– PANTHÉON - SORBONNE –
UNIVERSITÉ PARIS 1

*Centre de Recherche*

SAMOS - MATISSE

CNRS, UMR 8595

*Statistique Appliquée &*
*Modélisation Stochastique*

# On stability of nonlinear AR processes with Markov switching

J. YAO    J.G. ATTALI

**Prépublication du SAMOS N° 101**    ***(Version modifiée)***

**Avril 1999**

*Applied Probability Trust* (April 28, 1999)

# ON STABILITY OF NONLINEAR AR PROCESSES WITH MARKOV SWITCHING

J.-F. YAO,* *SAMOS, Université Paris 1*

J.-G. ATTALI,** *Labo. de Probabilités, Université Paris 6 and SAMOS, Université Paris 1*

**Abstract**

We investigate the stability problem for a nonlinear autoregressive model with Markov switching. First we give conditions for the existence and the uniqueness of a stationary ergodic solution. The existence of moments of such a solution is then examined and we establish a strong law of large numbers for a wide class of unbounded functions, as well as a central limit theorem under an irreductibility condition.

MARKOV SWITCHING; NONLINEAR AR; LYAPOUNOV FUNCTIONS, STABILITY

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J20

SECONDARY 62M10

## 1. Introduction

Let $\mathbf{X} := (X_n)_{n\geq 0}$ be a positive recurrent Markov chain on a finite set $E = \{1, \ldots, m\}$, with transition probability matrix $P$ and invariant probability measure (hereafter i.p.m.) $\mu$. We consider a nonlinear AR process with Markov switching $\mathbf{Y} = (Y_n)$ (abbreviated as NAR-MS) defined for integers $n \geq 1$, by

$$Y_n = f_{X_n}(Y_{n-1}) + \varepsilon_n, \quad Y_n \in \mathbb{R}^d. \tag{1}$$

* Postal address: 90, rue de Tolbiac, F-75634 Paris cedex 13, France.

* Postal address: 4, place Jussieu, F-75252 Paris cedex 05, France.

Here the error process $\varepsilon := (\varepsilon_n)_{n \geq 0}$ is an i.i.d $\mathbb{R}^d$-valued sequence of random variables (our results can be extended to the case where the error variable $\varepsilon_n$ at time $n$ depend also on the current value $X_n$) and $(f_k)$ is a family of nonlinear autoregressive functions. We assume that $\varepsilon$, $\mathbf{X}$ and the variable $Y_0$ are independent.

The use of the Markov switching offers new possibilities for modeling time series "subject to discret shifts in regime - episodes across which the dynamic behaviour of the series is markedly differents", as noted by Hamilton [9] who first introduced such a model to analyse the US annual GNP (gross national product) series. These models have since then attracted a considerable interest in the statistical community, especially for econometrical series modeling ([10], [11], [12], [14]). However, very few is proved about their theoretical properties including stationarity or stability/ergodicity, especially in the nonlinear case (i.e. the $f_k$'s are nonlinear). On the other hand, nonlinear autoregression functions has been widely employed in applications (see e.g. [8], [18], [16]).

The main aim of this paper is the stability of the model $\mathbf{Y}$. We give conditions which ensure respectively

- the existence and the uniqueness of a strictly stationary ergodic solution for the model $\mathbf{Y}$ – **problem A**;
- the existence of moment of order $s \geq 1$ of the stationary distribution $\nu$ (i.e. the common marginal distribution of the stationary solution) – **problem B**;
- limit theorems including strong law of large numbers (SLLN) and the central limit theorem (CLT) – **problem C**.

Such limit theorems are a basic tool for an asymptotic estimation theory of the model $\mathbf{Y}$ (see e.g. [17], [19] for nonlinear AR models), especially when they can be applied to unbounded functions like $g(y) = ||y||^s$, with $s \geq 1$.

We shall consider two situations. First in Section 3, the autoregressive functions $f_k$ are *sublinear*. By using the Lyapounov method, we establish conditions ensuring the stability of the model and solve the Problems **A-B-C**. In the second situation (Section 4), the $f_k$'s are *Lipschitz*. When the regime chain $\mathbf{X}$ is stationary, the model (1) is then a particular case of iterative Lipschitz models considered by Bougerol [3]. This author has established accurate conditions for the existence and the uniqueness of a strictly stationary ergodic solution (see also [2], and [5], [4] for linear models with random coefficients). Therefore our solution to the problem **A** will be based on the results of [3] and it improves a previous result reported in [7]. We also provide solutions to the problems **B** and **C** in this case.

The paper is organised as follows. In Section 2, we introduce an extended Markov chain associated to the model (1) and establish some preliminary results. Then we examine the sublinear case and the Lipschitz case in Sections 3 and 4, respectively. Our main tool is a modified Lyapounov criterion of stability for Markov chains which is recalled in Appendix A.

## 2. Preliminary results on the Markov chain $Z_n = (X_n, Y_n^T)^T$

Let us denote $\lambda(\phi)$ the integral of a function $\phi$ with respect to some positive measure $\lambda$ and set $(\pi\phi)(x) := \int \pi(x, dy)\phi(x)$ if $\pi(x, dy)$ is a Markov transition probability kernel. The $m$-dimensional vector with constant component 1 is denoted by $\mathbf{1}$ and the spectral radius of a real matrix $A$ by $\rho(A)$.

The NAR-MS process $\mathbf{Y}$ is not a Markov process in general. However the extended process $\mathbf{Z} = (Z_n)$ with $Z_n := (X_n, Y_n^T)^T$ is a Markov chain as suggested by the following Lemma 1. For general results on the stability theory of Markov chains, we refer to Meyn and Tweedie [15] and Duflo [6] where we have extracted some of mostly used results in Appendix A for the reader's convenience. Throughout the paper, the transition kernel of $\mathbf{Z}$ is denoted $\Pi(z, dz')$.

*Lemma 1* *The extended process* $\mathbf{Z}$ *is a Markov chain on* $\mathcal{Z} = E \times \mathbb{R}^d$*. It is a Feller chain if the* $m$ *functions* $f_k$, $1 \leq k \leq m$ *are continuous. It is a srong Feller chain if in addition the random variable* $\varepsilon_1$ *has a density with respect to the Lebesgue measure on* $\mathbb{R}^d$*.*

**Proof.** There is a measurable function $h : E \times \mathbb{R} \to E$ and a sequence of i.i.d., real and centered random variables $(u_n\,,\, n \geq 1)$ which are independent from the error process $(\varepsilon_n)$ such that

$$X_n = h(X_{n-1}\,, u_n)\ .$$

$(Z_n)_{n\geq 0}$ is then an iterative Markov model (see [6], Section 6.1.2)

$$Z_n = \begin{pmatrix} X_n \\ Y_n \end{pmatrix} = \begin{pmatrix} h(X_{n-1}\,, u_n) \\ f_{h(X_{n-1}\,, u_n)}(Y_{n-1}) + \varepsilon_n \end{pmatrix} .$$

The announced results can be proved in the same way as for NAR models (*cf.* [6], §6.1). ■

