

# *Limit theorems for bifurcating integer-valued autoregressive processes*

**Bernard Bercu & Vassili Blandin**

## **Statistical Inference for Stochastic Processes**

An International Journal devoted to Time Series Analysis and the Statistics of Continuous Time Processes and Dynamical Systems

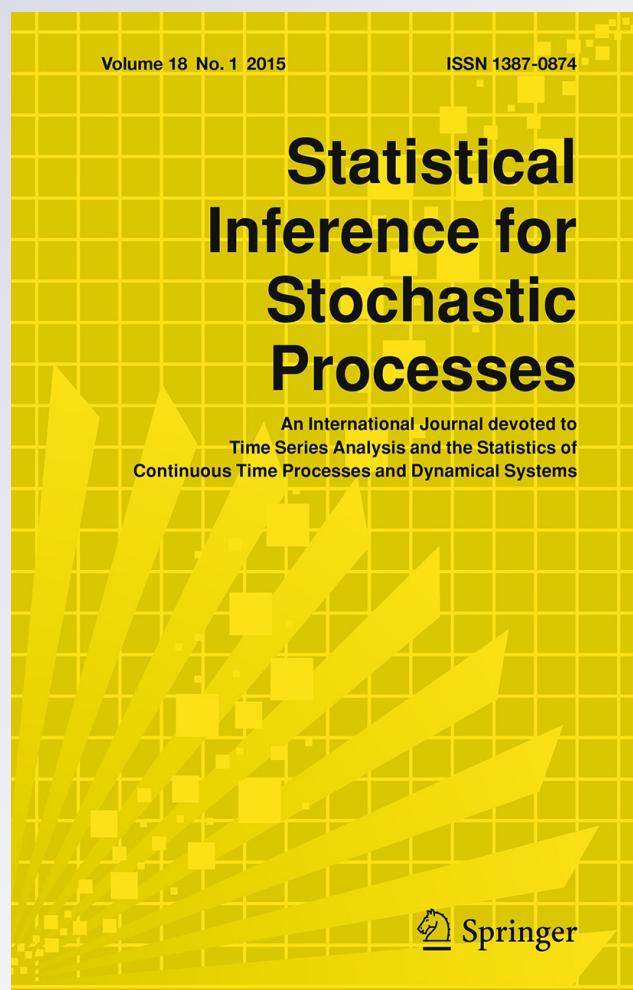
ISSN 1387-0874

Volume 18

Number 1

Stat Inference Stoch Process (2015)  
18:33-67

DOI 10.1007/s11203-014-9105-6



**Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Limit theorems for bifurcating integer-valued autoregressive processes

Bernard Bercu · Vassili Blandin

Received: 17 January 2013 / Accepted: 28 June 2014 / Published online: 29 August 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** We study the asymptotic behavior of the weighted least squares estimators of the unknown parameters of bifurcating integer-valued autoregressive processes. Under suitable assumptions on the immigration, we establish the almost sure convergence of our estimators, together with a quadratic strong law and central limit theorems. All our investigation relies on asymptotic results for vector-valued martingales.

**Keywords** Bifurcating autoregressive process · Integer-valued process · Weighted least squares · Martingale · Almost sure convergence · Central limit theorem

**Mathematics Subject Classification** Primary 60F15 · Secondary 60F05, 60G42

## 1 Introduction

Bifurcating integer-valued autoregressive (BINAR) processes are an adaptation of integer-valued autoregressive (INAR) processes to binary tree structured data. It can also be seen as the combination of INAR processes and bifurcating autoregressive (BAR) processes. BAR processes have been first introduced by [Cowan and Staudte \(1986\)](#) while INAR processes have been first investigated by [Alzaid and Al-Osh \(1987\)](#), [\(1990\)](#) and [McKenzie \(1985\)](#). BINAR processes take into account both inherited and environmental effects to explain the evolution of the integer-valued characteristic under study. To the best of our knowledge, this is the first paper devoted to BINAR processes.

We can easily see cell division as an example of binary tree structure, the integer-valued characteristic could then be, as an example, the number of parasites in a cell. Keeping this

---

B. Bercu (✉) · V. Blandin

Institut de Mathématiques de Bordeaux, Université de Bordeaux, UMR CNRS 5251, 351 cours de la libération, 33405 Talence Cedex, France  
e-mail: [bernard.bercu@math.u-bordeaux1.fr](mailto:bernard.bercu@math.u-bordeaux1.fr)

V. Blandin  
e-mail: [vassili.blandin@math.u-bordeaux1.fr](mailto:vassili.blandin@math.u-bordeaux1.fr)

example in mind, we consider that each time a cell is dividing, the two sister cells inherits both some parasites depending on the number of parasites of the mother, and some parasites from the environment. [Bansaye \(2008\)](#) used a Kimmel branching process to model this division process. This Kimmel process can be seen as the inheritance part of our BINAR process, where the parasites in the mother cell divide and then the offspring are distributed among the two sister cell. However, this model does not allow any environmental effect.

The first-order BINAR process is defined as follows. The initial cell is labelled 1 and the offspring of the cell labelled  $n$  are labelled  $2n$  and  $2n + 1$ . Denote by  $X_n$  the integer-valued characteristic of individual  $n$ . Then, the first-order BINAR process is given, for all  $n \geq 1$ , by

$$\begin{cases} X_{2n} &= a \circ X_n + \varepsilon_{2n} \\ X_{2n+1} &= b \circ X_n + \varepsilon_{2n+1} \end{cases} \quad (1.1)$$

where the thinning operator  $\circ$  will be defined in (2.1). The immigration sequence  $(\varepsilon_{2n}, \varepsilon_{2n+1})_{n \geq 1}$  represents the environmental effect, while the thinning operator represents the inherited effect. The example of the cell division incites us to suppose that  $\varepsilon_{2n}$  and  $\varepsilon_{2n+1}$  are correlated since the environmental effect on two sister cells can reasonably be seen as correlated.

The purpose of this paper is to study the asymptotic behavior of the weighted least squares (WLS) estimators of first-order BINAR process via a martingale approach. The martingale approach has been first proposed by [Bercu et al. \(2009\)](#). We also refer to [Wei and Winnicki \(1990\)](#) and [Winnicki \(1991\)](#) for the WLS estimation of parameters associated to branching processes. We shall make use of the strong law of large numbers [Duflo \(1997\)](#) as well as the central limit theorem ([Duflo 1997](#); [Hall and Heyde 1980](#)) for martingales, in order to investigate the asymptotic behavior of the WLS estimators, as previously done by [Basawa and Zhou \(2004\)](#), [Zhou and Basawa \(2005\)](#), [Zhou and Basawa \(2005\)](#). In contrast with [Bercu et al. \(2009\)](#), we investigate the asymptotic behavior of a WLS estimator instead of a least squares one. On the one hand, it enables us to reduce the moment assumption on the immigration sequence. On the other hand, it also allows us to reduce the asymptotic variance in the central limit theorem for our estimates. This gain of efficiency is entirely due to the weighted sequence which was inspired by [Wei and Winnicki \(1990\)](#). The fact that we consider an integer-valued process also forced us to adapt the proofs because of the thinning operator which needs to be manipulated more carefully than the classical product.

Several points of view appeared for both BAR and INAR processes and we tried to make a link between those approaches. On the one hand, for the BAR side of the BINAR process, we had a look to classical BAR studies as done by [Huggins and Basawa \(1999\)](#), [Huggins and Basawa \(2000\)](#) and [Huggins and Staudte \(1994\)](#) who studied the evolution of cell diameters and lifetimes, but also to bifurcating Markov chains models introduced by [Guyon \(2007\)](#) and used in [Delmas and Marsalle \(2010\)](#). However, we did not put aside the analogy with the Galton–Watson processes as studied in [Delmas and Marsalle \(2010\)](#) and [Heyde and Seneta \(1972\)](#). We also refer to the interesting contribution of [De Saporta et al. \(2011\)](#), [De Saporta et al. \(2012\)](#), [De Saporta et al. \(2014\)](#) dealing with statistical inference for BAR processes with missing data. On the other hand, concerning the INAR side of the BINAR process, we used the classical INAR definition but also had a look to [Bansaye \(2008\)](#), [Bansaye and Tran \(2011\)](#) who studied an integer-valued process on a binary tree without using an INAR model, and also [Kachour and Yao \(2009\)](#) who decided to study an integer-valued autoregressive process by a rounding approach instead of the classical INAR one. The approach of this paper has also been used for the study of random coefficient bifurcating autoregressive (RCBAR) process as in [Blandin \(2014\)](#) and [Bercu and Blandin \(2014\)](#). RCBAR processes is the combination of

BAR processes and random coefficient autoregressive processes. They have been previously investigated by Nicholls and Quinn (1980), Nicholls and Quinn (1982), Quinn and Nicholls (1981).

