A Rademacher-Menchov approach for random coefficient bifurcating autoregressive processes

Bernard Bercu, Vassili Blandin

A RADEMACHER-MENCHOV APPROACH FOR RANDOM COEFFICIENT BIFURCATING AUTOREGRESSIVE PROCESSES

BERNARD BERCU AND VASSILI BLANDIN

Abstract. We investigate the asymptotic behavior of the least squares estimator of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on inherited and environmental effects, we establish the almost sure convergence of our estimates. In addition, we also prove a quadratic strong law and central limit theorems. Our approach mainly relies on asymptotic results for vector-valued martingales together with the well-known Rademacher-Menchov theorem.

1. Introduction

The purpose of this paper is to study random coefficient bifurcating autoregressive processes (RCBAR). One can see those processes in two different ways. The first one is to see them as random coefficient autoregressive processes (RCAR) adapted to binary tree structured data, the second one is to consider those processes as the association of RCAR processes and bifurcating autoregressive processes (BAR). BAR processes have been first studied by Cowan and Staudte [10] while RCAR processes have been first investigated by Nicholls and Quinn [26, 27]. The RCBAR structure allows us to reckon with environmental and inherited effects in order to better take into account the evolution of the characteristic under study. One shall see cell division as an example of binary tree structured data.

Let us detail what a RCBAR process is. The first individual is designated as the individual 1 and each individual n leads to individuals $2n$ and $2n+1$. The random variable X_n will stand for the characteristic under study of individual n. We can now make explicit the first-order RCBAR process which is given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} &= a_n X_n + \varepsilon_{2n}, \\ X_{2n+1} &= b_n X_n + \varepsilon_{2n+1}, \end{cases} \tag{1.1}$$

where the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$ represents the environmental effect while the random coefficient sequence (a_n, b_n) represents the inherited effect. The example of the cell division incites us to suppose that ε_{2n} ans ε_{2n+1} are correlated since the environmental effect on two sister cells can reasonably be seen as correlated. Denote by a and b the conditional means of the random coefficient sequences (a_n) and (b_n). Moreover, let c and d be the conditional means of the driven noises (ε_{2n}) and (ε_{2n+1}), respectively. If $\widehat{\theta}_n$ stands for the least squares estimator of the unknown vector of means $\theta^t = (a, b, c, d)$, we shall prove, under suitable assumptions on inherited and

2010 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F05, 60G42.

Key words and phrases. bifurcating autoregressive process; random coefficient; least squares; martingale; almost sure convergence; central limit theorem.

environmental effects, that $\widehat{\theta}_n$ converges almost surely to θ with the almost sure rate of convergence

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{2^n}\right) \qquad \text{a.s.}$$

In addition, we shall also establish the asymptotic normality

$$\sqrt{2^n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^{-1}L\Gamma^{-1})$$

where the matrices L and Γ will be explicitely calculated.

Our theoretical approach is motivated by experiments on the single celled organism *Escherichia coli* which reproduces by dividing itself into two poles, one being called the new pole, the other being called the old pole. We refer the reader to the pioneer work on statistical analysis of *Escherichia coli* carried out by Stewart et al. [30] and Guyon et al. [18], as well as to the recent contribution of De Saporta et al. [12, 14]. It was empirically shown in these statistical analysis of experimental data that some variables among cell lines, such as the life span of the cells, does not evolve in the same way whether it is related to the new or the old pole. The difference in the evolution leads us to consider an asymmetric RCBAR. Considering a RCBAR process instead of a BAR process [13] allows us to assume that the inherited effect is no more deterministic, as randomness often appears in nature. Moreover, we can consider both deterministic and random inherited effects since we also allow the random variables modeling the inherited effect to be deterministic, making this study usable for RCBAR as well as BAR.

Our goal is to investigate the asymptotic behavior of the least squares estimators of the unknown parameters of first-order RCBAR processes. In contrast with the previous work of Blandin [9] where the asymptotic behavior of weighted least squares estimators were investigated, we propose here to make use of a totally different strategy based on the standard least squares (LS) estimators together with the well-known Rademacher-Menchov theorem. The martingale approach for BAR processes has been first suggested by Bercu et al. [6], followed by De Saporta et al. [11]. We also refer the reader to Bercu and Blandin [5] for the study of bifurcating integer-valued autoregressive processes and to Bansaye [3] for the study of its asymptotic behavior, as well as to the recent contribution of Djellout and Bitseki Penda [8] on moderate deviation principles for the LS estimators of the unknown parameters of BAR processes. Our approach relies on the Rademacher-Menchov theorem which allows us to study the LS estimates in a different way as in de Saporta et al. [13]. In particular, it enables us to significantly reduce the moment assumptions on the random coefficient sequence (a_n, b_n) and on the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$. We shall also make use of the strong law of large numbers and the central limit theorem for martingales [16, 19] in order to study the asymptotic behavior of our LS estimates. The martingale approach of this paper has also been used by Basawa and Zhou [4, 37, 38].

Since several methods have been proposed for the study of BAR processes, we tried to take into consideration each of them. In this way, we took into account the classical BAR approach as used by Huggins and Basawa [21, 22] and by Huggins and Staudte [23] who investigated the evolution of cell diameters and lifetimes. We were also inspired by the bifurcating Markov chain approach brought in by Guyon [17] and applied by Delmas and Marsalle [15]. We also reckoned with the analogy with the Galton-Watson processes as in Delmas and Marsalle [15] and Heyde and Seneta [20]. Even though we chose to use LS estimates, different methods have been investigated for parameter estimation in RCAR processes. We have chosen to make use of the least squares approach instead of quasi-maximum likelihood method proposed by Aue et al. [1, 2] as well as Berkes et al. [7]. The reason why we have chosen LS estimates is towfold. On the one hand, the LS method allows us to avoid the maximization step of the quasi-likelihood method, which could be time-consuming and costly. On the other hand, except in Theorem 3 of [7], it is necessary to know the variance σ^2 of the driven noise, in order to estimate the RCA parameters. Moreover, σ^2 cannot be estimated by quasi-maximum likelihood method. Via our least squares approach, it is possible to consistently estimate all the conditional variances of the random coefficients and driven noises. We also refer the reader to Koul and Schick [24] for the M-estimation method, see also Vanecek [35] and Schick [29].

The paper is organized as follows. We will explain more accurately the model we will consider in Section 2, leading to Section 3 where we will give explicitly our LS estimates of the unknown parameters under study. The martingale point of view chosen in this paper will be highlighted in Section 4. All our results about the asymptotic behavior of our LS estimates will be stated in Section 5, in particular the almost sure convergence, the quadratic strong law and the asymptotic normality. Section 6 is devoted to the Rademacher-Menchov theorem. All technical proofs are postponed to the last sections.

2. Random coefficient bifurcating autoregressive processes

We will study the first-order RCBAR process given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} &= a_n X_n + \varepsilon_{2n}, \\ X_{2n+1} &= b_n X_n + \varepsilon_{2n+1}, \end{cases} \tag{2.1}$$

where X_1 is the ancestor of the process and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ is the driven noise of the process. We will suppose that $\mathbb{E}[X_1^{16}] < \infty$ and we will also assume that the two sequences (a_n, b_n) and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ are independent and identically distributed and that X_1, (a_n, b_n) and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ are mutually independent. RCBAR processes can be seen as a first-order random coefficient autoregressive process on a binary tree, each node of this tree representing an individual and the first node being the ancestor. For all $n \geq 0$, $\mathbb{G}_n$ will stand for the n-th generation, that is to say $\mathbb{G}_n = \{2^n, 2^n + 1, \ldots, 2^{n+1} - 1\}$. We will also denote by $\mathbb{T}_n$ the set of all individuals up to the n-th generation, namely

$$\mathbb{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_k.$$

One can see that the cardinality $|\mathbb{G}_n|$ of $\mathbb{G}_n$ is 2^n, while that of $\mathbb{T}_n$ is $2^{n+1}-1$. $\mathbb{G}_{r_n}$ will denote the generation of individual n with $r_n = [\log_2(n)]$ where $[x]$ stands for the integer part of x. Let us recall that the two offspring of individual n are individuals $2n$ and $2n+1$.

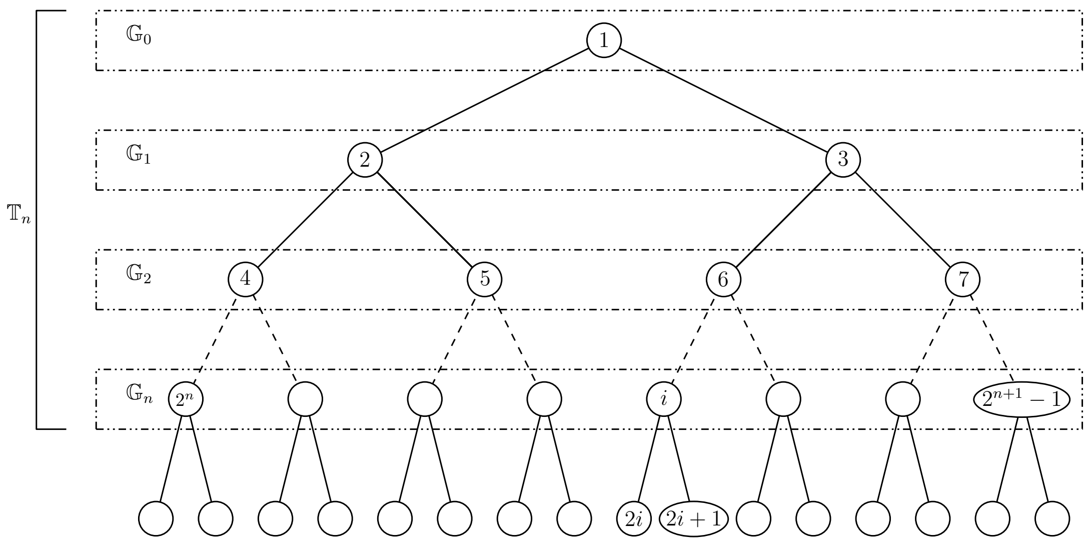

FIGURE 1. The tree associated with the RCBAR

3. LEAST SQUARES ESTIMATORS

Let $(\mathcal{F}_n)$ be the natural filtration associated with the generations of our first-order RCBAR (X_n), namely $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$ for all $n \in \mathbb{N}$. In all the sequel, we will assume that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{aligned} \mathbb{E}[a_k|\mathcal{F}_n] = a, \qquad \mathbb{E}[b_k|\mathcal{F}_n] = b,& \\ \mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c, \qquad \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d& \qquad \text{a.s.} \end{aligned} \tag{3.1}$$