We first study the properties of $\mathbf{Z}$ and then derive the properties of the marginal process $\mathbf{Y}$. For the problem $\mathbf{A}$, it is understood that the chain $\mathbf{X}$ is stationary, i.e. $X_0 \sim \mu$. Therefore such a stationary solution exists if and only if the chain $\mathbf{Z}$ has an i.p.m. $\sigma$. In this case, the i.p.m. $\mu$ (of $\mathbf{X}$) and the stationary distribution $\nu$ are just the marginals of $\sigma$:

$$(2) \qquad d\mu(x) = \int_{\mathbb{R}^d} d\sigma(x,y)\ , \qquad d\nu(y) = \sum_{x\in E} d\sigma(x,y)\ .$$

Furthermore, this stationary solution is unique if and only if the i.p.m. $\sigma$ of $\mathbf{Z}$ is unique. For the problem $\mathbf{B}$, by (2), $\nu(||y||^s) < \infty$ if and only if $\sigma(||y||^s) < \infty$. Finally for the problem $\mathbf{C}$, we consider the following SLLN and CLT for the extended chain $\mathbf{Z}$ : for all initial distribution $\sigma_0$ of $(X_0, Y_0)$, it holds

$$(3) \qquad \text{a.s.,} \quad \forall \varphi \in \mathcal{F}, \quad \frac{1}{n}\sum_{k=1}^{n} \varphi(X_k, Y_k) \quad \longrightarrow \quad \sigma(\varphi)$$

$$(4) \qquad \forall \varphi \in \mathcal{F}, \quad \frac{1}{\sqrt{n}}\sum_{k=1}^{n} [\varphi(X_k, Y_k) - \sigma(\varphi)] \quad \stackrel{\mathcal{D}}{\longrightarrow} \quad \mathcal{N}(0, a_\varphi)\ .$$

Here $\stackrel{\mathcal{D}}{\longrightarrow}$ stands for convergence in distribution, $a_\varphi$ some positive number (asymptotic variance) and $\mathcal{F}$ some class of functions defined on $\mathcal{Z}$ which will be specified later. By using the associated canonical space $(\mathcal{Z}^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}})$ (see e.g. [6], p. 185), the convergence (3) will hold for any arbitrary initial distribution $\sigma_0$ if and only if it holds $\mathbb{P}_{(x,y)}$-a.s. for all $(x,y) \in \mathcal{Z}$, where $\mathbb{P}_{(x,y)}$ is the probability distribution defined on the canonical space with the initial condition $X_0 \equiv x$ and $Y_0 \equiv y$ (the underlying expectation is denoted by $\mathbb{E}_{(x,y)}$). Let us also define for $q > 0$

$$(5)\ \mathcal{B}(q) \quad := \quad \left\{\varphi \ :\ \mathcal{Z} \to \mathbb{R} \ \text{ measurable s.t. } \ \forall (x,y) \in \mathcal{Z}\ , |\varphi(x,y)| \leq \text{const.}(1 + ||y||^q)\right\} .$$

Also for two metric spaces $F$ and $G$, the space of bounded continuous functions from $F$ to $G$ is denoted $\mathcal{C}_b(F, G)$.

Furthermore, we will invoke the $V$-uniform ergodicity for a Markov chain as defined in [15] (chapter 16). It is worth noting that any $V$-uniformly ergodic chain is in particular geometrically ergodic.

## 3. The sublinear case

The NAR-MS model (1) is called *sublinear* if the $m$ functions $f_k$ are continuous and there are positive constants $(a_k,\, b_k)$ such that for $k = 1, \dots, m$, $y \in \mathbb{R}^d$ and some norm $||\cdot||$ on $\mathbb{R}^d$, it holds

$$||f_k(y)|| \le a_k\, ||y|| + b_k \ . \tag{6}$$

For any measurable function $\varphi : \mathcal{Z} \to \mathbb{R}$, let $\bar{\varphi} = \varphi - \sigma(\varphi)$ and

$$\gamma_\varphi^2 = \mathbb{E}_\sigma \left[\bar{\varphi}^2(Z_0)\right] + 2\sum_{k=1}^{\infty} \mathbb{E}_\sigma \left[\bar{\varphi}(Z_0)\bar{\varphi}(Z_k)\right] \ . \tag{7}$$

*Theorem 1* ***(existence and uniqueness of a stationary ergodic solution)*** *Consider a sublinear NAR-MS model in the sense of (6). Assume*

*(i).* $\mathbb{E}\, ||\varepsilon_1||^\delta < \infty$ *for some* $\delta > 0$.

*(ii).* $\beta := \sum_{k=1}^m \mu_k \log a_k < 0$.

*(iii). (a). the variable* $\varepsilon_1$ *has a density with respect to the Lebesgue measure on* $\mathbb{R}^d$ *and*

*(b). this density is everywhere positive on* $\mathbb{R}^d$.

*Then, there is a* $\gamma_0 \in ]0, \delta]$ *such that*

*(1) The chain* $\mathbf{Z}$ *is* $V$*-uniformly ergodic with* $V(x,y) = ||y||^{\gamma_0} + 1$.

*(2) The i.p.m.* $\sigma$ *of* $\mathbf{Z}$ *satisfies* $\sigma(||y||^{\gamma_0}) < \infty$.

*(3) The SLLN (3) holds for all* $\varphi \in \mathcal{F} = \mathcal{B}(\gamma_0)$.

*(4) For all* $\varphi \in \mathcal{F} = \mathcal{B}(\gamma_0/2)$, *the constant* $\gamma_\varphi^2$ *is well defined, non-negative and finite, Furthermore, if* $\gamma_\varphi^2 > 0$, *the CLT (4) holds for* $\varphi$ *with asymptotic variance* $\gamma_\varphi^2$.

**Proof.** For $a, b \geq 0$, and any positive integer $n$, $(a+b)^{\frac{1}{n}} \leq a^{\frac{1}{n}} + b^{\frac{1}{n}}$. Then we have

$$\begin{aligned}
||Y_n||^{\frac{1}{n}} &= ||f_{X_n}(Y_{n-1}) + \varepsilon_n||^{\frac{1}{n}} \\
&\leq (a_{X_n}||Y_{n-1}|| + b_{X_n} + ||\varepsilon_n||)^{\frac{1}{n}} \\
&\leq (a_{X_n})^{\frac{1}{n}} ||Y_{n-1}||^{\frac{1}{n}} + (b_{X_n} + ||\varepsilon_n||)^{\frac{1}{n}} \\
&\leq \left(a_{X_n} a_{X_{n-1}}\right)^{\frac{1}{n}} ||Y_{n-2}||^{\frac{1}{n}} + (a_{X_n})^{\frac{1}{n}} \left(b_{X_{n-1}} + ||\varepsilon_{n-1}||\right)^{\frac{1}{n}} + (b_{X_n} + ||\varepsilon_n||)^{\frac{1}{n}} \\
&\vdots \\
&\leq (a_{X_n} \cdots a_{X_1})^{\frac{1}{n}} ||Y_0||^{\frac{1}{n}} + (b_{X_n} + ||\varepsilon_n||)^{\frac{1}{n}} + + \sum_{i=1}^{n-1} \left(a_{X_n} \cdots a_{X_{i+1}}\right)^{\frac{1}{n}} (b_{X_i} + ||\varepsilon_i||)^{\frac{1}{n}} . \qquad (8)
\end{aligned}$$