The paper is organised as follows. Section 2 is devoted to the presentation of the first-order BINAR process while Sect. 3 deals with the WLS estimators of the unknown parameters. Section 4 allows us to detail our approach based on martingales. Section 5 gathers the main results about the asymptotic properties of the WLS estimators. More precisely, we will propose the almost sure convergence, the quadratic strong law and the central limit theorem for our estimates. The rest of the paper is devoted to the proofs of our main results.

### 2 Bifurcating integer-valued autoregressive processes

Consider the first-order BINAR process given by (1.1) where the initial integer-valued state  $X_1$  is the ancestor of the process and  $(\varepsilon_{2n}, \varepsilon_{2n+1})$  represents the immigration which takes nonnegative integer values. In all the sequel, we shall assume that  $\mathbb{E}[X_1^8] < \infty$ . Moreover,

$$a \circ X_n = \sum_{i=1}^{X_n} Y_{n,i} \quad \text{and} \quad b \circ X_n = \sum_{i=1}^{X_n} Z_{n,i} \tag{2.1}$$

where  $(Y_{n,i})_{n,i \geq 1}$  and  $(Z_{n,i})_{n,i \geq 1}$  are two independent sequences of i.i.d., nonnegative integer-valued random variables with means  $a$  and  $b$  and positive variances  $\sigma_a^2$  and  $\sigma_b^2$  respectively. Moreover,  $\mu_a^4, \mu_b^4$  and  $\tau_a^6, \tau_b^6$  are the fourth-order and the sixth-order centered moments of  $(Y_{n,i})$  and  $(Z_{n,i})$ , respectively, and  $(Y_{n,i})$  and  $(Z_{n,i})$  admit eighth-order moments. We also assume that the two offspring sequences  $(Y_{n,i})$  and  $(Z_{n,i})$  are independent of the immigration  $(\varepsilon_{2n}, \varepsilon_{2n+1})$ . In addition, as in the literature concerning BAR processes, we shall assume that

$$0 < \max(a, b) < 1.$$

One can see this BINAR process as a first-order integer-valued autoregressive process on a binary tree, where each node represents an individual, node 1 being the original ancestor. For all  $n \geq 1$ , denote the  $n$ -th generation by

$$\mathbb{G}_n = \{2^n, 2^n + 1, \dots, 2^{n+1} - 1\}.$$

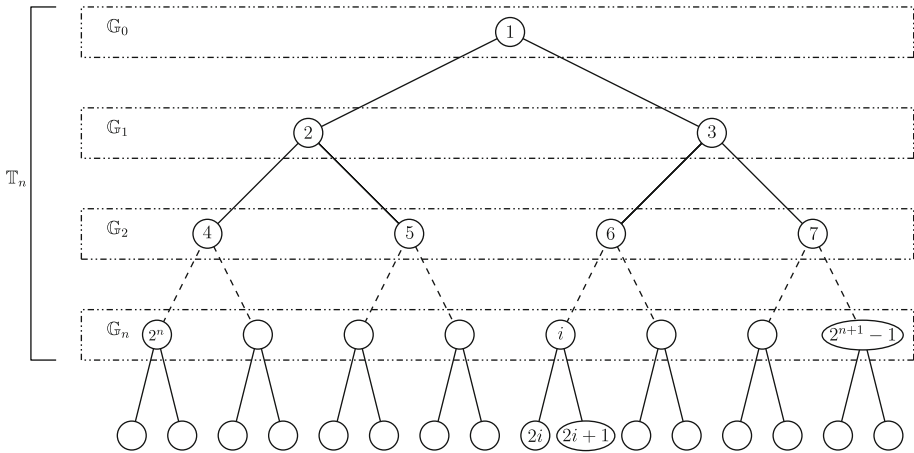
In particular,  $\mathbb{G}_0 = \{1\}$  is the initial generation and  $\mathbb{G}_1 = \{2, 3\}$  is the first generation of offspring from the first ancestor. Let  $\mathbb{G}_{r_n}$  be the generation of individual  $n$ , which means that  $r_n = \lceil \log_2(n) \rceil$  (Fig. 1). Recall that the two offspring of individual  $n$  are labelled  $2n$  and  $2n + 1$ , or conversely, the mother of individual  $n$  is  $\lfloor n/2 \rfloor$  where  $\lfloor x \rfloor$  stands for the largest integer less than or equal to  $x$ . Finally, denote by

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

the sub-tree of all individuals from the original individual up to the  $n$ -th generation. One can observe that the cardinality  $|\mathbb{G}_n|$  of  $\mathbb{G}_n$  is  $2^n$  while that of  $\mathbb{T}_n$  is  $|\mathbb{T}_n| = 2^{n+1} - 1$ .

### 3 Weighted least-squares estimation

Denote by  $\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}$  the natural filtration associated with the first-order BINAR process, which means that  $\mathcal{F}_n$  is the  $\sigma$ -algebra generated by all individuals up to the  $n$ -th generation,



**Fig. 1** The tree associated with the BINAR

in other words  $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$ . We will assume in all the sequel that, for all  $n \geq 0$  and for all  $k \in \mathbb{G}_n$ ,

$$\begin{cases} \mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c & \text{a.s.} \\ \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d & \text{a.s.} \end{cases}$$

Consequently, we deduce from (1.1) that, for all  $n \geq 0$  and for all  $k \in \mathbb{G}_n$ ,

$$\begin{cases} X_{2k} = aX_k + c + V_{2k}, \\ X_{2k+1} = bX_k + d + V_{2k+1}, \end{cases} \tag{3.1}$$

where  $V_{2k} = X_{2k} - \mathbb{E}[X_{2k}|\mathcal{F}_n]$  and  $V_{2k+1} = X_{2k+1} - \mathbb{E}[X_{2k+1}|\mathcal{F}_n]$ . Therefore, the two relations given by (3.1) can be rewritten in the matrix form

$$\chi_n = \theta^t \Phi_n + W_n \tag{3.2}$$

where

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

and the matrix parameter

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

Our goal is to estimate  $\theta$  from the observation of all individuals up to  $\mathbb{T}_n$ . We propose to make use of the WLS estimator  $\hat{\theta}_n$  of  $\theta$  which minimizes

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

where the choice of the weighting sequence  $(c_n)_{n \geq 1}$  is crucial. We shall choose  $c_n = 1 + X_n$  and we will go back to this suitable choice in Sect. 4. Consequently, we obviously have for all  $n \geq 1$

$$\hat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \Phi_k \chi_k^t \tag{3.3}$$

where

$$S_n = \sum_{k \in \mathbb{T}_n} \frac{1}{c_k} \Phi_k \Phi_k^t.$$

In order to avoid useless invertibility assumption, we shall assume, without loss of generality, that for all  $n \geq 0$ ,  $S_n$  is invertible. Otherwise, we only have to add the identity matrix of order 2,  $I_2$  to  $S_n$ . In all what follows, we shall make a slight abuse of notation by identifying  $\theta$  as well as  $\widehat{\theta}_n$  to

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

Therefore, we deduce from (3.3) that

$$\begin{aligned} \widehat{\theta}_n &= \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \text{vec}(\Phi_k \chi_k^t), \\ &= \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k X_{2k} \\ X_{2k} \\ X_k X_{2k+1} \\ X_{2k+1} \end{pmatrix} \end{aligned}$$

where  $\Sigma_n = I_2 \otimes S_n$  and  $\otimes$  stands for the standard Kronecker product. Consequently, (3.2) yields to

$$\begin{aligned} \widehat{\theta}_n - \theta &= \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \text{vec}(\Phi_k W_k^t), \\ &= \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \end{aligned} \tag{3.4}$$

In all the sequel, we shall make use of the following moment hypotheses.