Consequently, (2.1) can be rewritten as

$$\begin{cases} X_{2n} &= aX_n + c + V_{2n}, \\ X_{2n+1} &= bX_n + d + V_{2n+1}, \end{cases} \tag{3.2}$$

where, for all $k \in \mathbb{G}_n$, $V_{2k} = X_{2k} - E[X_{2k}|\mathcal{F}_n]$ and $V_{2k+1} = X_{2k+1} - E[X_{2k+1}|\mathcal{F}_n]$. We can rewrite the system (3.2) in a classic autoregressive form

$$\chi_n = \theta^t \Phi_n + W_n \tag{3.3}$$

where

$$\chi_n = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}, \qquad \Phi_n = \begin{pmatrix} X_n \\ 1 \end{pmatrix}, \qquad W_n = \begin{pmatrix} V_{2n} \\ V_{2n+1} \end{pmatrix},$$

and the matrix parameter θ given by

$$\theta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

One of our goal is to estimate θ from the observation of the $n+1$ first generations, namely $\mathbb{T}_n$. We will use the least squares estimator $\widehat{\theta}_n$ of θ which minimizes

$$\Delta_n(\theta) = \sum_{k \in \mathbb{T}_{n-1}} \|\chi_k - \theta^t \Phi_k\|^2.$$

Hence, we clearly have

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \Phi_k \chi_k^t \qquad \text{where} \qquad S_n = \sum_{k \in \mathbb{T}_n} \Phi_k \Phi_k^t. \tag{3.4}$$

In order to avoid any invertibility assumption, we will suppose that S_n is invertible. Otherwise, we only have to add the identity matrix of order 2, I_2, to S_n. Moreover, we will make a slight abuse of notation by identifying θ and $\widehat{\theta}_n$ to

$$\operatorname{vec}(\theta) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} \qquad \text{and} \qquad \operatorname{vec}(\widehat{\theta}_n) = \begin{pmatrix} \widehat{a}_n \\ \widehat{c}_n \\ \widehat{b}_n \\ \widehat{d}_n \end{pmatrix}.$$

In this vectorial form, we have

$$\widehat{\theta}_n = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k X_{2k} \\ X_{2k} \\ X_k X_{2k+1} \\ X_{2k+1} \end{pmatrix},$$

where $\Sigma_n = I_2 \otimes S_n$ and $\otimes$ stands for the standard Kronecker product. Hence, (3.3) yields to

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \tag{3.5}$$

In all this paper, we will make use of the following hypotheses on the moments of the random coefficient sequence (a_n, b_n) and on the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$. One can observe that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$, the random coefficients a_k, b_k and the driven noise ε_{2k}, ε_{2k+1} are $\mathcal{F}_{n+1}$-measurable.

(H.1) For all $n \geq 1$,

$$\mathbb{E}[a_n^{16}] < 1 \qquad \text{and} \qquad \mathbb{E}[b_n^{16}] < 1,$$
$$\mathbb{E}[\varepsilon_{2n}^{16}] < \infty \qquad \text{and} \qquad \mathbb{E}[\varepsilon_{2n+1}^{16}] < \infty.$$

(H.2) For all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\operatorname{Var}[a_k | \mathcal{F}_n] = \sigma_a^2 \geq 0 \quad \text{and} \quad \operatorname{Var}[b_k | \mathcal{F}_n] = \sigma_b^2 \geq 0 \qquad \text{a.s.}$$

$$\operatorname{Var}[\varepsilon_{2k} | \mathcal{F}_n] = \sigma_c^2 > 0 \quad \text{and} \quad \operatorname{Var}[\varepsilon_{2k+1} | \mathcal{F}_n] = \sigma_d^2 > 0 \qquad \text{a.s.}$$

(H.3) It exists $\rho_{ab}^2 \leq \sigma_a^2\sigma_b^2$ and $\rho_{cd}^2 < \sigma_c^2\sigma_d^2$ such that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(a_k - a)(b_k - b)|\mathcal{F}_n] = \rho_{ab} \qquad \text{a.s.}$$

$$\mathbb{E}[(\varepsilon_{2k} - c)(\varepsilon_{2k+1} - d)|\mathcal{F}_n] = \rho_{cd} \qquad \text{a.s.}$$

Moreover, for all $n \geq 0$ and $k, l \in \mathbb{G}_n$ with $k \neq l$, $(\varepsilon_{2k}, \varepsilon_{2k+1})$ and $(\varepsilon_{2l}, \varepsilon_{2l+1})$ as well as (a_k, b_k) and (a_l, b_l) are conditionally independent given $\mathcal{F}_n$.

(H.4) One can find $\mu_a^4 \geq \sigma_a^4$, $\mu_b^4 \geq \sigma_b^4$, $\mu_c^4 > \sigma_c^4$ and $\mu_d^4 > \sigma_d^4$ such that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[(a_k - a)^4 \,|\mathcal{F}_n\right] = \mu_a^4 \qquad \text{and} \qquad \mathbb{E}\left[(b_k - b)^4 \,|\mathcal{F}_n\right] = \mu_b^4 \qquad \text{a.s.}$$

$$\mathbb{E}\left[(\varepsilon_{2k} - c)^4 \,|\mathcal{F}_n\right] = \mu_c^4 \qquad \text{and} \qquad \mathbb{E}\left[(\varepsilon_{2k+1} - d)^4 \,|\mathcal{F}_n\right] = \mu_d^4 \qquad \text{a.s.}$$

$$\mathbb{E}[\varepsilon_{2k}^4] > \mathbb{E}[\varepsilon_{2k}^2]^2 \qquad \text{and} \qquad \mathbb{E}[\varepsilon_{2k+1}^4] > \mathbb{E}[\varepsilon_{2k+1}^2]^2.$$

In addition, it exists $\nu_{ab}^2 \geq \rho_{ab}^2$ and $\nu_{cd}^2 > \rho_{cd}^2$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(a_k - a)^2(b_k - b)^2|\mathcal{F}_n] = \nu_{ab}^2 \qquad \text{and} \qquad \mathbb{E}[(\varepsilon_{2k} - c)^2(\varepsilon_{2k+1} - d)^2|\mathcal{F}_n] = \nu_{cd}^2 \qquad \text{a.s.}$$

(H.5) There exists some $\alpha > 4$ such that

$$\sup_{n\geq 0}\sup_{k\in\mathbb{G}_n} \mathbb{E}[|a_k - a|^\alpha|\mathcal{F}_n] < \infty, \qquad \sup_{n\geq 0}\sup_{k\in\mathbb{G}_n} \mathbb{E}[|b_k - b|^\alpha|\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

$$\sup_{n\geq 0}\sup_{k\in\mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k} - c|^\alpha|\mathcal{F}_n] < \infty, \qquad \sup_{n\geq 0}\sup_{k\in\mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k+1} - d|^\alpha|\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

One can observe that hypothesis **(H.2)** allows us to consider a classical BAR process where $a_k = a$ and $b_k = b$ a.s. Moreover, under assumptions **(H.2)** and **(H.3)**, we have for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2, \qquad \mathbb{E}[V_{2k+1}^2|\mathcal{F}_n] = \sigma_b^2 X_k^2 + \sigma_d^2 \qquad \text{a.s.} \tag{3.6}$$

$$\mathbb{E}[V_{2k}V_{2k+1}|\mathcal{F}_n] = \rho_{ab} X_k^2 + \rho_{cd} \qquad \text{a.s.} \tag{3.7}$$

We deduce from (3.6) that, for all $n \geq 1$, $V_{2n}^2 = \eta^t \psi_n + v_{2n}$ where $v_{2n} = V_{2n}^2 - \mathbb{E}\left[V_{2n}^2|\mathcal{F}_{r_n}\right]$,

$$\eta = \begin{pmatrix}\sigma_a^2 & \sigma_c^2\end{pmatrix}^t \qquad \text{and} \qquad \psi_n = \begin{pmatrix}X_n^2 & 1\end{pmatrix}^t.$$

It leads us to estimate the vector of variances η by the least squares estimator

$$\widehat{\eta}_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \widehat{V}_{2k}^2 \psi_k \qquad \text{where} \qquad Q_n = \sum_{k\in\mathbb{T}_n} \psi_k \psi_k^t \tag{3.8}$$

and for all $k \in \mathbb{G}_n$,

$$\begin{cases} \widehat{V}_{2k} &= X_{2k} - \widehat{a}_n X_k - \widehat{c}_n, \\ \widehat{V}_{2k+1} &= X_{2k+1} - \widehat{b}_n X_k - \widehat{d}_n. \end{cases}$$