Since $(X_n)$ is a positive recurrent chain on a finite set, it holds that for any initial condition $X_0 = x \in \{1, \ldots, m\}$

$$\text{a.s.} \quad \frac{1}{n}\left[\log(a_{X_1}) + \cdots + \log(a_{X_n})\right] \longrightarrow \sum_k \mu_k \log a_k < 0.$$

This means that

$$\text{a.s.} \quad (a_{X_n} \cdots a_{X_1})^{\frac{1}{n}} \longrightarrow a_1^{\mu_1} \cdots a_m^{\mu_m} < 1.$$

The variables $(a_{X_n} \cdots a_{X_1})^{\frac{1}{n}}$ being bounded by $\sup_k a_k$, it comes by the dominated convergence theorem

$$\forall x \in \{1, \ldots, m\}, \quad \lim_{n \to +\infty} \mathbb{E}\left( \left(a_{X_0} \cdots a_{X_{n-1}}\right)^{\frac{1}{n}} / X_0 = x \right) = a_1^{\mu_1} \cdots a_m^{\mu_m} < 1 .$$

Furthermore, there is a $p \geq \dfrac{1}{\delta}$ such that

$$\alpha_p := \sup_{x \in \{1,\ldots,m\}} \mathbb{E}\left( \left(a_{X_p} \cdots a_{X_1}\right)^{\frac{1}{p}} / X_0 = x \right) < 1,$$

Taking expectation $\mathbb{E}_{(x,y)}$ of the inequality (8), with an arbitrary $(x, y) \in E \times \mathbb{R}^d$, yields

$$\mathbb{E}_{(x,y)} \left( ||Y_p||^{\frac{1}{p}} \right) \leq \alpha_p ||y||^{\frac{1}{p}} + \beta_p ,$$

where $\beta_p := \mathbb{E}_{(x,y)} \left[ \left(b_{X_p} + ||\varepsilon_p||\right)^{\frac{1}{p}} + \sum_{i=1}^{p-1} \left(a_{X_p} \cdots a_{X_{i+1}}\right)^{\frac{1}{p}} (b_{X_i} + ||\varepsilon_i||)^{\frac{1}{p}} \right]$. The last relation implies that for the Lyapounov function $V(x, y) = ||y||^{\frac{1}{p}} + 1$ on $\mathcal{Z}$, it holds

$$\Pi^p \, V(x, y) \leq \alpha_p \, V(x, y) + \beta_p + 1 - \alpha_p. \qquad (9)$$

On the other hand, the condition (iii)-(b) ensures that the transition kernel $\Pi$ is $\phi$-irreducible, $\phi$ being the product measure (counting $\otimes$ Lebesgue) on $\mathcal{Z}$. As $\Pi$ is (strongly) Feller (Lemma 1), it follows, by Proposition 6.2.8 of [15], that all compact susbsets of $\mathcal{Z}$ are petite. Hence, by Theorem 16.1.2 of [15], the chain is $V$-uniformly ergodic. This etablishes the conclusion (1). The conclusions (2) and (3) follow from the Lyapounov criterion recalled in Theorem 5 (Appendix A). Finally the conclusion (4) is a straightforward consequence of Theorem 17.5.3 of [15]. ■

The condition (iii)-b in Theorem 1 is used to ensure the uniqueness of the i.p.m. $\sigma$. It can be weakened or replaced by any other condition that guarantees this uniqueness (through the irreducibility for example). For instance, the following is a sustitute

*(iii)-(b') The density of $\varepsilon_1$ is non null on some half-space for $\mathbb{R}^d$.*

However, the SLLN and CLT given in Theorem 1 have a limited interest because we do not know any explicit lower bound for the exponent $\gamma_0$ (which can be as small as the constant $\delta$ is). It is thus important to extend this SLLN to functions of greater order $s \geq 1$. To this end, let us define the $m \times m$ matrix

$$Q_s = ((a_j)^s\, p_{ij}) = \begin{pmatrix} (a_1)^s\, p_{11} & \cdots & (a_m)^s\, p_{1m} \\ \vdots & \ddots & \vdots \\ (a_1)^s\, p_{m1} & \cdots & (a_m)^s\, p_{mm} \end{pmatrix}, \tag{10}$$

where the $p_{ij}$'s are the elements of the transition matrix $P$, and the $a_k$'s are given in (6). This matrix was introduced by Francq and Roussignol [7] in the particular case $s = 1$.

*Theorem 2* ***(moments of order $s \geq 1$ and limit theorems)*** *Let $s \geq 1$. For the sublinear NAR-MS model considered in Theorem 1, we make the same assumptions (iii)-(a)-(b) and replace those of (i) and (ii) by the following*

*(i')* $\mathbb{E}||\varepsilon_1||^s < \infty$.

*(ii')* $\rho(Q_s) < 1$.

*Then, all the conclusions of Theorem 1 hold with $\gamma_0$ replaced by $s$.*

Besides, compared to Theorem 1, (i') clearly implies (i) and (ii') is stronger than (ii), since by Lemma 2 of Appendix B

$$s \sum_{k=1}^{m} \mu_k \log a_k \leq \log \rho(Q_s) < 0 \ .$$

**Proof.** As $\rho(Q_s) < 1$ and $Q_s$ is a nonnegative matrix, by Perron-Frobenius theorem, there is some positive integer $p$ for which

$$(Q_s)^p \mathbf{1} < \mathbf{1} \ . \tag{11}$$

Here we have written $\mathbf{u} < \mathbf{v}$ for two real vectors $\mathbf{u} = (u_i)$ and $\mathbf{v} = (v_i)$ when $u_i < v_i$ for all $i$. Let us prove the contraction inequality (9) for this exponent $p$ and the Lyapounov function $V(x,y) = ||y||^s + 1$ defined on $E \times \mathbb{R}^d$. We have

$$\begin{aligned} ||Y_p||^s &= \left(||f_{X_p}(Y_{p-1}) + \varepsilon_p||\right)^s \\ &\leq \left(a_{X_p}||Y_{p-1}|| + b_{X_p} + ||\varepsilon_p||\right)^s \\ &\ \ \vdots \\ &\leq \left(a_{X_p} \cdots a_{X_1}||Y_0|| + b_{X_p} + ||\varepsilon_p|| + \sum_{i=1}^{p-1} a_{X_p} \cdots a_{X_{i+1}} \left(b_{X_i} + ||\varepsilon_i||\right)\right)^s . \end{aligned} \tag{12}$$