**(H.1)** For all  $n \geq 0$  and for all  $k \in \mathbb{G}_n$

$$\mathbb{E}[\varepsilon_{2k} | \mathcal{F}_n] = c \text{ and } \mathbb{E}[\varepsilon_{2k+1} | \mathcal{F}_n] = d \text{ a.s.}$$

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

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

**(H.3)** For all  $n \geq 0$  and for all  $k, l \in \mathbb{G}_{n+1}$ , if  $[k/2] \neq [l/2]$ ,  $\varepsilon_k$  and  $\varepsilon_l$  are conditionally independent given  $\mathcal{F}_n$ , while otherwise it exists  $\rho^2 < \sigma_c^2 \sigma_d^2$  such that, for all  $k \in \mathbb{G}_n$

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

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

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

In addition, it exists  $v^4 \leq \mu_c^4 \mu_d^4$  such that, for all  $k \in \mathbb{G}_n$

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

(H.5) One can find  $\tau_c^6 > 0$  and  $\tau_d^6 > 0$  such that

$$\begin{aligned} \sup_{n \geq 1} \sup_{k \in \mathbb{G}_n} \mathbb{E}[\varepsilon_{2k}^6 | \mathcal{F}_n] = \tau_c^6 \quad \text{and} \quad \sup_{n \geq 1} \sup_{k \in \mathbb{G}_n} \mathbb{E}[\varepsilon_{2k+1}^6 | \mathcal{F}_n] = \tau_d^6 \quad \text{a.s.} \\ \sup_{n \geq 2} \mathbb{E}[\varepsilon_n^8] < \infty \end{aligned}$$

It follows from hypothesis (H.1) that  $V_{2n}$  and  $V_{2n+1}$  can be rewritten as

$$V_{2n} = \sum_{i=1}^{X_n} (Y_{n,i} - a) + (\varepsilon_{2n} - c) \quad \text{and} \quad V_{2n+1} = \sum_{i=1}^{X_n} (Z_{n,i} - b) + (\varepsilon_{2n} - d).$$

Hence, under assumption (H.2), we have for all  $n \geq 0$  and for all  $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = \sigma_a^2 X_k + \sigma_c^2 \quad \text{and} \quad \mathbb{E}[V_{2k+1}^2 | \mathcal{F}_n] = \sigma_b^2 X_k + \sigma_d^2 \quad \text{a.s.} \quad (3.5)$$

Consequently, if we choose  $c_n = 1 + X_n$  for all  $n \geq 1$ , we clearly have for all  $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] \leq \max(\sigma_a^2, \sigma_c^2) c_k \quad \text{and} \quad \mathbb{E}[V_{2k+1}^2 | \mathcal{F}_n] \leq \max(\sigma_b^2, \sigma_d^2) c_k \quad \text{a.s.}$$

It is exactly the reason why we have chosen this weighting sequence into (3.3). Similar WLS estimation approach for branching processes with immigration may be found in Wei and Winnicki (1990) and Winnicki (1991). We can also observe that, for all  $k \in \mathbb{G}_n$ , under the assumption (H.3)

$$\rho = \mathbb{E}[V_{2k} V_{2k+1} | \mathcal{F}_n] \quad \text{a.s.}$$

Hence, we propose to estimate the conditional covariance  $\rho$  by

$$\hat{\rho}_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} \hat{V}_{2k} \hat{V}_{2k+1} \quad (3.6)$$

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

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

For all  $n \geq 0$  and for all  $k \in \mathbb{G}_n$ , denote  $v_{2k} = V_{2k}^2 - \mathbb{E}[V_{2k}^2 | \mathcal{F}_n]$ . We deduce from (3.5) that for all  $n \geq 1$

$$V_{2n}^2 = \eta^t \Phi_n + v_{2n}$$

where  $\eta^t = (\sigma_a^2 \ \sigma_c^2)$ . It leads us to estimate the vector of variances  $\eta$  by the WLS estimator

$$\hat{\eta}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} \hat{V}_{2k}^2 \Phi_k \quad (3.7)$$

where

$$Q_n = \sum_{k \in \mathbb{T}_n} \frac{1}{d_k} \Phi_k \Phi_k^t$$



and the weighting sequence  $(d_n)_{n \geq 1}$  is given, for all  $n \geq 1$ , by  $d_n = (1 + X_n)^2$ . This choice is due to the fact that for all  $n \geq 1$  and for all  $k \in \mathbb{G}_n$

$$\begin{aligned} \mathbb{E}[v_{2k}^2 | \mathcal{F}_n] &= \mathbb{E}[V_{2k}^4 | \mathcal{F}_n] - (\mathbb{E}[V_{2k}^2 | \mathcal{F}_n])^2 \quad \text{a.s.} \\ &= 2\sigma_a^4 X_k^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2) X_k + \mu_c^4 - \sigma_c^4 \quad \text{a.s.} \end{aligned} \tag{3.8}$$

where we recall that  $\mu_a^4$  is the fourth-order centered moment of  $(Y_{n,i})$ . Consequently, as  $d_n \geq 1$ , we clearly have for all  $n \geq 1$  and for all  $k \in \mathbb{G}_n$

$$\mathbb{E}[v_{2k}^2 | \mathcal{F}_n] \leq (\mu_a^4 - \sigma_a^4 + 4\sigma_a^2 \sigma_c^2 + \mu_c^4 - \sigma_c^4) d_k \quad \text{a.s.}$$

We have a similar WLS estimator  $\widehat{\xi}_n^t$  of the vector of variances  $\xi^t = (\sigma_b^2 \sigma_d^2)$  by replacing  $\widehat{V}_{2k}^2$  by  $\widehat{V}_{2k+1}^2$  into (3.7).

### 4 A martingale approach

In order to establish all the asymptotic properties of our estimators, we shall make use of a martingale approach. For all  $n \geq 1$ , denote

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

We can clearly rewrite (3.4) as

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n. \tag{4.1}$$

As in [Bercu et al. \(2009\)](#), we make use of the notation  $M_n$  since it appears that  $(M_n)_{n \geq 1}$  a martingale. This fact is a crucial point of our study and it justifies the vector notation since most of asymptotic results for martingales were established for vector-valued martingales. Let us rewrite  $M_n$  in order to emphasize its martingale quality. Let  $\Psi_n = I_2 \otimes \varphi_n$  where  $\varphi_n$  is the matrix of dimension  $2 \times 2^n$  given by

$$\varphi_n = \begin{pmatrix} \frac{X_{2^n}}{\sqrt{c_{2^n}}} & \frac{X_{2^{n+1}}}{\sqrt{c_{2^{n+1}}}} & \cdots & \frac{X_{2^{n+1-1}}}{\sqrt{c_{2^{n+1-1}}}} \\ 1 & 1 & \cdots & 1 \\ \frac{X_{2^n}}{\sqrt{c_{2^n}}} & \frac{X_{2^{n+1}}}{\sqrt{c_{2^{n+1}}}} & \cdots & \frac{X_{2^{n+1-1}}}{\sqrt{c_{2^{n+1-1}}}} \end{pmatrix}.$$

It represents the individuals of the  $n$ -th generation which is also the collection of all  $\Phi_k / \sqrt{c_k}$  where  $k$  belongs to  $\mathbb{G}_n$ . Let  $\xi_n$  be the random vector of dimension  $2^n$

$$\xi_n^t = \left( \frac{V_{2^n}}{\sqrt{c_{2^n-1}}} \quad \frac{V_{2^{n+2}}}{\sqrt{c_{2^{n+1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1-2}}}{\sqrt{c_{2^n-1}}} \quad \frac{V_{2^{n+1}}}{\sqrt{c_{2^{n-1}}}} \quad \frac{V_{2^{n+3}}}{\sqrt{c_{2^{n+1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1-1}}}{\sqrt{c_{2^n-1}}} \right).$$

The vector  $\xi_n$  gathers the noise variables of  $\mathbb{G}_n$ . The special ordering separating odd and even indices has been made in [Bercu et al. \(2009\)](#) so that  $M_n$  can be rewritten as

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

Under **(H.1)**, we clearly have for all  $n \geq 0$ ,  $\mathbb{E}[\xi_{n+1} | \mathcal{F}_n] = 0$  a.s. and  $\Psi_n$  is  $\mathcal{F}_n$ -measurable. In addition it is not hard to see that under **(H.1)** to **(H.3)**,  $(M_n)$  is a locally square integrable

vector martingale with increasing process given, for all  $n \geq 1$ , by

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

where

$$L_k = \sum_{i \in G_k} \frac{1}{c_i^2} \begin{pmatrix} \sigma_a^2 X_i + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_i + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1 \end{pmatrix}. \tag{4.3}$$

It is necessary to establish the convergence of  $\langle M \rangle_n$ , properly normalized, in order to prove the asymptotic results for our BINAR estimators  $\widehat{\theta}_n$ ,  $\widehat{\eta}_n$  and  $\widehat{\zeta}_n$ . Since the sizes of  $\Psi_n$  and  $\xi_n$  double at each generation, we have to adapt the proof of vector-valued martingale convergence given in [Duflo \(1997\)](#) to our framework.