We clearly have a similar expression for the estimator of the vector of variances $\zeta = \begin{pmatrix}\sigma_b^2 & \sigma_d^2\end{pmatrix}^t$ by replacing $\widehat{V}_{2k}$ by $\widehat{V}_{2k+1}$ into (3.8). It also follows from (3.7) that, for all $n \geq 1$, $V_{2n}V_{2n+1} = \nu^t \psi_n + w_{2n}$ where $w_{2n} = V_{2n}V_{2n+1} - \mathbb{E}\left[V_{2n}V_{2n+1}|\mathcal{F}_{r_n}\right]$ and ν is the vector of covariances

$$\nu = \begin{pmatrix}\rho_{ab} & \rho_{cd}\end{pmatrix}^t.$$

Therefore, we can estimate ν by

$$\widehat{\nu}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \widehat{V}_{2k} \widehat{V}_{2k+1} \psi_k. \tag{3.9}$$

4. A martingale approach

We already saw that relation (3.5) can be rewritten as

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n \qquad \text{where} \qquad M_n = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \tag{4.1}$$

The key point of this study is to remark that (M_n) is a locally square integrable martingale, which allows us to make use of asymptotic results for martingales. This justifies our vectorial notation introduced previously since most of those asymptotic results have been established for vector-valued martingales. In order to study this martingale, let us rewrite M_n in a more convenient way. Let $\Psi_n = I_2 \otimes \varphi_n$ where φ_n is the 2×2^n matrix given by

$$\varphi_n = \begin{pmatrix} X_{2^n} & X_{2^n+1} & \dots & X_{2^{n+1}-1} \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

The first line of φ_n gathers the individuals of the n-th generation, φ_n can also be seen as the collection of all Φ_k, $k \in \mathbb{G}_n$. Let ξ_n be the random vector of dimension 2^n gathering the noise variables of $\mathbb{G}_n$, namely

$$\xi_n^t = \begin{pmatrix} V_{2^n} & V_{2^n+2} & \dots & V_{2^{n+1}-2} & V_{2^n+1} & V_{2^n+3} & \dots & V_{2^{n+1}-1} \end{pmatrix}.$$

The special ordering separating odd and even indices has been made in Bercu et al. [6] in order to rewrite M_n as

$$M_n = \sum_{k=1}^{n} \Psi_{k-1} \xi_k.$$

It clearly follows from **(H.1)** to **(H.3)** that (M_n) is a locally square integrable martingale with increasing process given, for all $n \geq 1$, by

$$\langle M \rangle_n = \sum_{k=0}^{n-1} \Psi_k \mathbb{E}[\xi_{k+1} \xi_{k+1}^t | \mathcal{F}_k] \Psi_k^t = \sum_{k=0}^{n-1} L_k \qquad \text{a.s.} \tag{4.2}$$

where

$$L_n = \sum_{k \in \mathbb{G}_n} \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix} \tag{4.3}$$

with

$$\begin{cases} P(X) = \sigma_a^2 X^2 + \sigma_c^2, \\ Q(X) = \rho_{ab} X^2 + \rho_{cd}, \\ R(X) = \sigma_b^2 X^2 + \sigma_d^2. \end{cases} \tag{4.4}$$

The first step of our approach will be to establish the convergence of $\langle M \rangle_n$ properly normalized, from which we will be able to deduce several asymptotic results for our RCBAR estimates.

5. Main results

Lemma 5.1. *Assume that* **(H.1)** *is satisfied. Then, we have for all* $p \in \{1, 2, \ldots, 8\}$,

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} X_k^p = s_p \qquad a.s. \tag{5.1}$$

where s_p *is a constant depending only on the moments of* a_1, b_1, ε_2 *and* ε_3 *up to the* p*-th order.*

Remark 5.2. *In particular, we have*

$$s_1 = \frac{c+d}{2-(a+b)},$$

$$s_2 = \frac{2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\left(\frac{(ac+bd)(c+d)}{2-(a+b)} + \frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2}\right),$$

and explicit expressions for s_3 *to* s_8 *are given at the end of Section 7.*

Proposition 5.3. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \frac{\langle M \rangle_n}{|\mathbb{T}_{n-1}|} = L \qquad a.s. \tag{5.2}$$

where L *is the positive definite matrix given by*

$$L = \begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix} \otimes C + \begin{pmatrix} \sigma_a^2 & \rho_{ab} \\ \rho_{ab} & \sigma_b^2 \end{pmatrix} \otimes D,$$

where

$$C = \begin{pmatrix} s_2 & s_1 \\ s_1 & 1 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} s_4 & s_3 \\ s_3 & s_2 \end{pmatrix}. \tag{5.3}$$

Remark 5.4. *One can observe in the proof of Lemma 5.1 that we only need to assume for convergence* (5.2) *that*

$$\mathbb{E}[a_n^8] < 1, \qquad \mathbb{E}[b_n^8] < 1, \qquad \sup_{n\geq 1} \mathbb{E}[\varepsilon_{2n}^8] < \infty, \qquad \sup_{n\geq 1} \mathbb{E}[\varepsilon_{2n+1}^8] < \infty.$$

Our first result deals with the almost sure convergence of the LS estimator $\widehat{\theta}_n$. We will denote by $\|x\|$ the euclidean norm of a vector x.

Theorem 5.5. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then,* $\widehat{\theta}_n$ *converges almost surely to* θ *with the almost sure rate of convergence*

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s.$$

In addition, we also have the quadratic strong law

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |\mathbb{T}_{k-1}|(\widehat{\theta}_k - \theta)^t \Gamma \Lambda^{-1} \Gamma (\widehat{\theta}_k - \theta) = Tr(\Lambda^{-1/2} L \Lambda^{-1/2}) \qquad a.s. \tag{5.4}$$

where

$$\Lambda = I_2 \otimes (C+D) \qquad \textit{and} \qquad \Gamma = I_2 \otimes C.$$

Our second result concerns the almost sure asymptotic properties of our least squares variance and covariance estimators $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$. We need to introduce some new variables

$$\eta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k}^2 \psi_k, \qquad \zeta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k+1}^2 \psi_k, \qquad \nu_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k}V_{2k+1} \psi_k.$$

Theorem 5.6. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then,* $\widehat{\eta}_n$ *and* $\widehat{\zeta}_n$ *both converge almost surely to* η *and* ζ *respectively. More precisely,*

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{5.5}$$

$$\|\widehat{\zeta}_n - \zeta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{5.6}$$

In addition, $\widehat{\nu}_n$ *converges almost surely to* ν *with*

$$\|\widehat{\nu}_n - \nu_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{5.7}$$

Remark 5.7. *We also have the less precise almost sure rates of convergence to the true parameters*

$$\|\widehat{\eta}_n-\eta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\zeta}_n-\zeta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\nu}_n-\nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s.$$

Finally, our last result is devoted to the asymptotic normality of our least squares estimates $\widehat{\theta}_n$, $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$.

Theorem 5.8. *Assume that* **(H.1)** *to* **(H.5)** *are satisfied. Then, we have the asymptotic normality*

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^{-1}L\Gamma^{-1}). \tag{5.8}$$

In addition, we also have

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}M_{ac}A^{-1}), \tag{5.9}$$

$$\sqrt{|\mathbb{T}_{n-1}|}\,\left(\widehat{\zeta}_n - \zeta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}M_{bd}A^{-1}), \tag{5.10}$$

where

$$A = \begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix},$$

$$M_{ac} = (\mu_a^4 - \sigma_a^4)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + 4\sigma_a^2\sigma_c^2\begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\mu_c^4 - \sigma_c^4)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix},$$

$$M_{bd} = (\mu_b^4 - \sigma_b^4)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + 4\sigma_b^2\sigma_d^2\begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\mu_d^4 - \sigma_d^4)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix}.$$

Finally,

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\nu}_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, A^{-1}HA^{-1}\right) \tag{5.11}$$

where

$$H = (\nu_{ab}^2 - \rho_{ab}^2)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + (\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2 + 2\rho_{ab}\rho_{cd})\begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\nu_{cd}^2 - \rho_{cd}^2)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix}.$$

The rest of the paper is dedicated to the proof of our main results.

6. On the Rademacher-Menchov theorem

Our almost sure convergence results rely on the Rademacher-Menchov theorem for orthonormal sequences of random variables given by Rademacher [28] and Menchoff [25], see Stout [31] and also Tandori [33, 34] and an unpublished note of Talagrand [32] for some extensions of this result.