Hence by taking expectation for arbitrary $(x,y) \in E \times \mathbb{R}^d$, we get

$$\left(\mathbb{E}_{(x,y)}||Y_p||^s\right)^{\frac{1}{s}} \leq \left(\mathbb{E}_{(x,y)} \left(a_{X_p} \cdots a_{X_1}||Y_0|| + b_{X_p} + ||\varepsilon_p|| + \sum_{i=1}^{p-1} a_{X_p} \cdots a_{X_{i+1}} \left(b_{X_i} + ||\varepsilon_i||\right)\right)^s\right)^{\frac{1}{s}}$$

It follows from the $L_s$ norm inequality that

$$\begin{aligned} \left(\mathbb{E}_{(x,y)}||Y_p||^s\right)^{\frac{1}{s}} &\leq \left(\mathbb{E}_{(x,y)} \left(a_{X_p} \cdots a_{X_1}||Y_0||\right)^s\right)^{\frac{1}{s}} + \\ &\quad + \left(\mathbb{E}_{(x,y)} \left(b_{X_p} + ||\varepsilon_p|| + \sum_{i=1}^{p-1} a_{X_p} \cdots a_{X_{i+1}} \left(b_{X_i} + ||\varepsilon_i||\right)\right)^s\right)^{\frac{1}{s}} , \end{aligned}$$

This is nothing but

$$\begin{aligned} \left(\mathbb{E}_{(x,y)}||Y_p||^s\right)^{\frac{1}{s}} &\leq \left(\mathbb{E}\left(a_{X_p} \cdots a_{X_1}\right)^s / X_0 = x\right)^{\frac{1}{s}} ||y|| + \\ &\quad + \left(\mathbb{E}_{(x,y)} \left(b_{X_p} + ||\varepsilon_p|| + \sum_{i=1}^{p-1} a_{X_p} \cdots a_{X_{i+1}} \left(b_{X_i} + ||\varepsilon_i||\right)\right)^s\right)^{\frac{1}{s}} , \end{aligned} \tag{13}$$

This inequality can be written as

$$\forall (x,y) \in E \times \mathbb{R}^d, \quad \mathbb{E}_{(x,y)} ||Y_p||^s \le (\alpha_p ||y|| + \beta_p)^s . \tag{14}$$

where we have used

$$\begin{cases} \alpha_p := \sup_{x \in \{1,\cdots,m\}} \left( \mathbb{E} \left( a_{X_p} \cdots a_{X_1} \right)^s / X_0 = x \right)^{\frac{1}{s}} , \\ \beta_p := \left( \mathbb{E}_{(x,y)} \left( b_{X_p} + ||\varepsilon_p|| + \sum_{i=1}^{p-1} a_{X_p} \cdots a_{X_{i+1}} \left( b_{X_i} + ||\varepsilon_i|| \right) \right)^s \right)^{\frac{1}{s}} , \end{cases}$$

On the other hand, a straightforward calculus yields

$$\begin{aligned} &\mathbb{E} \left[ \left( a_{X_p} \cdots a_{X_1} \right)^s \mid X_0 = x \right] \\ &= \left( p_{x1} (a_1)^s, \ldots p_{xm} (a_m)^s \right) (Q_s)^{p-1} \mathbf{1} = x\text{-th component of } (Q_s)^p \, \mathbf{1} \, . \end{aligned}$$

By (11), this component is smaller than 1. Hence $\alpha_p < 1$ and

$$\limsup_{||y|| \to +\infty} \frac{\Pi^p V(x,y)}{V(x,y)} \le \alpha_p^s < 1 , \qquad \text{with } V(x,y) := ||y||^s + 1. \tag{15}$$

This last relation is equivalent to (9), since $\Pi^p V$ is bounded on compact sets.

The end of the proof is the same as the one used to conclude the proof of Theorem 1 with here the new Lyapounov function $V(x,y) = ||y||^s + 1$. ■

## 4. The Lipschitz case

The model (1) is called *Lipschitz* if there are positive constants $a_k$ such that for $k = 1, \ldots m$, $y$ , $y' \in \mathbb{R}^d$ and a norm $|| \cdot ||$ on $\mathbb{R}^d$ :

$$||f_k(y) - f_k(y')|| \le a_k \, ||y - y'|| \, . \tag{16}$$

To solve the problem **A**, we apply the results of [3].

*Theorem 3* ***(Existence and uniqueness of a strictly stationary solution)*** *We consider the Lipschitz NAR-MS model (16) and assume*

*(i).* $\mathbb{E} \, \log^+ ||\varepsilon_1|| < \infty$ .

*(ii).* $\beta := \sum_{k=1}^{m} \mu_k \log a_k < 0.$

*Then,*

*(1) The model (1) has an unique strictly stationary and ergodic solution with stationary distribution* $\nu$.

*(2) For all* $(x,y) \in \mathcal{Z}$

$$\mathbb{P}_{(x,y)}\text{-}a.s., \quad \forall \varphi \in C_b(\mathbb{R}^d, \mathbb{R}), \quad \frac{1}{n}\sum_{i=1}^{n} \varphi(Y_i) \longrightarrow \nu(\varphi) \ . \tag{17}$$

**Proof.** *Step 1.* We start with a particular initial condition $X_0 \sim \mu$, $Y \equiv y$ with some fixed $y \in \mathbb{R}^d$. Then, the chain $\mathbf{X}$ is stationary and we can apply Theorem 3.1 of [3]. By extending the model (1) from $\mathbb{N}$ to $\mathbb{Z}$, the sequence

$$u \longmapsto \phi_n(u) = f_{X_n}(u) + \varepsilon_n = \phi_{X_n,\varepsilon_n}(u)$$

is a strictly stationary and ergodic sequence of random Lipschitz maps ([3]). Let $L(g)$ be the Lipschitz module of any Lipschitz map $g$ and $\log^+(x) = \max(0, \log x)$. Before checking the two assumptions used in that Theorem 3.1, we restate them below for clearness reasons:

(C1) For some $y \in \mathbb{R}^d$, $\mathbb{E}[\log^+ ||\phi_1(y) - y||]$ is finite.

(C2) The random variable $\log^+ L(\phi_1)$ is integrable, and for some integer $p > 0$, the real number

$$\alpha := p^{-1} \ \mathbb{E} \log L\left(\phi_p \circ \cdots \circ \phi_1\right)$$

is strictly negative.