### 5 Main results

In all the sequel, we will assume that the law of the immigration  $(\varepsilon_{2n}, \varepsilon_{2n+1})$  does not depend on  $n$ . However, we shall get rid of the standard assumption commonly used in the INAR literature that the offspring sequences  $(Y_{n,i})$  and  $(Z_{n,i})$  share the same Bernoulli distribution. The only assumption that we will use here is that the offspring sequences  $(Y_{n,i})$  and  $(Z_{n,i})$  admit eighth-order moments. We have to introduce some more notations in order to state our main results. From the original process  $(X_n)_{n \geq 1}$ , we shall define a new process  $(Y_n)_{n \geq 1}$  recursively given by  $Y_1 = X_1$ , and if  $Y_n = X_k$  with  $n, k \geq 1$ , then

$$Y_{n+1} = X_{2k+\kappa_n}$$

where  $(\kappa_n)_{n \geq 1}$  is a sequence of i.i.d. random variables with Bernoulli  $\mathcal{B}(1/2)$  distribution. Such a construction may be found in [Guyon \(2007\)](#) for the asymptotic analysis of BAR processes. The process  $(Y_n)$  gathers the values of the original process  $(X_n)$  along the random branch of the binary tree  $(\mathbb{T}_n)$  given by  $(\kappa_n)$ . Denote by  $k_n$  the unique  $k \geq 1$  such that  $Y_n = X_k$ . Then, for all  $n \geq 1$ , we have

$$Y_{n+1} = a_{n+1} \circ Y_n + e_{n+1} \tag{5.1}$$

where

$$a_{n+1} = \begin{cases} a & \text{if } \kappa_n = 0 \\ b & \text{otherwise} \end{cases} \quad \text{and} \quad e_n = \varepsilon_{k_n}. \tag{5.2}$$

**Lemma 5.1** *Assume that  $(\varepsilon_n)$  satisfies (H.1) to (H.4). Then, we have*

$$Y_n \xrightarrow{\mathcal{L}} T$$

where  $T$  is a positive non degenerate integer-valued random variable with  $\mathbb{E}[T^3] < \infty$ .

Denote  $\mathcal{C}_3^1(\mathbb{R}_+) = \left\{ f \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}) \mid \exists \gamma > 0, \forall x \geq 0, |f'(x)| + |f(x)| \leq \gamma(1+x^3) \right\}$ .

**Lemma 5.2** *Assume that  $(\varepsilon_n)$  satisfies (H.1) to (H.5). Then, for all  $f \in \mathcal{C}_3^1(\mathbb{R}_+)$ , we have*

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)] \quad \text{a.s.}$$

**Proposition 5.3** Assume that  $(\varepsilon_n)$  satisfies **(H.1)** to **(H.5)**. Then, we have

$$\lim_{n \rightarrow \infty} \frac{\langle M \rangle_n}{|\mathbb{T}_{n-1}|} = L \quad a.s. \tag{5.3}$$

where  $L$  is the positive definite matrix given by

$$L = \mathbb{E} \left[ \frac{1}{(1+T)^2} \begin{pmatrix} \sigma_a^2 T + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 T + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right].$$

Our first result deals with the almost sure convergence of our WLS estimator  $\widehat{\theta}_n$ .

**Theorem 5.4** Assume that  $(\varepsilon_n)$  satisfies **(H.1)** to **(H.5)**. Then,  $\widehat{\theta}_n$  converges almost surely to  $\theta$  with the rate of convergence

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O} \left( \frac{n}{|\mathbb{T}_{n-1}|} \right) \quad a.s. \tag{5.4}$$

In addition, we also have the quadratic strong law

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \Lambda (\widehat{\theta}_k - \theta) = \text{tr}(\Lambda^{-1/2} L \Lambda^{-1/2}) \quad a.s. \tag{5.5}$$

where

$$\Lambda = I_2 \otimes A \quad \text{and} \quad A = \mathbb{E} \left[ \frac{1}{1+T} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right]. \tag{5.6}$$

Our second result concerns the almost sure asymptotic properties of our WLS variance and covariance estimators  $\widehat{\eta}_n$ ,  $\widehat{\zeta}_n$  and  $\widehat{\rho}_n$ . Let

$$\begin{aligned} \eta_n &= Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k}^2 \Phi_k, \\ \zeta_n &= Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k+1}^2 \Phi_k, \\ \rho_n &= \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1}} V_{2k} V_{2k+1}. \end{aligned}$$

**Theorem 5.5** Assume that  $(\varepsilon_n)$  satisfies **(H.1)** to **(H.5)**. Then,  $\widehat{\eta}_n$  and  $\widehat{\zeta}_n$  converge almost surely to  $\eta$  and  $\zeta$  respectively. More precisely,

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

$$\|\widehat{\zeta}_n - \zeta_n\| = \mathcal{O} \left( \frac{n}{|\mathbb{T}_{n-1}|} \right) \quad a.s. \tag{5.8}$$

In addition,  $\widehat{\rho}_n$  converges almost surely to  $\rho$  with

$$\widehat{\rho}_n - \rho_n = \mathcal{O} \left( \frac{n}{|\mathbb{T}_{n-1}|} \right) \quad a.s. \tag{5.9}$$

**Remark 5.6** We also have the almost sure rates of convergence

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

Our last result is devoted to the asymptotic normality of our WLS estimators  $\widehat{\theta}_n, \widehat{\eta}_n, \widehat{\zeta}_n$  and  $\widehat{\rho}_n$ .

**Theorem 5.7** *Assume that  $(\varepsilon_n)$  satisfies (H.1) to (H.5). Then, we have the asymptotic normality*

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (I_2 \otimes A^{-1})L(I_2 \otimes A^{-1})). \tag{5.10}$$

In addition, we also have

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1}M_{ac}B^{-1}), \tag{5.11}$$

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\zeta}_n - \zeta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1}M_{bd}B^{-1}), \tag{5.12}$$

where

$$B = \mathbb{E} \left[ \frac{1}{(1+T)^2} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right],$$

$$M_{ac} = \mathbb{E} \left[ \frac{2\sigma_a^4 T^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2)T + \mu_c^4 - \sigma_c^4}{(1+T)^4} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right],$$

$$M_{bd} = \mathbb{E} \left[ \frac{2\sigma_b^4 T^2 + (\mu_b^4 - 3\sigma_b^4 + 4\sigma_b^2 \sigma_d^2)T + \mu_d^4 - \sigma_d^4}{(1+T)^4} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right].$$

Finally,

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2) \tag{5.13}$$

where

$$\sigma_\rho^2 = \sigma_a^2 \sigma_b^2 \mathbb{E}[T^2] + (\sigma_a^2 \sigma_d^2 + \sigma_b^2 \sigma_c^2) \frac{\bar{c}}{1-\bar{a}} + v^2 - \rho^2,$$

$$\mathbb{E}[T^2] = \frac{\Upsilon \bar{c}}{1-\bar{a}} + \frac{\bar{c}^2 - \Upsilon \bar{c}}{1-\bar{a}^2} + \frac{2(ac+bd)\bar{c}}{(1-\bar{a})(1-\bar{a}^2)},$$

$$\Upsilon = \frac{\sigma_a^2 + \sigma_b^2}{2(\bar{a} - \bar{a}^2)}, \quad \bar{a} = \frac{a+b}{2}, \quad \bar{a}^2 = \frac{a^2 + b^2}{2},$$

$$\bar{c} = \frac{c+d}{2}, \quad \bar{c}^2 = \frac{\sigma_c^2 + \sigma_d^2 + c^2 + d^2}{2}. \tag{5.14}$$

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

### 6 Proof of Lemma 5.1

We can reformulate (5.1) and (5.2) as

$$Y_n = a_n \circ a_{n-1} \circ \dots \circ a_2 \circ Y_1 + \sum_{k=2}^{n-1} a_n \circ a_{n-1} \circ \dots \circ a_{k+1} \circ e_k + e_n.$$

We already made the assumption that the law of the immigration  $(\varepsilon_{2n}, \varepsilon_{2n+1})$  does not depend on  $n$ . Consequently, the couples  $(a_k, e_k)$  and  $(a_{n-k+2}, e_{n-k+2})$  share the same distribution.