Theorem 6.1. *Let (X_n) be an orthonormal sequence of square integrable random variables which means that for all $n \neq k$, $\mathbb{E}[X_nX_k] = 0$ and $\mathbb{E}[X_n^2] = 1$. Assume that a sequence of real numbers (a_n) satisfies*

$$\sum_{n=1}^{\infty} a_n^2 (\log n)^2 < \infty. \tag{6.1}$$

Then, the following series converges almost surely

$$\sum_{n=1}^{\infty} a_n X_n. \tag{6.2}$$

Remark 6.2. *One can observe that (X_n) is a square integrable sequence but is neither a sequence of independent random variables nor a sequence of uncorrelated random variables since (X_n) is not necessarily centered. In addition, in the case where (X_n) is an orthogonal sequence of random variables, we have the same result* (6.2)*, replacing* (6.1) *by*

$$\sum_{n=1}^{\infty} a_n^2 \mathbb{E}[X_n^2] (\log n)^2 < \infty.$$

If $a_n = 1/n$, it follows from (6.2) and Kronecker's lemma that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = 0 \qquad \text{a.s.}$$

7. Proof of the keystone Lemma 5.1

We shall introduce some suitable notations. Let (β_n) be the sequence defined, for all $n \geq 1$, by $\beta_{2n} = a_n$ and $\beta_{2n+1} = b_n$. Then, (2.1) can be rewritten as

$$\begin{cases} X_{2n} &= \beta_{2n}X_n + \varepsilon_{2n}, \\ X_{2n+1} &= \beta_{2n+1}X_n + \varepsilon_{2n+1}. \end{cases}$$

Consequently, for all $n \geq 2$

$$X_n = \beta_n X_{[\frac{n}{2}]} + \varepsilon_n.$$

First of all, let us prove that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} L_n = s_1 \qquad \text{where} \qquad L_n = \sum_{k\in\mathbb{T}_n} X_k.$$

One can observe that

$$\begin{aligned} L_n &= X_1 + \sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \Big(\beta_k X_{[\frac{k}{2}]} + \varepsilon_k\Big) \\ &= X_1 + (a+b)L_{n-1} + A_{n-1} + B_{n-1} + E_{n-1}, \end{aligned}$$

where

$$A_n = \sum_{k\in\mathbb{T}_n} X_k(a_k - a), \qquad B_n = \sum_{k\in\mathbb{T}_n} X_k(b_k - b), \qquad E_n = \sum_{k\in\mathbb{T}_n} (\varepsilon_{2k} + \varepsilon_{2k+1}).$$

Hence, we obtain that

$$\begin{aligned} \frac{L_n}{2^{n+1}} &= \frac{X_1}{2^{n+1}} + \frac{a+b}{2}\frac{L_{n-1}}{2^n} + \frac{A_{n-1}}{2^{n+1}} + \frac{B_{n-1}}{2^{n+1}} + \frac{E_{n-1}}{2^{n+1}} \\ &= \left(\frac{a+b}{2}\right)^n \frac{L_0}{2} + \sum_{k=1}^{n} \left(\frac{a+b}{2}\right)^{n-k} \left(\frac{X_1}{2^{k+1}} + \frac{A_{k-1}}{2^{k+1}} + \frac{B_{k-1}}{2^{k+1}} + \frac{E_{k-1}}{2^{k+1}}\right). \end{aligned} \tag{7.1}$$

Recalling that $|\mathbb{T}_n| = 2^{n+1}-1$, the standard strong law of large numbers immediately implies that

$$\lim_{n\to\infty} \frac{E_n}{2^{n+1}} = \mathbb{E}[\varepsilon_2 + \varepsilon_3] = c + d \qquad \text{a.s.}$$

Let us tackle the limit of A_n using the Rademacher-Menchov theorem given in Theorem 6.1. Let Y_n and R_n be defined as

$$Y_n = X_n(a_n - a) \qquad \text{and} \qquad R_n = \sum_{k=1}^{n} Y_k.$$

For all $n \geq 0$ and for all $k \in \mathbb{G}_n$, $\mathbb{E}[a_k - a|\mathcal{F}_n] = 0$. Moreover, we clearly have for all $n \geq 2$ and for all different $k, l \in \mathbb{G}_n$,

$$\begin{aligned} \mathbb{E}[Y_k Y_l] &= \mathbb{E}\left[\mathbb{E}[X_k X_l (a_k - a)(a_l - a)|\mathcal{F}_n]\right] \\ &= \mathbb{E}\left[X_k X_l \mathbb{E}[a_k - a|\mathcal{F}_n]\mathbb{E}[a_l - a|\mathcal{F}_n]\right] = 0. \end{aligned}$$

It means that (Y_n) is a sequence of orthogonal random variables. In addition we have, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{aligned} \mathbb{E}[Y_k^2] &= \mathbb{E}\left[\mathbb{E}[X_k^2(a_k - a)^2|\mathcal{F}_n]\right] \\ &= \mathbb{E}\left[X_k^2 \mathbb{E}[(a_k - a)^2|\mathcal{F}_n]\right] = \sigma_a^2 \mathbb{E}[X_k^2]. \end{aligned}$$

In order to calculate $\mathbb{E}[X_n^2]$, let us remark, with the convention that a product over an empty set is equal to 0, that for all $n \geq 1$,

$$X_n = \left(\prod_{k=0}^{r_n-1} \beta_{\left[\frac{n}{2^k}\right]}\right) X_1 + \sum_{k=0}^{r_n-1} \left(\prod_{i=0}^{k-1} \beta_{\left[\frac{n}{2^i}\right]}\right) \varepsilon_{\left[\frac{n}{2^k}\right]}.$$

Consequently,

$$\begin{aligned}\mathbb{E}[X_n^2] = \mathbb{E}\left[\left(\prod_{k=0}^{r_n-1} \beta^2_{\left[\frac{n}{2^k}\right]}\right) X_1^2\right] + 2\sum_{k=0}^{r_n-1} \mathbb{E}\left[\left(\prod_{l=0}^{r_n-1} \beta_{\left[\frac{n}{2^l}\right]}\right) X_1 \left(\prod_{i=0}^{k-1} \beta_{\left[\frac{n}{2^i}\right]}\right) \varepsilon_{\left[\frac{n}{2^k}\right]}\right] \\ + \mathbb{E}\left[\left(\sum_{k=0}^{r_n-1} \left(\prod_{i=0}^{k-1} \beta_{\left[\frac{n}{2^i}\right]}\right) \varepsilon_{\left[\frac{n}{2^k}\right]}\right)^2\right].\end{aligned}$$

First of all,

$$\mathbb{E}\left[\left(\prod_{k=0}^{r_n-1} \beta^2_{\left[\frac{n}{2^k}\right]}\right) X_1^2\right] = \mathbb{E}[X_1^2] \prod_{k=0}^{r_n-1} \mathbb{E}\left[\beta^2_{\left[\frac{n}{2^k}\right]}\right] \leq \mathbb{E}[X_1^2] \max(\mathbb{E}[a_1^2], \mathbb{E}[b_1^2])^{r_n} \leq \mathbb{E}[X_1^2].$$

Next, for the cross term

$$\begin{aligned}&\left|\sum_{k=0}^{r_n-1} \mathbb{E}\left[\left(\prod_{l=0}^{r_n-1} \beta_{\left[\frac{n}{2^l}\right]}\right) X_1 \left(\prod_{i=0}^{k-1} \beta_{\left[\frac{n}{2^i}\right]}\right) \varepsilon_{\left[\frac{n}{2^k}\right]}\right]\right| \\ &= \left|\mathbb{E}[X_1] \sum_{k=0}^{r_n-1} \left(\prod_{i=0}^{k-1} \mathbb{E}\left[\beta^2_{\left[\frac{n}{2^i}\right]}\right]\right) \left(\prod_{l=k+1}^{r_n-1} \mathbb{E}\left[\beta_{\left[\frac{n}{2^l}\right]}\right]\right) \mathbb{E}\left[\beta_{\left[\frac{n}{2^k}\right]} \varepsilon_{\left[\frac{n}{2^k}\right]}\right]\right| \\ &\leq \mathbb{E}[|X_1|] \max(|ac|, |bd|) \frac{\max(|a|, |b|)^{r_n} - \max(\mathbb{E}[a_1^2], \mathbb{E}[b_1^2])^{r_n}}{\max(|a|, |b|) - \max(\mathbb{E}[a_1^2], \mathbb{E}[b_1^2])} \\ &\leq \mathbb{E}[|X_1|] \max(|ac|, |bd|) \frac{1}{|\max(|a|, |b|) - \max(\mathbb{E}[a_1^2], \mathbb{E}[b_1^2])|}.\end{aligned}$$

By the same token, it is not hard to see that the last term is also bounded. Consequently, we proved that it exists some positive constant μ such that, for all $n \geq 0$, $\mathbb{E}[X_n^2] \leq \mu$, leading to

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \mathbb{E}[Y_n^2] (\log n)^2 \leq \sigma_a^2 \mu \sum_{n=1}^{\infty} \frac{(\log n)^2}{n^2} < \infty.$$

Therefore, it follows from the Rademacher-Menchov theorem that the series

$$\sum_{n=1}^{\infty} \frac{1}{n} Y_n$$

converges a.s. Consequently, Kronecker's lemma implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} Y_k = \lim_{n\to\infty} \frac{1}{n} R_n = 0 \qquad \text{a.s.}$$

In particular

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} R_{|\mathbb{T}_n|} = \lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} A_n = 0 \qquad \text{a.s.}$$

Hence, we find that

$$\lim_{n\to\infty} \frac{1}{2^{n+1}} A_n = 0 \qquad \text{a.s.}$$

By the same token, we also have

$$\lim_{n\to\infty} \frac{1}{2^{n+1}} B_n = 0 \qquad \text{a.s.}$$

To sum up, we obtain that

$$\lim_{n\to\infty} \frac{X_1}{2^{n+1}} + \frac{A_{n-1}}{2^{n+1}} + \frac{B_{n-1}}{2^{n+1}} + \frac{E_{n-1}}{2^{n+1}} = \frac{c+d}{2} \qquad \text{a.s.} \tag{7.2}$$

Therefore, we deduce from (7.1) and (7.2) together with the assumption that $\max(|a|,|b|) < 1$ and Lemma A.3 of [6], that

$$\lim_{n\to\infty} \frac{L_n}{2^{n+1}} = \frac{c+d}{2} \frac{1}{1 - \dfrac{a+b}{2}} \qquad \text{a.s.} \tag{7.3}$$

which means that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} X_k = \frac{c+d}{2-(a+b)} \qquad \text{a.s.}$$