To check (C1), first note that $\log^+(a+b) \leq \log^+ a + \log^+ b + \log 2$ for any $a > 0$, $b > 0$. Since

$$||\phi_1(y) - y|| \leq ||f_{X_1}(y)|| + ||\varepsilon_1|| + ||y|| \ ,$$

and $\mathbb{E} \log^+ ||\varepsilon_1|| < \infty$, (C1) is fulfilled with any $y \in \mathbb{R}^d$. For the first part of (C2), as $L(\phi_1) = L(f_{X_1}) \leq a_{X_1}$, the variable $\log^+ L(\phi_1)$ is clearly integrable. Furthermore,

$$(\phi_p \circ \cdots \circ \phi_1)(y) = f_{X_p}\left(f_{X_{p-1}}\left(\cdots\ (f_{X_1}(y) + \varepsilon_1) + \varepsilon_2\right) + \cdots + \varepsilon_{p-1}\right) + \varepsilon_p \ .$$

Therefore

$$L\left(\phi_p \circ \cdots \circ \phi_1\right) \leq a_{X_p} a_{X_{p-1}} \cdots a_{X_1} .$$

Since $(a_{X_n})$ is stationary, we have

$$\frac{1}{p} \mathbb{E} \log \left(a_{X_p} a_{X_{p-1}} \cdots a_{X_1}\right) \equiv \mathbb{E} \log a_X = \beta.$$

It follows from the assumption ((ii)) that there is some positive integer $p$ for which $\alpha < 0$. The condition (C2) is then also fulfilled. Hence, the existence and the uniqueness of a strictly stationary ergodic solution follows from that Theorem 3.1. Moreover, by applying Corollary 3.2 and 3.3 in [3], we have both

*(a)* $Y_n$ converges weakly to $\nu$ ;

*(b)* almost surely, for all $\varphi \in C_b(\mathbb{R}^d, \mathbb{R})$, $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(Y_i) = \nu(\varphi)$ .

*Step 2.* Let us go back to the general (non stationary) situation with an arbitrary initial condition $X_0 \equiv x$, $Y_0 = y$ for some $(x, y) \in \mathcal{Z}$. Let $\varphi \in C_b(\mathbb{R}^d, \mathbb{R})$ and we define the event

$$A_\varphi = \left\{ \omega \ : \ \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(Y_i) = \nu(\varphi) \right\} .$$

By the previous step, $\mathbb{P}_{(\mu,y)}(A_\varphi) = 1$. Moreover,

$$\mathbb{P}_{(\mu,y)} = \sum_{x=1}^{m} \mathbb{P}_{(x,y)}(A_\varphi)\mu(x) .$$

Since $\mathbf{X}$ is positive recurrent, $\mu(x) > 0$ for all $x$. Hence $\mathbb{P}_{(x,y)}(A_\varphi) = 1$ for all $(x, y)$. Since $\varphi$ is arbitrary, the proof of (ii) is completed. ■.

It is worth comparing Theorem 3 to Theorem 1 for the sublinear case. First, the moment condition $\mathbb{E}\,||\varepsilon_1||^\delta < \infty$ is weakened on $\mathbb{E} \log^+ ||\varepsilon_1|| < \infty$. This condition on the logarithmic moment of the error process is optimal since for an one-dimensional AR(1) process (i.e. $m = 1$, $d = 1$, $f_1(y) = ay$), it is known to be the weakest condition for existence of a strictly stationary solution (hence an i.p.m. for the chain $\mathbf{Z}$). Besides, the Lipschitz property enables a direct method to establish this existence and guarantees its uniqueness. It is for these reasons that we no longer assume that $\varepsilon_1$ has an everywhere positive density (which ensured the uniqueness of $\sigma$ in the sublinear case), nor that $\varepsilon_1$ has a density (which guaranteed that the chain $\mathbf{Z}$ is strongly Feller).

On the other hand, Theorem 3 does not implies the stability of the extended chain $\mathbf{Z}$. The existence of moments for the stationary distribution $\nu$ and a SLLN (3) applicable to the chain $\mathbf{Z}$ are also lacking.

The Lipschitz property of the model yields also the weak convergence of the marginal distributions of $(Y_n)$.

*Proposition 1* *Under the same assumptions as in Theorem 3 and for all initial condition* $(X_0 = x, Y_0 = y)$, *the distribution of* $Y_n$ *converges weakly to* $\nu$.

**Proof.** By Theorem 3, there is a strictly stationary and ergodic solution $\{Y_t\,,\, t \in \mathbb{Z}\}$, with stationary distribution $\nu$. Let $(x,y) \in \mathcal{Z}$ be fixed. The idea is to compare the two processes $\left\{Y_n^{(x,y)},\, n \in \mathbb{N}\right\}$ and $\left\{\overline{Y}_n^{(\mu,y)},\, n \in \mathbb{N}\right\}$, starting from $(X_0, Y_0) = (x,y)$ and $(X_0, y)$ with $X_0 \sim \mu$, respectively. The associated regime chains are denoted $\{X_n^x\}$ and $\{X_n^\mu\}$. As $\{X_n^\mu\}$ is stationary, Theorem 3.1 of [3] can be applied to $\left\{\overline{Y}_n^{(\mu,y)}\right\}$.

Furthermore, since the regime chain $\mathbf{X}$ is positive recurrent, there is a (successful) coupling ([13]) $\{(\mathbb{X}_n^x, \mathbb{X}_n^\mu)\}$ of the two processes $\{X_n^x\}$ et $\{X_n^\mu\}$ for which the coupling time

$$T = \inf\left\{n \ge 1\, : \mathbb{X}_j^x = \mathbb{X}_j^\mu \quad \text{for all } j \ge n\right\}\ ,$$

is a.s. finite. Let us consider for $n \ge 1$, the following NAR-MS processes associated to this coupling

$$\begin{aligned}
\mathbb{Y}_n^{(x,y)} &= f_{\mathbb{X}_n^x}\left(\mathbb{Y}_{n-1}^{(x,y)}\right) + \varepsilon_n = f_{\mathbb{X}_n^x,\varepsilon_n}\left(\mathbb{Y}_{n-1}^{(x,y)}\right) \quad , \quad \text{with } \ \mathbb{Y}_0^{(x,y)} = y\ ,\\
\mathbb{Y}_n^{(\mu,y)} &= f_{\mathbb{X}_n^\mu}\left(\mathbb{Y}_{n-1}^{(\mu,y)}\right) + \varepsilon_n = f_{\mathbb{X}_n^\mu,\varepsilon_n}\left(\mathbb{Y}_{n-1}^{(\mu,y)}\right), \quad \text{with } \ \mathbb{Y}_0^{(\mu,y)} = y\ .
\end{aligned}$$

We obtain

$$\mathbb{Y}_n^{(x,y)} - \mathbb{Y}_n^{(\mu,y)} = \begin{cases} \left(f_{\mathbb{X}_n^x,\varepsilon_n}\circ\cdots\circ f_{\mathbb{X}_1^x,\varepsilon_1} - f_{\mathbb{X}_n^\mu,\varepsilon_n}\circ\cdots\circ f_{\mathbb{X}_1^\mu,\varepsilon_1}\right)(y\ )\ , & \text{if } n \le T\\ \\ \left(f_{\mathbb{X}_n^\mu,\varepsilon_n}\circ\cdots\circ f_{\mathbb{X}_{T+1}^\mu,\varepsilon_{T+1}}\right)\left(\mathbb{Y}_T^{(x,y)} - \mathbb{Y}_T^{(\mu,y)}\right)\ , & \text{if } n > T \end{cases}\ .$$

Since $T$ is a.s. finite, it holds from Theorem 3.1 in [3] that for $n \ge T$

$$\text{a.s.} \qquad \limsup \frac{1}{n}\log||\mathbb{Y}_n^{(x,y)} - \mathbb{Y}_n^{(\mu,y)}|| \le \beta < 0\ .$$

This shows in particular that a.s., $\mathbb{Y}_n^{(x,y)} - \mathbb{Y}_n^{(\mu,y)} \to 0$.