Hence, for all  $n \geq 2$ ,  $Y_n$  has the same distribution than the random variable

$$\begin{aligned} Z_n &= a_2 \circ \dots \circ a_n \circ Y_1 + \sum_{k=2}^{n-1} a_2 \circ a_3 \circ \dots \circ a_{n-k+1} \circ e_{n-k+2} + e_2, \\ &= a_2 \circ \dots \circ a_n \circ Y_1 + \sum_{k=3}^n a_2 \circ a_3 \circ \dots \circ a_{k-1} \circ e_k + e_2. \end{aligned}$$

For the sake of simplicity, we will denote

$$Z_n = a_2 \circ \dots \circ a_n \circ Y_1 + \sum_{k=2}^n a_2 \circ a_3 \circ \dots \circ a_{k-1} \circ e_k. \tag{6.1}$$

For all  $n \geq 2$  and for all  $2 \leq k \leq n$ , let  $\Sigma_n^{n-k+2} = a_k \circ \dots \circ a_n \circ Y_1$  and  $\Sigma_n^1 = Y_1$ . We clearly have  $\Sigma_n^{n-k+2} = a_k \circ \Sigma_n^{n-k+1}$ . Consequently, it follows from the tower property of the conditional expectation that

$$\mathbb{E}[\Sigma_n^n] = \mathbb{E}[a_2 \circ \Sigma_n^{n-1}] = \mathbb{E}[a \circ \Sigma_n^{n-1}] \mathbb{P}(a_2 = a) + \mathbb{E}[b \circ \Sigma_n^{n-1}] \mathbb{P}(a_2 = b),$$

leading to

$$\begin{aligned} \mathbb{E}[\Sigma_n^n] &= \frac{1}{2} \left( \mathbb{E} \left[ \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i} \mid \Sigma_n^{n-1} \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} Z_{2,i} \mid \Sigma_n^{n-1} \right] \right] \right), \\ &= \frac{1}{2} \left( \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}[Y_{2,i}] \right] + \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}[Z_{2,i}] \right] \right), \\ &= \frac{1}{2} (\mathbb{E}[a \Sigma_n^{n-1}] + \mathbb{E}[b \Sigma_n^{n-1}]) = \bar{a} \mathbb{E}[\Sigma_n^{n-1}] = \dots = \bar{a}^{n-1} \mathbb{E}[\Sigma_n^1] = \bar{a}^{n-1} \mathbb{E}[Y_1]. \end{aligned}$$

The stability hypothesis  $0 < \max(a, b) < 1$  implies that  $0 < \bar{a} < 1$  which leads to

$$\sum_{n=2}^{\infty} \mathbb{E}[\Sigma_n^n] = \mathbb{E}[Y_1] \sum_{n=2}^{\infty} \bar{a}^{n-1} = \frac{\mathbb{E}[Y_1] \bar{a}}{1 - \bar{a}}.$$

Then, we obtain from the monotone convergence theorem that

$$\lim_{n \rightarrow \infty} \Sigma_n^n = 0 \quad \text{a.s.} \tag{6.2}$$

It now remains to study the right-hand side sum in (6.1). For all  $n \geq 2$ , denote

$$T_n = \sum_{k=2}^n a_2 \circ \dots \circ a_{k-1} \circ e_k.$$

By the same calculation as before, we have for all  $n \geq 2$

$$\mathbb{E}[T_n] = \sum_{k=2}^n \bar{a}^{k-2} \mathbb{E}[e_k] = \bar{c} \sum_{k=0}^{n-2} \bar{a}^k,$$

which implies that

$$\lim_{n \rightarrow \infty} \mathbb{E}[T_n] = \frac{\bar{c}}{1 - \bar{a}}.$$

Hence, the positive increasing sequence  $(T_n)$  converges almost surely to

$$T = \sum_{k=2}^{\infty} a_2 \circ \dots \circ a_{k-1} \circ e_k$$

which is almost surely finite, thanks to the monotone convergence, as  $\mathbb{E}[T] < \infty$ . Therefore, we can conclude from (6.1) and (6.2) that

$$\lim_{n \rightarrow \infty} Z_n = T \quad \text{a.s.}$$

leading to

$$Y_n \xrightarrow{\mathcal{L}} T.$$

Let us prove that  $\mathbb{E}[T^3] < \infty$ . First of all, we already saw that

$$\mathbb{E}[a_2 \circ \dots \circ a_n \circ e_{n+1}] = \bar{a}^{n-1} \mathbb{E}[e_2] = \bar{a}^{n-1} \bar{c}.$$

In addition,

$$\begin{aligned} \mathbb{E}[(\Sigma_n^n)^2] &= \frac{1}{2} (\mathbb{E}[(a \circ \Sigma_n^{n-1})^2] + \mathbb{E}[(b \circ \Sigma_n^{n-1})^2]), \\ &= \frac{1}{2} \left( \mathbb{E} \left[ \mathbb{E} \left[ \left( \sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i} \right)^2 \middle| \Sigma_n^{n-1} \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ \left( \sum_{i=1}^{\Sigma_n^{n-1}} Z_{2,i} \right)^2 \middle| \Sigma_n^{n-1} \right] \right] \right), \end{aligned}$$

and the first expectation is

$$\begin{aligned} \mathbb{E} \left[ \mathbb{E} \left[ \left( \sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i} \right)^2 \middle| \Sigma_n^{n-1} \right] \right] &= \mathbb{E} \left[ \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i}^2 + \sum_{\substack{i=1 \\ j=1 \\ j \neq i}}^{\Sigma_n^{n-1}} \sum_{j=1}^{\Sigma_n^{n-1}} Y_{2,i} Y_{2,j} \middle| \Sigma_n^{n-1} \right] \right], \\ &= \mathbb{E} \left[ \sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}[Y_{2,i}^2] + \sum_{\substack{i=1 \\ j=1 \\ j \neq i}}^{\Sigma_n^{n-1}} \sum_{j=1}^{\Sigma_n^{n-1}} \mathbb{E}[Y_{2,i}] \mathbb{E}[Y_{2,j}] \right], \\ &= \mathbb{E}[\Sigma_n^{n-1}(\sigma_a^2 + a^2) + \Sigma_n^{n-1}(\Sigma_n^{n-1} - 1)a^2], \\ &= \mathbb{E}[\Sigma_n^{n-1}]\sigma_a^2 + a^2\mathbb{E}[(\Sigma_n^{n-1})^2]. \end{aligned}$$