Let us now tackle the study of

$$K_n = \sum_{k\in\mathbb{T}_n} X_k^2.$$

First, one can observe that

$$\begin{aligned}
K_n &= \sum_{k\in\mathbb{T}_n} X_k^2 = X_1^2 + \sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \left(\beta_k X_{[\frac{k}{2}]} + \varepsilon_k\right)^2 \\
&= X_1^2 + \left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \beta_k^2 X_{[\frac{k}{2}]}^2\right) + 2\left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \beta_k \varepsilon_k X_{[\frac{k}{2}]}\right) + \left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \varepsilon_k^2\right) \\
&= X_1^2 + (\sigma_a^2 + \sigma_b^2 + a^2 + b^2) K_{n-1} + 2(ac+bd) L_{n-1} + A_{n-1} + B_{n-1} + E_{n-1},
\end{aligned}$$

where

$$A_n = \sum_{k\in\mathbb{T}_n} X_k^2 (a_k^2 + b_k^2 - (\sigma_a^2 + \sigma_b^2 + a^2 + b^2)),$$

$$B_n = \sum_{k\in\mathbb{T}_n} X_k (a_k \varepsilon_{2k} + b_k \varepsilon_{2k+1} - (ac+bd)) \qquad \text{and} \qquad E_n = \sum_{k\in\mathbb{T}_n} (\varepsilon_{2k}^2 + \varepsilon_{2k+1}^2).$$

Hence, we obtain, as for L_n

$$\frac{K_n}{2^{n+1}} = \mu^n \frac{K_0}{2} + \sum_{k=1}^{n} \mu^{n-k} \left(\frac{X_1^2}{2^{k+1}} + \nu \frac{L_{k-1}}{2^k} + \frac{A_{k-1}}{2^{k+1}} + \frac{B_{k-1}}{2^{k+1}} + \frac{E_{k-1}}{2^{k+1}}\right),$$

where, since $\mathbb{E}[a_k^2] = \sigma_a^2 + a^2 < 1$ and $\mathbb{E}[b_k^2] = \sigma_b^2 + b^2 < 1$,

$$\mu = \frac{\sigma_a^2 + \sigma_b^2 + a^2 + b^2}{2} < 1 \qquad \text{and} \qquad \nu = ac + bd.$$

As previously, the strong law of large numbers leads to

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} E_n = \sigma_c^2 + \sigma_d^2 + c^2 + d^2 \qquad \text{a.s.} \tag{7.4}$$

Moreover, it follows once again from the Rademacher-Menchov theorem with Kronecker's lemma, (7.3), (7.4) and Lemma A.3 of [6] that

$$\lim_{n\to\infty} \frac{K_n}{2^{n+1}} = \frac{1}{1-\mu}\left(\nu \frac{c+d}{2-(a+b)} + \frac{\sigma_c^2 + \sigma_d^2 + c^2 + d^2}{2}\right) \qquad \text{a.s.}$$

leading to convergence (5.1) for $p = 2$. We shall not carry out the proof of (5.1) for $3 \le p \le 8$ inasmuch as it follows essentially the same lines that those for $p = 2$. One can observe that, in order to prove (5.1) for $3 \le p \le 8$, it is necessary to assume that $\mathbb{E}[a_1^{2p}] < 1$, $\mathbb{E}[b_1^{2p}] < 1$, $\mathbb{E}[\varepsilon_2^{2p}] < \infty$ and $\mathbb{E}[\varepsilon_3^{2p}] < \infty$. The limiting values s_3 to s_8 may be explicitly calculated. More precisely, for all $p \in \{1, 2, \ldots, 8\}$, denote

$$A_p = \mathbb{E}[a_1^p], \qquad B_p = \mathbb{E}[b_1^p], \qquad C_p = \mathbb{E}[\varepsilon_2^p], \qquad D_p = \mathbb{E}[\varepsilon_3^p].$$

We already saw that

$$s_1 = \frac{C_1 + D_1}{2 - (A_1 + B_1)} \quad \text{and} \quad s_2 = \frac{2}{2 - (A_2 + B_2)}\left((A_1C_1 + B_1D_1)s_1 + \frac{C_2 + D_2}{2}\right).$$

The other limiting values s_3 to s_8 of convergence (5.1) can be recursively calculated via the linear relation

$$s_p = \frac{2}{2 - (A_p + B_p)}\left(\sum_{k=1}^{p-1} \frac{1}{2}\binom{p}{k}(A_kC_{p-k} + B_kD_{p-k})s_k + \frac{C_p + D_p}{2}\right).$$

8. Proof of Proposition 5.3

The almost sure convergence (5.2) is immediate through (4.2), (4.3) and Lemma 5.1. Let us now prove that L is a positive definite matrix. First, the matrices

$$\begin{pmatrix} \sigma_a^2 & \rho_{ab} \\ \rho_{ab} & \sigma_b^2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix}$$

are clearly positive semidefinite and positive definite under **(H.3)**. Moreover, D is clearly positive semidefinite since

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} \begin{pmatrix} X_k^4 & X_k^3 \\ X_k^3 & X_k^2 \end{pmatrix} = D \qquad \text{a.s.}$$

Finally, let us prove that C is positive definite. Its trace is clearly greater than 1, hence we only have to prove that its determinant is positive. Its determinant is given by

$$s_2 - s_1^2 = \frac{\sigma_c^2 + \sigma_d^2}{2 - (\sigma_a^2 + \sigma_b^2 + a^2 + b^2)} + \left(\frac{c+d}{2-(a+b)}\right)^2 \frac{\sigma_a^2 + \sigma_b^2}{2 - (\sigma_a^2 + \sigma_b^2 + a^2 + b^2)} + \frac{2}{2 - (\sigma_a^2 + \sigma_b^2 + a^2 + b^2)} \frac{(ad - bc + c - d)^2}{(2-(a+b))^2}.$$

The first term of this sum is positive since under **(H.1)** $\sigma_a^2 + \sigma_b^2 + a^2 + b^2 < 2$ and since under **(H.2)** $\sigma_c^2 + \sigma_d^2 > 0$. Moreover, the two other terms are clearly nonnegative, which proves that this matrix is positive definite. Since the Kronecker product of two positive semidefinite (respectively positive definite) matrices is a positive semidefinite (respectively positive definite) matrix, we can conclude that L is positive definite.

9. Proofs of the almost sure convergence results

We shall make use of a martingale approach, as the one developed by Bercu et al. [6] or de Saporta et al. [13]. For all $n \geq 1$, let

$$\mathcal{V}_n = M_n^t P_{n-1}^{-1} M_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1} (\widehat{\theta}_n - \theta)$$

where

$$P_n = \sum_{k \in \mathbb{T}_n} (1 + X_k^2) I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

By the same calculations as in [6], we can easily see that if $\Delta M_n = M_n - M_{n-1}$,

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1}, \tag{9.1}$$

where

$$\mathcal{A}_n = \sum_{k=1}^{n} M_k^t (P_{k-1}^{-1} - P_k^{-1}) M_k,$$

$$\mathcal{B}_{n+1} = 2 \sum_{k=1}^{n} M_k^t P_k^{-1} \Delta M_{k+1} \qquad \text{and} \qquad \mathcal{W}_{n+1} = \sum_{k=1}^{n} \Delta M_{k+1}^t P_k^{-1} \Delta M_{k+1}.$$

Lemma 9.1. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n \to \infty} \frac{\mathcal{W}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad a.s. \tag{9.2}$$

where C and D are the matrices given by (5.3). *In addition, we also have*

$$\mathcal{B}_{n+1} = o(n) \qquad a.s. \tag{9.3}$$

and

$$\lim_{n \to \infty} \frac{\mathcal{V}_{n+1} + \mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad a.s. \tag{9.4}$$

Proof. The proof of convergence (9.2) is left to the reader inasmuch as it follows essentially the same arguments as the proof of Lemma B.3 in [6]. In particular, we also find that

$$\lim_{n\to\infty} \frac{P_n}{|\mathbb{T}_n|} = I_2 \otimes (C + D) \qquad \text{a.s.} \tag{9.5}$$

We shall proceed to the proof of (9.3). We clearly have

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^{n} M_k^t P_k^{-1} \Delta M_{k+1} = 2\sum_{k=1}^{n} M_k^t P_k^{-1} \Psi_k \xi_{k+1}.$$

Hence, $(\mathcal{B}_n)$ is a square integrable martingale. In addition, we have

$$\Delta\mathcal{B}_{n+1} = 2M_n^t P_n^{-1} \Delta M_{n+1}.$$

Consequently,

$$\begin{aligned}
\mathbb{E}[(\Delta\mathcal{B}_{n+1})^2|\mathcal{F}_n] &= 4\mathbb{E}[M_n^t P_n^{-1} \Delta M_{n+1} \Delta M_{n+1}^t P_n^{-1} M_n|\mathcal{F}_n] \qquad \text{a.s.} \\
&= 4M_n^t P_n^{-1} \mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t|\mathcal{F}_n] P_n^{-1} M_n \qquad \text{a.s.} \\
&= 4M_n^t P_n^{-1} L_n P_n^{-1} M_n \qquad \text{a.s.}
\end{aligned}$$

However, we already saw from (4.3) that

$$L_n = \sum_{k\in\mathbb{G}_n} \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

Moreover,

$$\Delta P_n = P_n - P_{n-1} = \sum_{k\in\mathbb{G}_n} (1+X_k)^2 I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

