_{I_h} \dots \rho_{I_1}(X_0)$ owing to (8). The first sum of (15) contains only terms of the form

$$E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_l \rangle)$$

or

$$E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \dots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_1}(X_0), e_l \rangle).$$

Terms of the following form vanish for $p \neq k$: in fact,

$$\begin{aligned} & E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_l \rangle) \\ &= E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle) E(\langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_l \rangle) \end{aligned}$$

owing to independence properties. But

$$\begin{aligned} & E(\langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_l \rangle) \\ &= E[E[\langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_l \rangle \mid (I_n)]] \\ &= E[E[\langle \rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \mid (I_n)] E[\langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_l \rangle \mid (I_n)]] \end{aligned}$$

because the random variables $\rho_{I_h} \dots \rho_{I_p}(\varepsilon_{p-1})$ and $\rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1})$ are conditionally independent with respect to (I_n) by the independence of ε_{p-1} and ε_{k-1} . Now

$$\begin{aligned} E(\langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_j \rangle) &= E[E[\langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_j \rangle \mid (I_n)]] \\ &= E[\langle E[\rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}) \mid (I_n)], e_j \rangle] \\ &= E[\langle \rho_{I_h} \dots \rho_{I_k} E[\varepsilon_{k-1} \mid (I_n)], e_j \rangle] \end{aligned} \quad (16)$$

by Proposition 1. Owing to independence properties and knowing that ε_0 is centred,

$$E[\varepsilon_{k-1} \mid (I_n)] = E[\varepsilon_{k-1}] = 0.$$

Terms of the form

$$E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \dots \rho_{I_k}(\varepsilon_{k-1}), e_j \rangle \langle \rho_{I_h} \dots \rho_{I_1}(X_0), e_l \rangle)$$

also vanish due to independence properties, in a way similar to (16).

For $p = k$, consider the terms of the form

$$\begin{aligned} & E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_l \rangle) \\ &= E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle) E(\langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_l \rangle) \end{aligned}$$

as a result of independence properties. But

$$\begin{aligned} \langle C(e_j), e_l \rangle &= E(\langle X_h, e_j \rangle \langle X_h, e_l \rangle) \\ &= E[\langle \varepsilon_h, e_j \rangle \langle \varepsilon_h, e_l \rangle] + \cdots + E(\langle \rho_{I_h} \cdots \rho_{I_2}(\varepsilon_1), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_2}(\varepsilon_1), e_l \rangle) \\ &\quad + E(\langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle) \end{aligned}$$

by (8), the other terms involved in this relation yield vanishing expectations by use of techniques of the kind seen before. Accordingly,

$$\begin{aligned} & \sum_p E(\langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_p}(\varepsilon_{p-1}), e_l \rangle) \\ &= \langle C(e_j), e_l \rangle - E(\langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle). \end{aligned}$$

Thus,

$$\begin{aligned} E\langle Z_0, Z_h \rangle_{\mathcal{S}} &= \sum_{j,l} E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle) \\ &\quad - \sum_{j,l} \langle C(e_j), e_l \rangle E(\langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle). \end{aligned}$$

Let us find upper bounds for each sum in absolute value. For the first one, note that

$$\begin{aligned} & \left| \sum_{j,l} E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle) \right| \\ &= \sum_{j,l} E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1}) \langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2] \\ &= \sum_{j,l} E[E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1}) \langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2 \mid \sigma(I_1, \dots, I_h)]] \\ &= \sum_{j,l} E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1}) E[\langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2 \mid \sigma(I_1, \dots, I_h)]] \\ &= \sum_{j,l} E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1}) E[\langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2]] \\ &= \sum_{j,l} E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1})] E[\langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2] \\ &\leq [(1 - q)\|\rho_0\|^2 + q\|\rho_1\|^2]^h \sum_{j,l} E[\langle X_0, e_j \rangle^2 \langle X_0, e_l \rangle^2] \\ &\leq [(1 - q)\|\rho_0\|^2 + q\|\rho_1\|^2]^h E\|X_0\|^4. \end{aligned}$$

For the second sum, we reason as previously and obtain

$$\begin{aligned} & E(\langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle) \\ &= E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1})(X_0, e_j) \langle X_0, e_l \rangle] \\ &= E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1})] E[\langle X_0, e_j \rangle \langle X_0, e_l \rangle] \\ &= E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1})] \langle C(e_j), e_l \rangle. \end{aligned}$$

Hence,

$$\begin{aligned} & \left| \sum_{j,l} \langle C(e_j), e_l \rangle E(\langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_j \rangle \langle \rho_{I_h} \cdots \rho_{I_1}(X_0), e_l \rangle) \right| \\ &= \sum_{j,l} \langle C(e_j), e_l \rangle^2 E[(A_{j,I_h} \cdots A_{j,I_1})(A_{l,I_h} \cdots A_{l,I_1})] \\ &\leq [(1 - q) \|\rho_0\|^2 + q \|\rho_1\|^2]^h \|C\|_{\mathfrak{H}}^2. \end{aligned}$$