Since the computation of the second expectation is exactly the same, we obtain

$$\begin{aligned} \mathbb{E}[(\Sigma_n^n)^2] &= \mathbb{E}[\Sigma_n^{n-1}] \frac{\sigma_a^2 + \sigma_b^2}{2} + \bar{a}^2 \mathbb{E}[(\Sigma_n^{n-1})^2], \\ &= \bar{a}^{n-2} \frac{\sigma_a^2 + \sigma_b^2}{2} \mathbb{E}[Y_1] + \bar{a}^2 \mathbb{E}[(\Sigma_n^{n-1})^2] = \dots \\ &= \left( \sum_{i=0}^{n-2} \bar{a}^{n-i-2} \bar{a}^{2i} \right) \frac{\sigma_a^2 + \sigma_b^2}{2} \mathbb{E}[Y_1] + \bar{a}^{2n-1} \mathbb{E}[(\Sigma_n^1)^2], \\ &= \frac{\bar{a}^{n-1} - \bar{a}^{2n-1}}{\bar{a} - \bar{a}^2} \frac{\sigma_a^2 + \sigma_b^2}{2} \mathbb{E}[Y_1] + \bar{a}^{2n-1} \mathbb{E}[Y_1^2], \end{aligned}$$

$$= (\bar{a}^{n-1} - \bar{a}^{2^{n-1}}) \Upsilon \mathbb{E}[Y_1] + \bar{a}^{2^{n-1}} \mathbb{E}[Y_1^2]$$

where

$$\Upsilon = \frac{\sigma_a^2 + \sigma_b^2}{2(\bar{a} - a^2)}.$$

In the same way, we can prove that

$$\mathbb{E}[(a_2 \circ \dots \circ a_n \circ e_{n+1})^2] = (\bar{a}^{n-1} - \bar{a}^{2^{n-1}}) \Upsilon \bar{c} + \bar{a}^{2^{n-1}} \bar{c}^2.$$

Consequently, as  $(e_n)$  is an integer-valued random variable,

$$\mathbb{E}[(a_2 \circ \dots \circ a_n \circ e_{n+1})^2] \leq \bar{a}^{n-1} (\Upsilon \bar{c} + \bar{c}^2) \leq \bar{a}^{n-1} (\Upsilon + 1) \bar{c}^2.$$

Furthermore, we obtain from tedious but straightforward calculations that it exists some constant  $\xi > 0$  such that for all  $2 \leq p \leq 8$

$$\mathbb{E}[(a_2 \circ \dots \circ a_n \circ e_{n+1})^p] \leq \xi \mathbb{E}[e_2^p] \bar{a}^{n-1}. \tag{6.3}$$

One can observe that the constant  $\xi$  only depends on the moments of  $(Y_{n,i})$  and  $(Z_{n,i})$  up to order 8. Hence, as  $0 < \bar{a} < 1$ , we deduce from (6.3) and the triangle inequality that

$$\begin{aligned} \mathbb{E}[T^3]^{1/3} &\leq \sum_{k=2}^{\infty} \mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k)^3]^{1/3}, \\ &\leq \xi^{1/3} \mathbb{E}[e_2^3]^{1/3} \sum_{k=2}^{\infty} \bar{a}^{(k-2)/3} < \infty \end{aligned}$$

which immediately leads to  $\mathbb{E}[T^3] < \infty$ . Finally, let us compute  $\text{Var}(T)$  in order to prove that  $T$  is not degenerate. First, one can observe that

$$\begin{aligned} \mathbb{E}[T^2] &= \mathbb{E} \left[ \left( \sum_{k=2}^{\infty} a_2 \circ \dots \circ a_{k-1} \circ e_k \right)^2 \right], \\ &= \sum_{k=2}^{\infty} \mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k)^2] \\ &\quad + 2 \sum_{k=2}^{\infty} \sum_{l=k+1}^{\infty} \mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k) (a_2 \circ \dots \circ a_{l-1} \circ e_l)] \end{aligned}$$

We already saw that

$$\mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k)^2] = (\bar{a}^{k-2} - \bar{a}^{2^{k-2}}) \Upsilon \bar{c} + \bar{a}^{2^{k-2}} \bar{c}^2.$$

Moreover, we have, for all  $l \geq 3$

$$\begin{aligned} \mathbb{E} [e_2(a_2 \circ \dots \circ a_{l-1} \circ e_l)] &= \frac{1}{2} \mathbb{E} [\varepsilon_2(a \circ \dots \circ a_{l-1} \circ e_l)] + \frac{1}{2} \mathbb{E} [\varepsilon_3(b \circ \dots \circ a_{l-1} \circ e_l)], \\ &= \frac{1}{2} (\mathbb{E} [\varepsilon_2] \mathbb{E} [(a \circ \dots \circ a_{l-1} \circ e_l)] \\ &\quad + \mathbb{E} [\varepsilon_3] \mathbb{E} [(b \circ \dots \circ a_{l-1} \circ e_l)]), \\ &= \frac{1}{2} (c(a\bar{a}^{l-3}\bar{c}) + d(b\bar{a}^{l-3}\bar{c})), \\ &= \frac{(ac + bd)\bar{c}}{2} \bar{a}^{l-3}. \end{aligned}$$

In addition, for all  $k \geq 2$  and for all  $l \geq k + 1$

$$\begin{aligned} &\mathbb{E} [(a_2 \circ \dots \circ a_{k-1} \circ e_k) (a_2 \circ \dots \circ a_{l-1} \circ e_l)] \\ &= \frac{1}{2} \mathbb{E} [(a \circ \dots \circ a_{k-1} \circ e_k) (a \circ \dots \circ a_{l-1} \circ e_l)] \\ &\quad + \frac{1}{2} \mathbb{E} [(b \circ \dots \circ a_{k-1} \circ e_k) (b \circ \dots \circ a_{l-1} \circ e_l)]. \end{aligned}$$

Let us tackle the first term

$$\begin{aligned} &\mathbb{E} [(a \circ \dots \circ a_{k-1} \circ e_k) (a \circ \dots \circ a_{l-1} \circ e_l)] \\ &= \mathbb{E} \left[ \left( \sum_{i=1}^{a_3 \circ \dots \circ a_{k-1} \circ e_k} Y_{k,i} \right) \left( \sum_{j=1}^{a_3 \circ \dots \circ a_{l-1} \circ e_l} Y_{l,j} \right) \right], \\ &= \mathbb{E} \left[ \sum_{i=1}^{a_3 \circ \dots \circ a_{k-1} \circ e_k} \sum_{j=1}^{a_3 \circ \dots \circ a_{l-1} \circ e_l} \mathbb{E} [Y_{k,i} Y_{l,j} | a_3 \circ \dots \circ a_{k-1} \circ e_k, a_3 \circ \dots \circ a_{l-1} \circ e_l] \right], \\ &= \mathbb{E} \left[ \sum_{i=1}^{a_3 \circ \dots \circ a_{k-1} \circ e_k} \sum_{j=1}^{a_3 \circ \dots \circ a_{l-1} \circ e_l} \mathbb{E} [Y_{k,i} Y_{l,j}] \right], \\ &= \mathbb{E} \left[ \sum_{i=1}^{a_3 \circ \dots \circ a_{k-1} \circ e_k} \sum_{j=1}^{a_3 \circ \dots \circ a_{l-1} \circ e_l} a^2 \right], \\ &= a^2 \mathbb{E} [(a_3 \circ \dots \circ a_{k-1} \circ e_k) (a_3 \circ \dots \circ a_{l-1} \circ e_l)]. \end{aligned}$$

Hence, we obtained that

$$\begin{aligned} &\mathbb{E} [(a_2 \circ \dots \circ a_{k-1} \circ e_k) (a_2 \circ \dots \circ a_{l-1} \circ e_l)] \\ &= \bar{a}^2 \mathbb{E} [(a_3 \circ \dots \circ a_{k-1} \circ e_k) (a_3 \circ \dots \circ a_{l-1} \circ e_l)], \\ &= \bar{a}^{2k-2} \mathbb{E} [e_k (a_k \circ \dots \circ a_{l-1} \circ e_l)], \\ &= \bar{a}^{2k-2} \frac{(ac + bd)\bar{c}}{2} \bar{a}^{l-k-1}. \end{aligned}$$



Finally, we have

$$\begin{aligned} \mathbb{E}[T^2] &= \sum_{k=2}^{\infty} \left( (\bar{a}^{k-2} - \bar{a}^{2k-2}) \Upsilon \bar{c} + \bar{a}^{2k-2} \bar{c}^2 \right) + 2 \sum_{l=3}^{\infty} \frac{(ac + bd)\bar{c}}{2} \bar{a}^{l-3} \\ &\quad + 2 \sum_{k=3}^{\infty} \sum_{l=k+1}^{\infty} \frac{\bar{a}^{2k-2} (ac + bd)\bar{c}}{2} \bar{a}^{l-k-1}, \\ &= \Upsilon \bar{c} \left( \frac{1}{1-\bar{a}} - \frac{1}{1-\bar{a}^2} \right) + \frac{\bar{c}^2}{1-\bar{a}^2} + \frac{(ac + bd)\bar{c}}{1-\bar{a}} \left( 1 + \frac{\bar{a}^2}{1-\bar{a}^2} \right), \\ &= \Upsilon \bar{c} \left( \frac{1}{1-\bar{a}} - \frac{1}{1-\bar{a}^2} \right) + \frac{\bar{c}^2}{1-\bar{a}^2} + \frac{(ac + bd)\bar{c}}{(1-\bar{a})(1-\bar{a}^2)}. \end{aligned}$$