For $\alpha = \max(\sigma_a^2, \sigma_c^2) + \max(\sigma_b^2, \sigma_d^2) + \max(|\rho_{ab}|, |\rho_{cd}|)$, denote

$$\Delta_n = \begin{pmatrix} \alpha(1+X_n^2) - P(X_n) & -Q(X_n) \\ -Q(X_n) & \alpha(1+X_n^2) - R(X_n) \end{pmatrix}$$

where $P(X_n)$, $Q(X_n)$ and $R(X_n)$ are given by (4.4). It is not hard to see that

$$\alpha\Delta P_n - L_n = \sum_{k\in\mathbb{G}_n} \Delta_k \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}$$

and that Δ_n is positive definite which immediately implies that $L_n \leq \alpha\Delta P_n$. Moreover, we can use Lemma B.1 of [6] to say that

$$P_{n-1}^{-1} \Delta P_n P_{n-1}^{-1} \leq P_{n-1}^{-1} - P_n^{-1}.$$

Hence

$$\begin{aligned}
\mathbb{E}[(\Delta\mathcal{B}_{n+1})^2|\mathcal{F}_n] &= 4M_n^t P_n^{-1} L_n P_n^{-1} M_n \qquad \text{a.s.} \\
&\leq 4\alpha M_n^t P_n^{-1} \Delta P_n P_n^{-1} M_n \qquad \text{a.s.} \\
&\leq 4\alpha M_n^t (P_{n-1}^{-1} - P_n^{-1}) M_n \qquad \text{a.s.}
\end{aligned}$$

leading to $\langle \mathcal{B} \rangle_n \leq 4\alpha \mathcal{A}_n$. Therefore it follows from the strong law of large numbers for martingales that $\mathcal{B}_n = o(\mathcal{A}_n)$. Hence, we deduce from decomposition (9.1) that

$$\mathcal{V}_{n+1} + \mathcal{A}_n = o(\mathcal{A}_n) + \mathcal{O}(n) \qquad \text{a.s.}$$

leading to $\mathcal{V}_{n+1} = \mathcal{O}(n)$ and $\mathcal{A}_n = \mathcal{O}(n)$ a.s. which implies that $\mathcal{B}_n = o(n)$ a.s. Finally we clearly obtain convergence (9.4) from the main decomposition (9.1) together with (9.2) and (9.3), which completes the proof of Lemma 9.1. $\square$

Lemma 9.2. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. For all* $\delta > 1/2$, *we have*

$$\|M_n\|^2 = o(|\mathbb{T}_n| n^{\delta}) \qquad a.s.$$

Proof. Let us recall that

$$M_n = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}.$$

Denote

$$T_n = \sum_{k \in \mathbb{T}_{n-1}} X_k V_{2k} \qquad \text{and} \qquad U_n = \sum_{k \in \mathbb{T}_{n-1}} V_{2k}.$$

On the one hand, T_n can be rewritten as

$$T_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|} f_k \qquad \text{where} \qquad f_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k \in \mathbb{G}_{n-1}} X_k V_{2k}.$$

We already saw in Section 3 that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\mathbb{E}[V_{2k} | \mathcal{F}_n] = 0 \qquad \text{and} \qquad \mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2 = P(X_k) \qquad \text{a.s.}$$

In addition, for all $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k}^4 | \mathcal{F}_n] = \mu_a^4 X_k^4 + 6\sigma_a^2 \sigma_c^2 X_k^2 + \mu_c^4$ a.s. which implies that

$$\mathbb{E}[V_{2k}^4 | \mathcal{F}_n] \leq \mu_{ac}^4 (1 + X_k^2)^2 \qquad \text{a.s.} \tag{9.6}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$. Consequently, $\mathbb{E}[f_{n+1} | \mathcal{F}_n] = 0$ a.s. In addition,

$$\mathbb{E}[f_{n+1}^4 | \mathcal{F}_n] = \frac{1}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} X_k^4 \mathbb{E}[V_{2k}^4 | \mathcal{F}_n] + \frac{3}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} \sum_{\substack{l \in \mathbb{G}_n \\ l \neq k}} X_k^2 X_l^2 \mathbb{E}[V_{2k}^2 | \mathcal{F}_n] \mathbb{E}[V_{2l}^2 | \mathcal{F}_n],$$

which implies from (9.6) together with the Cauchy-Schwarz inequality that

$$\mathbb{E}[f_{n+1}^4 | \mathcal{F}_n] \leq \frac{\mu_{ac}^4}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} X_k^4 (1 + X_k^2)^2 + 3 \max(\sigma_a^2, \sigma_c^2)^2 \left(\frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} X_k^2 (1 + X_k^2) \right)^2.$$

Therefore, we infer from Lemma 5.1 that

$$\sup_{n \geq 0} \mathbb{E}[f_{n+1}^4 | \mathcal{F}_n] < \infty \qquad \text{a.s.}$$

Hence, we obtain from Wei's lemma given in [36] page 1672, together with Lemma A.2 of [6], that for all $\delta > 1/2$,

$$T_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

On the other hand, U_n can be rewritten as

$$U_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|}g_k \qquad \text{where} \qquad g_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k \in \mathbb{G}_{n-1}} V_{2k}.$$

Via the same calculation as before, $\mathbb{E}[g_{n+1}|\mathcal{F}_n] = 0$ a.s. and

$$\mathbb{E}[g_{n+1}^4|\mathcal{F}_n] \leq \frac{\mu_{bd}^4}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} (1+X_k^2)^2 + 3\max(\sigma_b^2, \sigma_d^2)^2 \left(\frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} (1+X_k^2) \right)^2.$$

where $\mu_{bd}^4 = \max(\mu_b^4, 3\sigma_b^2\sigma_d^2, \mu_d^4)$. Hence, we deduce once again from Lemma 5.1 and Wei's Lemma, together with Lemma A.2 of [6], that for all $\delta > 1/2$,

$$U_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

In the same way, we obtain the same result for the two last components of M_n, which completes the proof of Lemma 9.2. $\square$

9.1. **Proof of Theorem 5.5.** We recall that $\mathcal{V}_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1} (\widehat{\theta}_n - \theta)$ which implies that

$$\|\widehat{\theta}_n - \theta\|^2 \leq \frac{\mathcal{V}_n}{\lambda_{min}(\Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1})}.$$

where $\lambda_{min}(A)$ stands for the smallest eigenvalue of A. On the one hand, it follows from (9.4) that $\mathcal{V}_n = \mathcal{O}(n)$ a.s. On the other hand, we deduce from Lemma 5.1 that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C = \Gamma \qquad \text{a.s.} \tag{9.7}$$

where C is the positive definite matrix given by (5.3). Therefore, we obtain from (9.5) and (9.7) that

$$\lim_{n\to\infty} \frac{\lambda_{min}(\Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1})}{|\mathbb{T}_{n-1}|} = \lambda_{min}(C(C+D)^{-1}C) > 0 \qquad \text{a.s.}$$

Consequently, we find that

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|} \right) \qquad \text{a.s.}$$

We are now in position to prove the quadratic strong law (5.4). First of all, a direct application of Lemma 9.2 ensures that $\mathcal{V}_n = o(n^\delta)$ a.s. for all $\delta > 1/2$. Hence, we obtain from (9.4) that

$$\lim_{n\to\infty} \frac{\mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad \text{a.s.} \tag{9.8}$$

Let us rewrite $\mathcal{A}_n$ as

$$\mathcal{A}_n = \sum_{k=1}^{n} M_k^t \left(P_{k-1}^{-1} - P_k^{-1}\right) M_k = \sum_{k=1}^{n} M_k^t P_{k-1}^{-1/2} A_k P_{k-1}^{-1/2} M_k,$$

where $A_k = I_4 - P_{k-1}^{1/2} P_k^{-1} P_{k-1}^{1/2}$. We already saw from (9.5) that

$$\lim_{n\to\infty} \frac{P_n}{|\mathbb{T}_n|} = I_2 \otimes (C + D) \qquad \text{a.s.} \tag{9.9}$$

which ensures that

$$\lim_{n\to\infty} A_n = \frac{1}{2} I_4 \qquad \text{a.s.}$$

In addition, we deduce from (9.4) that $\mathcal{A}_n = \mathcal{O}(n)$ a.s. which implies that

$$\frac{\mathcal{A}_n}{n} = \left(\frac{1}{2n} \sum_{k=1}^{n} M_k^t P_{k-1}^{-1} M_k\right) + o(1) \qquad \text{a.s.} \tag{9.10}$$

Moreover, we also have from (9.7) and (9.9) that

$$\begin{aligned}
\frac{1}{n}\sum_{k=1}^{n} M_k^t P_{k-1}^{-1} M_k &= \frac{1}{n}\sum_{k=1}^{n} (\widehat{\theta}_k - \theta)^t \Sigma_{k-1} P_{k-1}^{-1} \Sigma_{k-1} (\widehat{\theta}_k - \theta) \\
&= \frac{1}{n}\sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} |\mathbb{T}_{k-1}| P_{k-1}^{-1} \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} (\widehat{\theta}_k - \theta) \\
&= \frac{1}{n}\sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \Gamma (I_2 \otimes (C+D)^{-1}) \Gamma (\widehat{\theta}_k - \theta) + o(1) \qquad \text{a.s.} \qquad (9.11)
\end{aligned}$$

Therefore, (9.8) together with (9.10) and (9.11) lead to (5.4). □

9.2. **Proof of Theorem 5.6.** We only prove (5.5) inasmuch as the proof of (5.6) follows exactly the same lines. Relation (3.8) immediately leads to

$$\begin{aligned}
Q_{n-1}(\widehat{\eta}_n - \eta_n) &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k}^2 - V_{2k}^2) \psi_k \\
&= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \left((\widehat{V}_{2k} - V_{2k})^2 + 2(\widehat{V}_{2k} - V_{2k}) V_{2k} \right) \psi_k. \qquad (9.12)
\end{aligned}$$