But

$$\|C\|_{\mathfrak{H}}^2 \leq \|C\|_{\mathcal{M}}^2 = (E\|X_0\|^2)^2,$$

so

$$|E\langle Z_0, Z_h \rangle_{\mathfrak{H}}| \leq c_2^h [E\|X_0\|^4 + (E\|X_0\|^2)^2] = M c_2^h.$$

Consequently, as $c_2 < 1$,

$$\begin{aligned} \left| 2 \sum_{h=1}^{p-1} (p-h) E\langle Z_0, Z_h \rangle_{\mathfrak{H}} \right| &\leq 2M \sum_{h=1}^{p-1} (p-h) c_2^h \\ &\leq 2Mp \sum_{h=1}^{p-1} c_2^h = 2Mp c_2 \frac{1 - c_2^{p-1}}{1 - c_2} \\ &\leq 2Mp \frac{c_2}{c_2 - 1}. \end{aligned}$$

In conclusion,

$$E\|Z_n + \cdots + Z_{n+p-1}\|_{\mathfrak{H}}^2 = O(p),$$

and therefore we get (13) directly, and from [5, Corollary 2.3] we obtain (14).

Acknowledgements

I would like to thank the referee for several ideas which improved the presentation and the results of this paper. Furthermore, I am indebted to Jerome Dedecker and Julien Damon for their technical advice.

References

- [1] BASS, J. (1955). Sur la compatibilité des fonctions de répartition. *C. R. Acad. Sci. Paris* **240**, 839–841.
- [2] BESSE, P. AND CARDOT, H. (1996). Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Canad. J. Statist.* **24**, 467–487.
- [3] BESSE, P., CARDOT, H. AND STEPHENSON, D. (2000). Autoregressive forecasting of some functional climatic variations. *Scand. J. Statist.* **27**, 673–687.

- [4] BOSQ, D. (1991). Modelization, nonparametric estimation and prediction for continuous time processes. In *Nonparametric Functional Estimation and Related Topics* (NATO ASI Ser. C **335**), ed. G. Roussas, Kluwer, Dordrecht, pp. 509–529.
- [5] BOSQ, D. (2000). *Linear Processes in Function Spaces* (Lecture Notes Statist. **149**). Springer, New York.
- [6] BOSQ, D. (2002). Estimation of mean and covariance operator of autoregressive processes in Banach spaces. *Statist. Inf. Stoch. Process.* **7**, 1–20.
- [7] BOUGEROL, P. AND PICARD, N. (1992). Strict stationarity of generalized autoregressive processes. *Ann. Prob.* **20**, 1714–1730.
- [8] BRANDT, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Prob.* **18**, 211–220.
- [9] CAVALLINI, A. *et al.* (1994). Nonparametric prediction of harmonic levels in electrical networks. In *Proc. IEEE ICHPS 6, Bologna*.
- [10] DAMON, J. AND GUILLAS, S. (2002). The inclusion of exogenous variables in functional autoregressive ozone forecasting. To appear in *Environmetrics*.
- [11] DE HAAN, L., RESNICK, S., ROOTZEN, H. AND DE VRIES, C. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Process. Appl.* **32**, 213–224.
- [12] FRANCO, C. AND ZAKOÏAN, J.-M. (2001). Stationarity of multivariate Markov-switching ARMA models. *J. Econometrics* **102**, 339–364.
- [13] GUILLAS, S. (2000). Non-causalité et discrétisation fonctionnelle, théorèmes limites pour un processus ARHX(1). *C. R. Acad. Sci. Paris Math.* **331**, 91–94.
- [14] HORST, U. (2001). The stochastic equation $Y_{t+1} = A_t Y_t + B_t$ with non-stationary coefficients. *J. Appl. Prob.* **38**, 80–94.
- [15] MÉLARD, G. AND ROY, R. (1988). Modèles de séries chronologiques avec seuils. *Rev. Statist. Appl.* **36**, 5–24.
- [16] NICHOLLS, D. F. AND QUINN, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction* (Lecture Notes Statist. **11**). Springer, New York.
- [17] PETRUCCELLI, J. D. AND WOOLFORD, S. W. (1984). A threshold AR(1) model. *J. Appl. Prob.* **21**, 270–286.
- [18] POURAHMADI, M. (1986). On stationarity of the solution of a doubly stochastic model. *J. Time Ser. Anal.* **7**, 123–131.
- [19] RIO, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants* (Math. Appl. **31**). Springer, New York.
- [20] TJØSTEIM, D. (1986). Some doubly stochastic time series models. *J. Time Ser. Anal.* **7**, 51–72.
- [21] TJØSTEIM, D. (1994). Non-linear time series: a selective review. *Scand. J. Statist.* **21**, 97–130.
- [22] TONG, H. (1983). *Threshold Models in Nonlinear Time Series Analysis* (Lecture Notes Statist. **21**). Springer, Berlin.
- [23] TONG, H. (1990). *Nonlinear Times Series*. Oxford University Press.
- [24] YAO, J. (2001). On square-integrability of an AR process with Markov switching. *Statist. Prob. Lett.* **52**, 265–270.
- [25] YAO, J.-F. AND ATTALI, J.-G. (2000). On stability of nonlinear AR processes with Markov switching. *Adv. Appl. Prob.* **32**, 394–407.