To conclude, we can compute the variance of  $T$

$$\begin{aligned} \text{Var}(T) &= \mathbb{E}[T^2] - \mathbb{E}[T]^2, \\ &= \Upsilon \bar{c} \left( \frac{1}{1-\bar{a}} - \frac{1}{1-\bar{a}^2} \right) + \frac{\sigma_c^2 + \sigma_d^2}{2(1-\bar{a}^2)} \\ &\quad + \frac{c^2 + d^2}{2(1-\bar{a}^2)} + \frac{(ac + bd)\bar{c}}{(1-\bar{a})(1-\bar{a}^2)} - \left( \frac{\bar{c}}{1-\bar{a}} \right)^2, \\ &= \Upsilon \bar{c} \left( \frac{1}{1-\bar{a}} - \frac{1}{1-\bar{a}^2} \right) + \frac{\sigma_c^2 + \sigma_d^2}{2(1-\bar{a}^2)} + \frac{2(ad - bc + c - d)^2}{(2 - (a^2 + b^2))(2 - (a + b))^2}. \end{aligned}$$

The first and the third terms of this sum are clearly nonnegative since  $\max(a, b) < 1$ , and the second term is clearly positive under **(H.2)**. This allows us to say that the variance of  $T^2$  is positive and  $T$  is not degenerate.

### 7 Proof of Lemma 5.2

We shall now prove that for all  $f \in \mathcal{C}_3^1(\mathbb{R}_+)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)]. \tag{7.1}$$

Denote  $g = f - \mathbb{E}[f(T)]$ ,

$$\bar{M}_{\mathbb{T}_n}(f) = \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} f(X_k) \quad \text{and} \quad \bar{M}_{\mathbb{G}_n}(f) = \frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} f(X_k).$$

Via Lemma A.2 of [Bercu et al. \(2009\)](#), it is only necessary to prove that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} g(X_k) = 0 \quad \text{a.s.}$$

We shall follow the induced Markov chain approach, originally proposed by in [Guyon \(2007\)](#). Let  $Q$  be the transition probability of  $(Y_n)$ ,  $Q^p$  the  $p$ -th iterated of  $Q$ . In addition, denote by  $\nu$  the distribution of  $Y_1 = X_1$  and  $\nu Q^p$  the law of  $Y_p$ . Finally, let  $P$  be the transition

probability of  $(X_n)$  as defined in Guyon (2007). We obtain from relation (7) of Guyon (2007) that for all  $n \geq 0$

$$\mathbb{E}[\overline{M}_{G_n}(g)^2] = \frac{1}{2^n} \nu Q^n g^2 + \sum_{k=0}^{n-1} \frac{1}{2^{k+1}} \nu Q^k P(Q^{n-k-1} g \star Q^{n-k-1} g)$$

where, for all  $x, y \in \mathbb{N}$ ,  $(f \star g)(x, y) = f(x)g(y)$ . Consequently,

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{E}[\overline{M}_{G_n}(g)^2] &= \sum_{n=0}^{\infty} \frac{1}{2^n} \nu Q^n g^2 + \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \frac{1}{2^{k+1}} \nu Q^k P(Q^{n-k-1} g \star Q^{n-k-1} g), \\ &\leq \sum_{k=0}^{\infty} \frac{1}{2^k} \nu Q^k \left( g^2 + P \left( \sum_{l=0}^{\infty} |Q^l g \star Q^l g| \right) \right). \end{aligned}$$

However, for all  $x \in \mathbb{N}$ ,

$$Q^n g(x) = Q^n f(x) - \mathbb{E}[f(T)] = \mathbb{E}_x[f(Y_n) - f(T)] = \mathbb{E}_x[f(Z_n) - f(T)]$$

where  $Z_n$  is given by (6.1). Hence, we deduce from the mean value theorem and Cauchy–Schwarz inequality that

$$|Q^n g(x)| \leq \mathbb{E}_x[W_n |Z_n - T|] \leq \mathbb{E}_x[W_n^2]^{1/2} \mathbb{E}_x[(Z_n - T)^2]^{1/2} \tag{7.2}$$

where

$$W_n = \sup_{z \in [Z_n, T]} |f'(z)|.$$

By the very definition of  $C_3^1(\mathbb{R}_+)$ , one can find some constant  $\alpha > 0$  such that  $|f'(z)| \leq \alpha(1 + z^6)$ . Hence, it exists some constant  $\beta > 0$  such that

$$\begin{aligned} \mathbb{E}_x[W_n^2] &\leq \alpha \mathbb{E}_x[1 + Z_n^6 + T^6] = \alpha(1 + \mathbb{E}_x[Z_n^6] + \mathbb{E}[T^6]), \\ &\leq \beta(1 + x^6). \end{aligned} \tag{7.3}$$

As a matter of fact, under hypotheses **(H.1)** to **(H.5)**,  $\mathbb{E}[T^6] < \infty$  and it exists some constant  $\gamma > 0$  such that  $\mathbb{E}_x[Z_n^6] < \gamma(1 + x^6)$ . Let us first deal with  $\mathbb{E}[T^6]$ . The triangle inequality, together with  $0 < \bar{a} < 1$  and (6.3) allow us to say that

$$\mathbb{E}[T^6]^{1/6} \leq \sum_{k=2}^{\infty} \mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k)^6]^{1/6} \leq \xi^{1/6} \mathbb{E}[e_2^6]^{1/6} \sum_{k=2}^{\infty} \bar{a}^{(k-2)/6} < \infty$$

which immediately leads to  $\mathbb{E}[T^6] < \infty$ . One the other hand, we infer from (6.1) that

$$\begin{aligned} \mathbb{E}_x[Z_n^6]^{1/6} &\leq \mathbb{E}_x[(a_2 \circ \dots \circ a_n \circ Y_1)^6]^{1/6} + \sum_{k=2}^n \mathbb{E}_x[(a_2 \circ a_3 \circ \dots \circ a_{k-1} \circ e_k)^6]^{1/6}, \\ &\leq \xi^{1/6} \mathbb{E}_x[Y_1^6]^{1/6} \bar{a}^{n-1} + \sum_{k=2}^{\infty} \mathbb{E}[(a_2 \circ a_3 \circ \dots \circ a_{k-1} \circ e_k)^6]^{1/6}, \\ &\leq \xi^{1/6} x + \sum_{k=2}^{\infty} \mathbb{E}[(a_2 \circ a_3 \circ \dots \circ a_{k-1} \circ e_k)^6]^{1/6} \end{aligned}$$

and we have already proved that the sum in the right-hand term is finite. So we can conclude that there exists some constant  $\gamma > 0$  such that  $\mathbb{E}_x[Z_n^6] < \gamma(1 + x^6)$ . Furthermore

$$Z_n - T = a_2 \circ \dots \circ a_n \circ Y_1 - \sum_{k=n+1}^{\infty} a_2 \circ \dots \circ a_k \circ e_{k+1}$$

and the triangle inequality allows us to say that

$$\mathbb{E}_x[(Z_n - T)^2]^{1/2} \leq \mathbb{E}_x[(a_2 \circ \dots \circ a_n \circ Y_1)^2]^{1/2} + \sum_{k=n+1}^{\infty} \mathbb{E}_x[(a_2 \circ \dots \circ a_k \circ e_{k+1})^2]^{1/2}.$$

We already saw in Sect. 6 that

$$\begin{aligned} \mathbb{E}_x[(a_2 \circ \dots \circ a_n \circ Y_1)^2] &= (\bar{a}^{n-1} - \bar{a}^{2^{n-1}})\Upsilon \mathbb{E}_x[Y_1] + \bar{a}^{2^{n-1}} \mathbb{E}_x[Y_1^2], \\ &= (\bar{a}^{n-1} - \bar{a}^{2^{n-1}})\Upsilon x + \bar{a}^{2^{n-1}} x^2 = x(\Upsilon \bar{a}^{n-1} + \bar{a}^{2^{n-1}}(x - \Upsilon)) \end{aligned}$$

and

$$\mathbb{E}_x[(a_2 \circ \dots \circ a_k \circ e_{k+1})^2] = (\bar{a}^{k-1} - \bar{a}^{2^{k-1}})\Upsilon \bar{c} + \bar{a}^{2^{k-1}} \bar{c}^2.$$