On the other hand, by Corollary 3.3 in [3], $\mathbb{Y}_n^{(\mu,y)}$ converges weakly to $\nu$. Since $\mathbb{Y}_n^{(x,y)} - \mathbb{Y}_n^{(\mu,y)} \to 0$ a.s., $\mathbb{Y}_n^{(x,y)}$ also converges weakly to $\nu$. The same is still true for $Y_n^{(x,y)}$ , since the probability distributions of $\left(\mathbb{Y}_n^{(x,y)}\right)$ and $\left(Y_n^{(x,y)}\right)$ are the same. ■

Let us still denote by $Q_s$ the matrix (10) where the coefficients $a_k$'s are now the Lipschitz modules given in (16).

**Comment.** In [7], the existence of a strictly stationary solution is established under

$$\rho(Q_1) < 1 . \tag{18}$$

This condition is stronger than the condition ((ii)) used in Theorem 3. Indeed, by Lemma 2 (Appendix B), we have

$$\sum \mu_k \log a_k \leq \log \rho(Q_1) .$$

Let us examine this difference in more details. Let $m = 2$ and

$$P = \begin{pmatrix} 1-\tilde{\alpha} & \tilde{\alpha} \\ \tilde{\beta} & 1-\tilde{\beta} \end{pmatrix} .$$

The condition $\rho(Q_1) < 1$ becomes

$$\left\{ \begin{array}{rcl} (\tilde{\alpha}+\tilde{\beta}-1)\, a_1 a_2 + (1-\tilde{\alpha})a_1 + (1-\tilde{\beta})a_2 & < & 1 . \\ (1-\tilde{\alpha})a_1 + (1-\tilde{\beta})a_2 & \leq & 2 , \end{array} \right. \tag{19}$$

while the condition ((ii)) can be written as

$$\tilde{\beta} \log a_1 + \tilde{\alpha} \log a_2 < 0 . \tag{20}$$

Figure 1 depicts these conditions in three situations $\tilde{\alpha} = \tilde{\beta} = \frac{3}{4}$, $\tilde{\alpha} = \tilde{\beta} = \frac{2}{4}$ and $\tilde{\alpha} = \tilde{\beta} = \frac{1}{4}$. It is worth noting that under the condition (20), some of $a_k$'s are allowed to be arbitrary large while under the condition (19), all of $a_k$'s must be bounded. ■

The next result examines the existence of moments and a SLLN for unbounded functions.

*Theorem 4* (***Moments of order*** $s \geq 1$ ***and a SLLN***) *Let* $s \geq 1$. *We consider the Lipschitz NAR-MS model (1) and assume*

*(i')* $\mathbb{E}||\varepsilon_1||^s < \infty$.

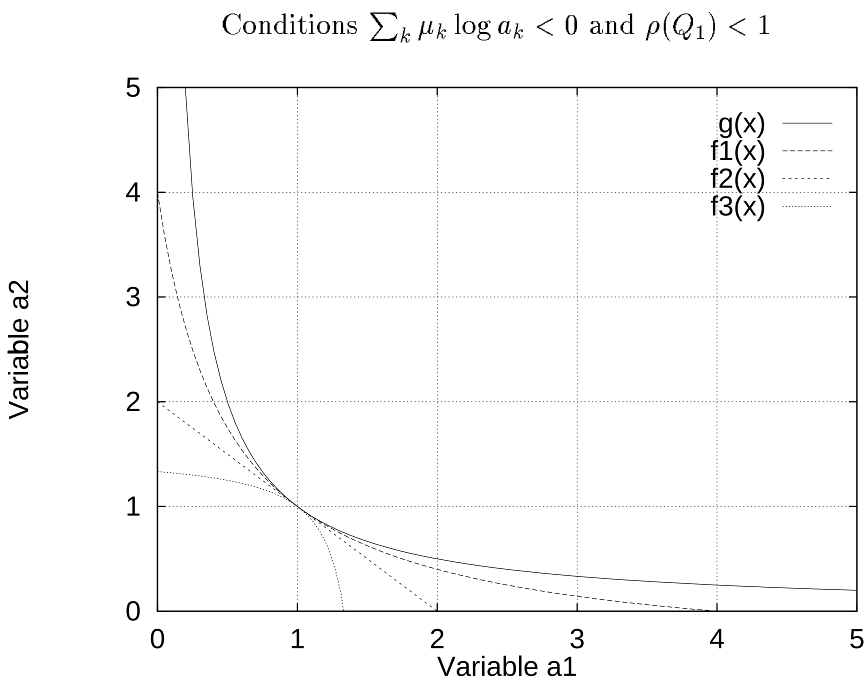


Figure 1. Plots of **g(x)**: $\tilde{\beta}\,\log a_1 + \tilde{\alpha}\log a_2 = 0$ and of ***f1(x), f2(x), f3(x)*** from Condition (12): $\rho(Q_1) = 1$ in three cases, respectively: (a). $\tilde{\alpha} = \tilde{\beta} = \frac{3}{4}$. (b). $\tilde{\alpha} = \tilde{\beta} = \frac{2}{4}$. (c). $\tilde{\alpha} = \tilde{\beta} = \frac{1}{4}$. Each of the acceptable domains (for $a_1$ and $a_2$) is included between the associated curve and the two axes $a_1$ and $a_2$.

*(ii') $\rho(Q_s) < 1$.*

*Then, in addition of the results given in Theorem 3, we have*

*(1) the unique stationary distribution $\nu$ of the process $\mathbf{Y}$ has a finite moment of order $s$, i.e. $\nu(||y||^s) < \infty$ ;*

*(2) the SLLN (17) is still valid for any continuous function $\varphi$ satisfying $\varphi(y) \le const.(1 + ||y||^s)$, $y \in \mathbb{R}^d$.*

**Proof.** The conclusions of Theorem 3 clearly hold. In particular the stationary distribution $\nu$ is unique. This implies that any i.p.m. $\sigma^*$ of the chain $\mathbf{Z}$ has the same marginal distributions $\mu$ and $\nu$.

*Part (1) :* we note that any Lipschitz map is sublinear and the contraction inequality (15) is still valid. This ensures that (see, e.g. Proposition 2.1.6 of [6]), almost surely, the family of probability measures $\{\Pi^k((x,y),dz)), k \geq 1\}$ is tight. Therefore there is some subsequence of this family which converges weakly to a probability measure $\sigma'$. Moreover, $\sigma'$ is necessarily an i.p.m of $\Pi$ and $\sigma'(||y||^s) < \infty$. Since $\nu$ is the $y$-marginal of $\sigma'$, we have $\nu(||y||^s) < \infty$.