Moreover, we clearly have from Section 3 that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\widehat{V}_{2k} - V_{2k} = - \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}^t \Phi_k,$$

which implies that

$$(\widehat{V}_{2k} - V_{2k})^2 \leq \left((\widehat{a}_n - a)^2 + (\widehat{c}_n - c)^2\right) \|\Phi_k\|^2 = \left((\widehat{a}_n - a)^2 + (\widehat{c}_n - c)^2\right) (1 + X_k^2).$$

In addition, since $\|\psi_k\|^2 = 1 + X_k^4 \leq (1 + X_k^2)^2$, we have

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k})^2 \psi_k \right\| \leq \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \right) \sum_{k \in \mathbb{G}_l} (1 + X_k^2)^2.$$

However, it follows from Lemma 5.1 that

$$\sum_{k \in \mathbb{G}_l} (1 + X_k^2)^2 = \mathcal{O}(|\mathbb{G}_l|) \qquad \text{a.s.}$$

and since Λ is positive definite, (5.4) leads to

$$\sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \right) |\mathbb{G}_l| = \mathcal{O}(n) \qquad \text{a.s.}$$

Hence, we find that

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k})^2 \psi_k \right\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{9.13}$$

Let us now tackle

$$P_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k}) V_{2k} \psi_k.$$

It is clear that

$$\Delta P_{n+1} = P_{n+1} - P_n = \sum_{k \in \mathbb{G}_n} (\widehat{V}_{2k} - V_{2k}) V_{2k} \psi_k = - \sum_{k \in \mathbb{G}_n} V_{2k} \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}.$$

Since, for al $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k} | \mathcal{F}_n] = 0$ a.s. and $\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = P(X_k)$ a.s., we have

$$\mathbb{E}[\Delta P_{n+1} \Delta P_{n+1}^t | \mathcal{F}_n] = \sum_{k \in \mathbb{G}_n} P(X_k) \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix} \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}^t \Phi_k \psi_k^t \qquad \text{a.s.}$$

which allows to say that (P_n) is a square integrable martingale with increasing process $\langle P \rangle_n$ given by

$$\begin{aligned} \langle P \rangle_n &= \sum_{l=0}^{n-1} \mathbb{E}[\Delta P_{l+1} \Delta P_{l+1}^t | \mathcal{F}_n] \\ &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} P(X_k) \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \psi_k^t \qquad \text{a.s.} \end{aligned}$$

Consequently, if $\alpha = \max(\sigma_a^2, \sigma_c^2)$, we obtain that

$$\begin{aligned} \|\langle P \rangle_n\| &\leq \alpha \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \right) \sum_{k \in \mathbb{G}_l} (1 + X_k^2) \|\psi_k\|^2 \|\Phi_k\|^2 \qquad \text{a.s.} \\ &\leq \alpha \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \right) \sum_{k \in \mathbb{G}_l} (1 + X_k^2)^4 \qquad \text{a.s.} \end{aligned}$$

leading, as previously via Lemma 5.1 and (5.4), to $\|\langle P\rangle_n\| = \mathcal{O}(n)$ a.s. The strong law of large numbers for martingale given e.g. in Theorem 1.3.15 of [16] implies that

$$P_n = o(n) \qquad \text{a.s.} \tag{9.14}$$

Then, we deduce from (9.12), (9.13) and (9.14) that

$$\|Q_{n-1}(\widehat{\eta}_n - \eta_n)\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{9.15}$$

Moreover, we obtain through Lemma 5.1 that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} Q_n = \begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix} = A \qquad \text{a.s.} \tag{9.16}$$

and we can prove, through tedious calculations, that this limiting matrix is positive definite. Therefore, (9.15) immediately implies (5.5). We shall now proceed to the proof of (5.7). Denote

$$R_n = \sum_{k\in\mathbb{T}_{n-1}} (\widehat{W}_k - W_k)^t J W_k \psi_k,$$

where

$$\widehat{W}_k = \begin{pmatrix} \widehat{V}_{2k} \\ \widehat{V}_{2k+1} \end{pmatrix} \qquad \text{and} \qquad J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It follows from (3.9) that

$$Q_n(\widehat{\nu}_n - \nu_n) = \sum_{k\in\mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k + R_n.$$

Furthermore, one can observe that (R_n) is a square integrable martingale with increasing process

$$\begin{aligned}
\langle R\rangle_n &= \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} \mathbb{E}[(\widehat{W}_k - W_k)^t J W_k W_k^t J(\widehat{W}_k - W_k)\psi_k\psi_k^t|\mathcal{F}_l] \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t J\mathbb{E}[W_k W_k^t|\mathcal{F}_l] J(\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t J \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} J(\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t \begin{pmatrix} R(X_k) & Q(X_k) \\ Q(X_k) & P(X_k) \end{pmatrix} (\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}
\end{aligned}$$

Then, as previously, Lemma 5.1 and (5.4) lead to $\|\langle R\rangle_n\| = \mathcal{O}(n)$ a.s. which allows us to say that $R_n = o(n)$ a.s. Furthermore

$$\begin{aligned}\Big\| \sum_{k\in\mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k \Big\| & \\ &\leq \frac{1}{2}\sum_{k\in\mathbb{T}_{n-1}} \Big((\widehat{V}_{2k} - V_{2k})^2 + (\widehat{V}_{2k+1} - V_{2k+1})^2 \Big) \|\psi_k\| \\ &\leq \frac{1}{2}\sum_{l=0}^{n-1} \|\widehat{\theta}_l - \theta\|^2 \sum_{k\in\mathbb{G}_l} \|\Phi_k\|^2 \|\psi_k\|,\end{aligned}$$

which implies, thanks to Lemma 5.1 and (5.4), that

$$\Big\| \sum_{k\in\mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k \Big\| = \mathcal{O}(n) \qquad \text{a.s.}$$

Finally, we infer from (9.16) that

$$\|\widehat{\nu}_n - \nu_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

It remains to prove the a.s. convergence of η_n, ζ_n and ν_n to η, ζ and ν, respectively which would immediately imply the a.s. convergence of our estimates through (5.5), (5.6) and (5.7). Denote

$$N_n = Q_{n-1}(\eta_n - \eta) = \sum_{k\in\mathbb{T}_{n-1}} v_{2k}\psi_k \tag{9.17}$$

where $v_{2n} = V_{2n}^2 - \eta^t\psi_n$. One can observe that (N_n) is a square integrable martingale with increasing process $\langle N\rangle_n$ given by

$$\langle N\rangle_n = \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \mathbb{E}[v_{2k}^2|\mathcal{F}_l]\psi_k\psi_k^t \qquad \text{a.s.}$$

Hence, if $\gamma = \max(\mu_a^4 - \sigma_a^4, 2\sigma_a^2\sigma_c^2, \mu_c^4 - \sigma_c^4)$, we obtain that

$$\begin{aligned}\|\langle N\rangle_n\| &\leq \Big\| \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \gamma(1+X_k^2)^2\psi_k\psi_k^t \Big\| \qquad \text{a.s.} \\ &\leq \gamma \sum_{k\in\mathbb{T}_{n-1}} (1+X_k^2)^2\|\psi_k\|^2 = \gamma \sum_{k\in\mathbb{T}_{n-1}} (1+X_k^2)^4 \qquad \text{a.s.}\end{aligned}$$

which leads, via Lemma 5.1, to $\|\langle N\rangle_n\| = \mathcal{O}(|\mathbb{T}_{n-1}|)$ a.s. Consequently,

$$\|N_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{a.s.}$$

Then, we deduce from (9.16) and (9.17) that η_n converges a.s. to η with the a.s. rate of convergence given in Remark 5.7. The proof concerning the a.s. convergence of ζ_n to ζ and the second rate of convergence in Remark 5.7 is exactly the same. Hereafter, denote

$$H_n = Q_{n-1}(\nu_n - \nu) = \sum_{k \in \mathbb{T}_{n-1}} w_{2k} \psi_k \tag{9.18}$$

where $w_{2n} = V_{2n}V_{2n+1} - \nu^t \psi_n$. Once again, the sequence (H_n) is a square integrable martingale with increasing process

$$\langle H \rangle_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \mathbb{E}[w_{2k}^2 | \mathcal{F}_l] \psi_k \psi_k^t \qquad \text{a.s.}$$

Moreover, if $\alpha = \max(\nu_{ab}^2, \nu_{cd}^2, (\sigma_a^2 + \sigma_c^2)(\sigma_b^2 + \sigma_d^2))$, we find that

$$\begin{aligned} \|\langle H \rangle_n\| &\leq \left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \alpha (1 + X_k^2)^2 \psi_k \psi_k^t \right\| \qquad \text{a.s.} \\ &\leq \alpha \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^2 \|\psi_k\|^2 = \alpha \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^4 \qquad \text{a.s.} \end{aligned}$$

which allows us to say, as previously, that

$$\|H_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{and} \qquad \|\nu_n - \nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

It clearly proves the a.s. convergence of ν_n to ν with the last a.s. rate of convergence given in Remark 5.7, which completes the proof of Theorem 5.6. □

10. Proofs of the asymptotic normalities

The key point of the proof of the asymptotic normality of our estimators is the central limit theorem for triangular array of vector martingale given e.g. in Theorem 2.1.9 of [16]. With this aim in mind, we will change the filtration considering, instead of the generation wise filtration $(\mathcal{F}_n)$, the sister-pair wise filtration $(\mathcal{G}_n)$ given by

$$\mathcal{G}_n = \sigma \left\{ X_1, (X_{2k}, X_{2k+1}), 1 \leq k \leq n \right\}.$$

10.1. **Proof of convergence** (5.8).