*Part (2) :* From Theorem 3, there is a strictly stationary ergodic solution $(Y_n^{(\mu,\nu)})$ for the NAR-MS model (1) (for this solution, $X_0 \sim \mu$ and $Y_0 \sim \nu$). Furthermore, let $(Y_n^{(\mu,y)})$ be the process which starts from an arbitrary fixed point $Y_0 \equiv y$ while keeping the stationary regime $X_0 \sim \mu$. We have for all $y$ (see Theorem 3.1 in [3])

$$\text{almost surely,} \quad Y_n^{(\mu,y)} - Y_n^{(\mu,\nu)} \longrightarrow 0. \tag{21}$$

We will prove the result by induction on the order $s$. Let us first assume $s = 1$. Since $\nu(||y||) < \infty$, by the ergodic theorem,

$$\frac{1}{n}\sum_{i=1}^{n} ||Y_i^{(\mu,\nu)}|| \longrightarrow \nu(||y||), \quad a.s.$$

By (21),

$$\frac{1}{n}\sum_{i=1}^{n} \Big| ||Y_i^{(\mu,\nu)}|| - ||Y_i^{(\mu,y)}|| \Big| \longrightarrow 0, \quad a.s.$$

Hence for all $y$,

$$\frac{1}{n}\sum_{i=1}^{n} ||Y_i^{(\mu,y)}|| \longrightarrow \nu(||y||), \quad a.s. \tag{22}$$

Furthermore, it is already proved (see point (b) in the proof of Theorem 3), that for all $y \in \mathbb{R}^d$, $\mathbb{P}_{(\mu,y)}$-a.s., the sequence of empirical mesures of $(Y_n^{(\mu,y)})$ converges weakly to $\nu$. Therefore the above SLLN (22) can be extended to any continuous function $\varphi$ satisfying $|\varphi| \leq \text{const.}(1+||\cdot||)$ (see e.g. Proposition 2.1.6 of [6]). To conclude the case $s = 1$, the condition "for all $y \in \mathbb{R}^d$, $\mathbb{P}_{(\mu,y)}$-a.s.," can be replaced by "for all $(x,y) \in E \times \mathbb{R}^d$, $\mathbb{P}_{(x,y)}$-a.s.," exactly as at the end of the proof of Theorem 3.

Next, let us assume $1 < s \leq 2$. We have for any $y, y'$

$$\Big| ||y||^s - ||y'||^s \Big| \leq s||y - y'|| \left(1 + ||y||^{s-1} + ||y'||^{s-1}\right).$$

Therefore,

(23)

$$\frac{1}{n}\sum_{i=1}^{n}\Big|||Y_i^{(\mu,\nu)}||^s - ||Y_i^{(\mu,y)}||^s\Big| \leq \frac{s}{n}\sum_{i=1}^{n}||Y_i^{(\mu,\nu)} - Y_i^{(\mu,y)}||\left(1 + ||Y_i^{(\mu,\nu)}||^{s-1} + ||Y_i^{(\mu,y)}||^{s-1}\right).$$

Since $s-1 \leq 1$, by the previous step, both $\frac{1}{n}\sum_{i=1}^{n}||Y_i^{(\mu,y)}||^{s-1}$ and $\frac{1}{n}\sum_{i=1}^{n}||Y_i^{(\mu,\nu)}||^{s-1}$ converge almost surely to $\nu(||y||^{s-1})$. Hence by (21) and (23)

$$\frac{1}{n}\sum_{i=1}^{n}\Big|||Y_i^{(\mu,\nu)}||^s - ||Y_i^{(\mu,y)}||^s\Big| \longrightarrow 0, \quad a.s.$$

This is the main point in the current case and the remainning conclusions can be established exactly as in the previous case $s=1$.

It is easily seen that the induction above can be applied to an arbitrary order $s$. ■.

One should point out that under the conditions of Theorem 4, we are not able to establish the ergodicity of the extended process $\mathbf{Z}$. The reason is that we do not know whether the Markov chain $\mathbf{Z}$ could have more than one invariant probability measure. If instead, the noise $\boldsymbol{\varepsilon}$ has a positive density as in Theorem 2, this uniqueness is ensured and Theorem 2 applies (since any Lipschitz function is sublinear). This difficulty is well illustrated by the following example. Let $\mathbf{Y} := (Y_n)_{n\geq 0}$ be the standard nonlinear AR process - without Markov switching or equivalently, with $m=1$ regime-, defined on $\mathbb{R}$ by

$$Y_n = \frac{1}{2}Y_{n-1} + \varepsilon_n\,, \quad Y_n \in \mathbb{R}\,, \tag{24}$$

where $\boldsymbol{\varepsilon} := (\varepsilon_n)_{n\geq 0}$ is an i.i.d $\mathbb{R}$-valued sequence of Bernouilli $\mathcal{B}\left(\{0,1\},\frac{1}{2}\right)$. We are clearly under the conditions of Theorem 4. However, when starting in $\mathbb{Q}$, the chain concentrates in $\mathbb{Q}$ and when starting in $\mathbb{R}\backslash\mathbb{Q}$, it concentrates in $\mathbb{R}\backslash\mathbb{Q}$. Being not irreducible, the chain is not Harris recurrent.

Consequently, Theorems 3 and 4 are useful for a NAR-MS process generated by a *poor* noise process $(\varepsilon_n)$, a discrete noise for example.

**Acknowledgement**

The authors are grateful to Xavier GUYON for helpful comments on this work.

**References**

[1] ATTALI, J. (1999). *(1). Méthodes de stabilité pour des chaînes de Markov non felleriennes. (2). Sur quelques autres problèmes issus des réseaux de neurones.* Thèse, Université Paris I.

[2] BOROVKOV, A. (1992). On ergodicity of iterations of random Lipschitz transformations. *Russian Acad. Sci. Dokl. Math.* **45,** 551–555.

[3] BOUGEROL, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM. J. Control. Optimiz.* **31,** 942–959.

[4] BOUGEROL, P. AND PICARD, N. (1992). Strict stationarity of generalized autoregressive processes. *Ann. Proba.* **20,** 1714–1730.

[5] BRANDT, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Prob.* **18,** 211–220.

[6] DUFLO, M. (1997). *Random Iterative Models.* Springer-Verlag.

[7] FRANCQ, C. AND ROUSSIGNOL, M. Consistency of MLE for AR models with Markov switching. *Prépublication No. 10/96*. Université de Marne-La-Vallée, to appear in *Statistics.*

[8] GALLANT, A. R. (1987). *Nonlinear Statistical Models.* Wiley.

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

[10] HAMILTON, J. D. (1990). Analysis of time series subject to changes in regime. *J. Economet.* **45,** 39–70.

[11] HAMILTON, J. D. (1996). Specification testing in Markov-switching time-series models. *J. Economet.* **70,** 127–157.

[12] Holst, U., Lingren, G., Holst, J. and Thuvesholmen, M. (1994). Recursive estimation in switching autoregressions with a Markov regime. *J. Times Series Anal.* **15,** 489–506.

[13] Lindvall, T. (1992). *Lectures on Coupling Methods.* Wiley.

[14] McCulloch, R. E. and Tsay, R. S. (1994). Statistical analysis of econometric time series via Markov switching models. *J. Time Series Anal.* **15,** 523–539.