We will consider the triangular array of vector martingale $(M_k^{(n)})$ defined as

$$M_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{k} D_l \qquad \text{where} \qquad D_l = \begin{pmatrix} X_l V_{2l} \\ V_{2l} \\ X_l V_{2l+1} \\ V_{2l+1} \end{pmatrix}. \tag{10.1}$$

It is obvious that $\left(M^{(n)}\right)$ is a square integrable vector valued martingale with respect to the filtration $(\mathcal{G}_k)$. Moreover, we can observe that

$$M_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{t_n} D_l = \frac{1}{\sqrt{|\mathbb{T}_n|}} M_{n+1} \tag{10.2}$$

where $t_n = |\mathbb{T}_n| = 2^{n+1} - 1$. In addition, the increasing process of this square integrable martingale is given by

$$\begin{aligned}\langle M^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \mathbb{E}[D_l D_l^t | \mathcal{G}_{l-1}] \\ &= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \begin{pmatrix} P(X_l) & Q(X_l) \\ Q(X_l) & R(X_l) \end{pmatrix} \otimes \begin{pmatrix} X_l^2 & X_l \\ X_l & 1 \end{pmatrix} \qquad \text{a.s.}\end{aligned}$$

Then, (5.2) leads to

$$\lim_{n \to \infty} \langle M^{(n)} \rangle_{t_n} = L \qquad \text{a.s.}$$

We will now establish Lindeberg's condition thanks to Lyapunov's condition. Let

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^4 \Big| \mathcal{G}_{k-1} \right].$$

It follows from (10.1) that

$$\begin{aligned}\phi_n &= \frac{1}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[(1 + X_k^2)^2 (V_{2k}^2 + V_{2k+1}^2)^2 \big| \mathcal{G}_{k-1} \right] \\ &\leq \frac{2}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[(1 + X_k^2)^2 (V_{2k}^4 + V_{2k+1}^4) \big| \mathcal{G}_{k-1} \right].\end{aligned}$$

Since we already saw in Section 9 that

$$\mathbb{E}\left[V_{2k}^4 | \mathcal{F}_n \right] \leq \mu_{ac}^4 (1 + X_k^2)^2 \qquad \text{and} \qquad \mathbb{E}[V_{2k+1}^4 | \mathcal{F}_n] \leq \mu_{bc}^4 (1 + X_k^2)^2 \qquad \text{a.s.}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$ and $\mu_{bd}^4 = \max(\mu_b^4, 3\sigma_b^2\sigma_d^2, \mu_d^4)$, we have that

$$\phi_n \leq \frac{2(\mu_{ac}^4 + \mu_{bd}^4)}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} (1 + X_k^2)^4 \qquad \text{a.s.}$$

leading, via Lemma 5.1, to the a.s. convergence of ϕ_n to 0. Consequently, Lyapunov's condition is satisfied and Theorem 2.1.9 of [16] together with (10.2) imply that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, L).$$

Moreover, we easily obtain from Lemma 5.1 that

$$\lim_{n \to \infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C = \Gamma \qquad \text{a.s.} \tag{10.3}$$

where C is the positive definite matrix given by (5.3). Finally, we deduce from (4.1) together with (10.3) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|} (\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^{-1} L \Gamma^{-1}).$$

10.2. **Proof of convergences** (5.9) **and** (5.11)**.** The proof of convergences (5.9) and (5.11) are left to the reader as it follows essentially the same lines as the one of convergence (5.8).

Acknowledgement. The authors are deeply grateful to the Associate Editor and the two anonymous Reviewers for their suggestions and constructive comments which helped to improve the paper substantially.

References

[1] Aue, A., and Horváth, L. Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Statist. Sinica 21*, 3 (2011), 973–999.

[2] Aue, A., Horváth, L., and Steinebach, J. Estimation in random coefficient autoregressive models. *J. Time Ser. Anal. 27*, 1 (2006), 61–76.

[3] Bansaye, V. Cell contamination and branching processes in a random environment with immigration. *Adv. in Appl. Probab. 41*, 4 (2009), 1059–1081.

[4] Basawa, I. V., and Zhou, J. Non-Gaussian bifurcating models and quasi-likelihood estimation. *J. Appl. Probab. 41* (2004), 55–64. Stochastic methods and their applications.

[5] Bercu, B., and Blandin, V. Limit theorems for bifurcating integer-valued autoregressive processes. *To appear in Stat. Inference Stoch. Process. 17* (2014).

[6] Bercu, B., de Saporta, B., and Gégout-Petit, A. Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron. J. Probab. 14* (2009), no. 87, 2492–2526.

[7] Berkes, I., Horváth, L., and Ling, S. Estimation in nonstationary random coefficient autoregressive models. *J. Time Series Anal. 30*, 4 (2009), 395–416.

[8] Bitseki Penda, S. V., and Djellout, H. Deviation inequalities and moderate deviations for estimators of parameters in bifurcating autoregressive models. *Ann. Inst. Henri Poincaré Probab. Stat. 50*, 3 (2014), 806–844.

[9] Blandin, V. Asymptotic results for bifurcating random coefficient autoregressive processes. *To appear in Statistics 48* (2014).

[10] Cowan, R., and Staudte, R. G. The bifurcating autoregressive model in cell lineage studies. *Biometrics 42* (1986), 769–783.

[11] De Saporta, B., Gégout-Petit, A., and Marsalle, L. Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electron. J. Stat. 5* (2011), 1313–1353.

[12] De Saporta, B., Gégout-Petit, A., and Marsalle, L. Asymmetry tests for bifurcating auto-regressive processes with missing data. *Statist. Probab. Lett. 82*, 7 (2012), 1439–1444.

[13] de Saporta, B., Gégout-Petit, A., and Marsalle, L. Random coefficients bifurcating autoregressive processes. *To appear in ESAIM PS 18* (2014).

[14] De Saporta, B., Gégout-Petit, A., and Marsalle, L. Statistical study of asymmetry in cell lineage data. *Comput. Statist. Data Anal. 69* (2014), 15–39.

[15] Delmas, J.-F., and Marsalle, L. Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl. 120*, 12 (2010), 2495–2519.

[16] Duflo, M. *Random iterative models*, vol. 34 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. Translated from the 1990 French original by Stephen S. Wilson and revised by the author.

[17] Guyon, J. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab. 17*, 5-6 (2007), 1538–1569.

[18] Guyon, J., Bize, A., Paul, G., Stewart, E., Delmas, J.-F., and Taddéi, F. Statistical study of cellular aging. In *CEMRACS 2004—mathematics and applications to biology and medicine*, vol. 14 of *ESAIM Proc.* EDP Sci., Les Ulis, 2005, pp. 100–114 (electronic).

[19] Hall, P., and Heyde, C. C. *Martingale limit theory and its application.* Academic Press Inc., New York, 1980. Probability and Mathematical Statistics.

[20] Heyde, C. C., and Seneta, E. Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab. 9* (1972), 235–256.

[21] Huggins, R. M., and Basawa, I. V. Extensions of the bifurcating autoregressive model for cell lineage studies. *J. Appl. Probab. 36*, 4 (1999), 1225–1233.

[22] Huggins, R. M., and Basawa, I. V. Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust. N. Z. J. Stat. 42*, 4 (2000), 423–432.

[23] Huggins, R. M., and Staudte, R. G. Variance components models for dependent cell populations. *J. Amer. Statist. Assoc. 89*, 425 (1994), 19–29.

[24] Koul, H. L., and Schick, A. Adaptive estimation in a random coefficient autoregressive model. *Ann. Statist. 24*, 3 (1996), 1025–1052.

[25] Menchoff, D. E. Sur les séries de fonctions orthogonales. *Fundamenta Math. 4* (1923), 82–105.

[26] Nicholls, D. F., and Quinn, B. G. *Random coefficient autoregressive models: an introduction*, vol. 11 of *Lecture Notes in Statistics.* Springer-Verlag, New York, 1982. Lecture Notes in Physics, 151.

[27] Quinn, B. G., and Nicholls, D. F. The estimation of random coefficient autoregressive models. II. *J. Time Ser. Anal. 2*, 3 (1981), 185–203.

[28] Rademacher, H. Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. *Math. Ann. 87*, 1-2 (1922), 112–138.

[29] Schick, A. $\sqrt{n}$-consistent estimation in a random coefficient autoregressive model. *Austral. J. Statist. 38*, 2 (1996), 155–160.

[30] Stewart, E. J., Madden, R., Paul, G., and Taddéi, F. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol 3(2)*, e45 (2005).

[31] Stout, W. F. *Almost sure convergence.* Academic Press, New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

[32] Talagrand, M. Convergence of orthogonal series using stochastic processes. *Unpublished note* (1996).

[33] Tandori, K. Über die Divergenz der Orthogonalreihen. *Publ. Math. Debrecen 8* (1961), 291–307.

[34] Tandori, K. Bemerkung zur Konvergenz der Orthogonalreihen. *Acta Sci. Math. (Szeged) 26* (1965), 249–251.

[35] Vaněček, P. Rate of convergence for a class of RCA estimators. *Kybernetika (Prague) 42*, 6 (2006), 699–709.

[36] Wei, C. Z. Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. *Ann. Statist. 15*, 4 (1987), 1667–1682.

[37] Zhou, J., and Basawa, I. V. Least-squares estimation for bifurcating autoregressive processes. *Statist. Probab. Lett. 74*, 1 (2005), 77–88.

[38] Zhou, J., and Basawa, I. V. Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J. Time Ser. Anal. 26*, 6 (2005), 825–842.

Université de Bordeaux, Institut de Mathématiques de Bordeaux, UMR CNRS 5251, 351 cours de la libération, 33405 Talence cedex, France.

E-mail address: bernard.bercu@math.u-bordeaux1.fr

E-mail address: vassili.blandin@math.u-bordeaux1.fr