[15] Meyn, S. and Tweedie, R. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag, Berlin.

[16] Pötscher, B. M. and Prucha, L. R. (1997). *Dynamic Nonlinear Econometric Models.* Springer-Verlag.

[17] Tjøstheim, D. (1990). Nonlinear time series and Markov chains. *Adv. Appl. Prob.* **22,** 587–611.

[18] Tong, H. (1990). *Non-linear Time Series (A Dynamic System Approach).* Oxford Univ. Press, New York.

[19] Yao, J. (1998). On least squares estimation for stable nonlinear autoregressive processes. *Ann. Inst. Math. Statist.* to appear in.

## A. A Lyapounov criterion for stability and positive recurrence

For various definitions used in this appendix, we refer to [15] and [6]. The presentation is based on [6], Chap. 6.

Let $\mathbf{Z} = (Z_n)_{n\geq 0}$ be a Markov chain defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in a Polish (metric, separable and complete) space $(E, \mathcal{E})$ ($\mathcal{E}$ is the Borel algebra). It is said to be *stable* if there is a probability measure $\sigma$ on $E$ such that for all initial distribution $\nu_0$ of $Z_0$, it holds $\mathbb{P}$-almost surely that the sequence of empirical distributions

$$\Lambda_n(\omega, dz) = \frac{1}{n}\sum_{k=1}^{n} \delta_{Z_k(\omega)}(dz) \ , \tag{25}$$

converges weakly to $\sigma$. Here $\delta_w(dz)$ is the Dirac mass at $w$. In other words, the following SLLN holds

$$\text{(26)} \qquad \text{a.s.} \quad \forall f \in \mathcal{C}_b(E, \mathbb{R}), \quad \frac{1}{n} \sum_{k=1}^{n} f(Z_k(\omega)) \longrightarrow \sigma(f) \ ,$$

where $\mathcal{C}_b(E, \mathbb{R})$ denotes the space of bounded continuous functions from $E$ to $\mathbb{R}$. By using the canonical space $(E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}})$ associated to the Markov chain, it is stable if and only if the above SLLN holds $\mathbb{P}_z$-a.s. for all $z \in E$.

The chain $\mathbf{Z}$ is said to be *positive recurrent* if there is a probability measure $\sigma$ on $E$ such that for all bounded measurable function $f : E \to \mathbb{R}$,

$$\text{(27)} \qquad \text{a.s.} \quad \frac{1}{n} \sum_{k=1}^{n} f(Z_k) \longrightarrow \sigma(f) \ .$$

This positive recurrence does not depend on the topology of $E$ and it implies the stability.

Let $\pi(z, dy)$ be the transition probability kernel of the chain. The chain is a Feller chain if for any $g \in \mathcal{C}_b(E, \mathbb{R})$, $\pi g \in \mathcal{C}_b(E, \mathbb{R})$. It is a strong Feller chain if $\pi g \in \mathcal{C}_b(E, \mathbb{R})$ for any bounded measurable function $g$. A function $V : E \to \mathbb{R}$ is a Lyapounov function if it is continuous, positive and tends to infinity as $z$ tends to infinity. The following Lyapounov criterion is used throughout the paper. The formulation given here is due to [6] and [1].

*Theorem 5* ***(Lyapounov criterion of stability and recurrence)*** *Let* $\mathbf{Z}$ *be a Markov chain on* $E$ *with transition probability kernel* $\pi$. *Assume that*

*(i). The chain is Feller.*

*(ii). There is a Lyapounov function* $V$ *on* $E$, *an integer* $p \geq 1$ *and two constants* $\alpha \in [0, 1)$, $\beta \geq 0$ *for which*

$$\text{(28)} \qquad \pi^p V \leq \alpha V + \beta \ .$$

*(iii).* $\pi$ *has at most one i.p.m.*

*Then,* $\mathbf{Z}$ *is stable. Moreover the SLLN (26) also holds with* $\mathcal{C}_b(E, \mathbb{R})$ *replaced by*

$$\text{(29)} \qquad \mathcal{C}(V, \eta) := \left\{ f \ : \ E \to \mathbb{R} \text{ continuous s.t. } |f| \leq \text{const.} \left[1 + V^{1-\eta}\right] \right\} \ ,$$

*with some constant* $\eta > 0$.

*If in addition*

*(iv). The chain* $\mathbf{Z}$ *is strongly Feller,*

*then* $\mathbf{Z}$ *is positive recurrent and the SLLN (26) still holds with* $\mathcal{C}_b(E, \mathbb{R})$ *replaced by*

$$\mathcal{B}(V) := \{ f \ : \ E \to \mathbb{R} \ \ \textit{measurable s.t.} \ |f| \le \textit{const.} \ [1 + V] \} \ . \tag{30}$$

## B. An auxiliary lemma

*Lemma 2* *Let* $a_1, \ldots, a_m$ *be nonnegative numbers,* $P$ *some* $m \times m$ *transition probability matrix for which* $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)$ *is an i.p.m. and*

$$Q = (a_j \, p_{ij}) = \begin{pmatrix} a_1 \, p_{11} & \cdots & a_m \, p_{1m} \\ \vdots & \ddots & \vdots \\ a_1 \, p_{m1} & \cdots & a_m \, p_{mm} \end{pmatrix} \ . \tag{31}$$

*Then*

$$\sum_{k=1}^{m} \mu_k \, \log a_k \le \log \rho(Q) \ . \tag{32}$$

**Proof.** We use the norm $|\mathbf{u}|_1 := \sum |u_k|$ for $\mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}^m$ and the induced operator norm $|A|_1$ for any $m \times m$ real matrix $A$. Let $\mathbf{X} = (X_n, \, n \ge 0)$ be a *stationary* Markov chain with transition probability matrix $P$ (i.e. $X_0 \sim \mu$) and $\mathbf{w} = (a_1 \mu_1, \, \ldots, \, a_m \mu_m)^T$. A straightforward calculus ([7]) gives for $p \in \mathbb{N}^*$ ,

$$\mathbb{E} \, \left( a_{X_p} \ \cdots a_{X_1} \right) = \mathbf{w}^T \, Q^p \, \mathbf{1} = |Q^p \mathbf{w}|_1 \le |Q^p|_1 \, |\mathbf{w}|_1 \ . \tag{33}$$

Thus for $p \ge 1$,

$$\begin{aligned} & \limsup_p \frac{1}{p} \, \mathbb{E} \, \log \left( a_{X_p} \cdots a_{X_1} \right) \\ \le \quad & \limsup_p \frac{1}{p} \log \mathbb{E} \, \left( a_{X_p} \cdots a_{X_0} \right) \le \limsup_p \frac{1}{p} \, \left( \log |Q^p|_1 + \log \, |\mathbf{w}|_1 \right) = \log \rho \ . \end{aligned}$$

On the other hand, $(X_n)$ is stationary, so that

$$\frac{1}{p} \, \mathbb{E} \, \log \ \left( a_{X_p} \ \cdots \, a_{X_1} \right) \equiv \mathbb{E} \, \log a_X \ . \qquad \blacksquare$$