Hence

$$\begin{aligned} \sum_{k=n+1}^{\infty} \mathbb{E}_x[(a_2 \circ \dots \circ a_k \circ e_{k+1})^2]^{1/2} &= \sum_{k=n+1}^{\infty} \left( \bar{a}^{k-1} \Upsilon \bar{c} + \bar{a}^{2^{k-1}} (\bar{c}^2 - \Upsilon \bar{c}) \right)^{1/2}, \\ &\leq \sum_{k=n+1}^{\infty} \left( \bar{a}^{k-1} \bar{c} + \bar{a}^{k-1} |\bar{c}^2 - \Upsilon \bar{c}| \right)^{1/2}, \\ &\leq \sum_{k=n+1}^{\infty} \sqrt{\bar{a}}^{k-1} \delta = \delta \frac{\sqrt{\bar{a}}^n}{1 - \sqrt{\bar{a}}}. \end{aligned}$$

where

$$\delta = \sqrt{\max(\bar{c}^2, (1 + \Upsilon)\bar{c} - \bar{c}^2)}.$$

To sum up, we find that

$$\begin{aligned} \mathbb{E}_x[(Z_n - T)^2]^{1/2} &\leq \sqrt{x} \left( \Upsilon \bar{a}^{n-1} + \bar{a}^{2^{n-1}}(x - \Upsilon) \right)^{1/2} + \frac{\delta}{1 - \sqrt{\bar{a}}} \sqrt{\bar{a}}^n, \\ &\leq \begin{cases} \sqrt{x} (\Upsilon \bar{a}^{n-1} + \bar{a}^{2^{n-1}}(x - \Upsilon))^{1/2} + \frac{\delta}{1 - \sqrt{\bar{a}}} \sqrt{\bar{a}}^n & \text{if } x > \Upsilon, \\ \sqrt{x} \sqrt{\Upsilon} \sqrt{\bar{a}}^{n-1} + \frac{\delta}{1 - \sqrt{\bar{a}}} \sqrt{\bar{a}}^n & \text{if } x \leq \Upsilon, \end{cases} \\ &\leq \begin{cases} x \sqrt{\bar{a}}^{n-1} + \frac{\delta}{1 - \sqrt{\bar{a}}} \sqrt{\bar{a}}^n & \text{if } x > \Upsilon, \\ \frac{1+x}{2} \sqrt{\Upsilon} \sqrt{\bar{a}}^{n-1} + \frac{\delta}{1 - \sqrt{\bar{a}}} \sqrt{\bar{a}}^n & \text{if } x \leq \Upsilon, \end{cases} \\ &\leq \sqrt{\bar{a}}^n (1+x) \left( \frac{\sqrt{\Upsilon}}{2\sqrt{\bar{a}}} + \frac{\delta}{1 - \sqrt{\bar{a}}} \right). \end{aligned} \tag{7.4}$$

Finally, we obtain from (7.2) together with (7.3) and (7.4) that for some constant  $\kappa > 0$

$$|Q^n g(x)| \leq \sqrt{\beta}(1+x^6)^{1/2} \sqrt{a}^{n-1} (1+x) \left( \frac{\sqrt{\Upsilon}}{2} + \frac{\delta}{1-\sqrt{a}} \right) \leq \sqrt{a}^n \kappa (1+x^4).$$

Therefore,

$$P \left( \sum_{n=0}^{\infty} |Q^n g \star Q^n g| \right) \leq \frac{\kappa^2}{1-a} P(h \star h)$$

where, for all  $x \in \mathbb{N}$ ,  $h(x) = 1+x^4$ . We are now in position to prove that

$$\mathbb{E} \left[ \sum_{n=0}^{\infty} \overline{M}_{\mathbb{G}_n}(g)^2 \right] < \infty. \tag{7.5}$$

It is not hard to see that from hypothesis **(H.5)**, it exists some constant  $\lambda > 0$  such that for all  $x \in \mathbb{N}$ ,  $P(h \star h)(x) \leq \lambda(1+x^8)$ . Consequently, it exists some constant  $\mu > 0$  such that

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{E} [\overline{M}_{\mathbb{G}_n}(g)^2] &\leq \sum_{k=0}^{\infty} \frac{1}{2^k} \nu Q^k \left( g^2 + P \left( \sum_{l=0}^{\infty} |Q^l g \star Q^l g| \right) \right), \\ &\leq \sum_{k=0}^{\infty} \frac{1}{2^k} \left( \mathbb{E}[g^2(Y_k)] + \frac{\lambda \kappa^2}{1-a} (1 + \mathbb{E}[Y_k^8]) \right), \\ &\leq \left( 2\mu + \frac{\lambda \kappa^2}{1-a} \right) \left( 2 + \sum_{k=0}^{\infty} \frac{1}{2^k} \mathbb{E}[Y_k^8] \right). \end{aligned} \tag{7.6}$$

Furthermore, we can deduce from (6.3) that it exists some constant  $\zeta$  such that

$$\begin{aligned} \mathbb{E}[Y_n^8]^{1/8} &\leq \mathbb{E}[(a_2 \circ \dots \circ a_n \circ Y_1)^8]^{1/8} + \sum_{k=2}^n \mathbb{E}[(a_2 \circ \dots \circ a_{k-1} \circ e_k)^8]^{1/8}, \\ &\leq \mathbb{E}[(a_2 \circ \dots \circ a_n \circ Y_1)^8]^{1/8} + \xi^{1/8} \mathbb{E}[e_2^8]^{1/8} \sum_{k=2}^n a^{k-2}, \\ &\leq \zeta^{1/8} \mathbb{E}[Y_1^8]^{1/8} a^{n-1} + \frac{\xi^{1/8} \mathbb{E}[e_2^8]^{1/8}}{1-a}, \\ &\leq \frac{\zeta^{1/8} \mathbb{E}[Y_1^8]^{1/8} + \xi^{1/8} \mathbb{E}[e_2^8]^{1/8}}{1-a}. \end{aligned} \tag{7.7}$$

Then, (7.6) and (7.7) immediately lead to (7.5). Finally, the monotone convergence theorem implies that

$$\lim_{n \rightarrow \infty} \overline{M}_{\mathbb{G}_n}(g) = 0 \quad \text{a.s.}$$

which completes the proof of Lemma 5.2.

### 8 Proof of Proposition 5.3

The almost sure convergence (5.3) immediately follows from (4.2) and (4.3) together with Lemma 5.2. It only remains to prove that  $\det(L) > 0$  where the limiting matrix  $L$  can be

rewritten as

$$L = \mathbb{E}[\Gamma \otimes \mathcal{B}]$$

where

$$\Gamma = \begin{pmatrix} \sigma_a^2 T + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 T + \sigma_d^2 \end{pmatrix} \quad \text{and} \quad \mathcal{B} = \begin{pmatrix} \frac{T^2}{(1+T)^2} & \frac{T}{(1+T)^2} \\ \frac{T}{(1+T)^2} & \frac{1}{(1+T)^2} \end{pmatrix}.$$

We have

$$\begin{aligned} L &= \mathbb{E} \left[ \begin{pmatrix} \sigma_a^2 T & 0 \\ 0 & \sigma_b^2 T \end{pmatrix} \otimes \mathcal{B} \right] + \mathbb{E} \left[ \begin{pmatrix} \sigma_c^2 & \rho \\ \rho & \sigma_d^2 \end{pmatrix} \otimes \mathcal{B} \right], \\ &= \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \otimes \mathbb{E}[T\mathcal{B}] + \begin{pmatrix} \sigma_c^2 & \rho \\ \rho & \sigma_d^2 \end{pmatrix} \otimes \mathbb{E}[\mathcal{B}]. \end{aligned} \tag{8.1}$$

We shall prove that  $\mathbb{E}[\mathcal{B}]$  is a positive definite matrix and that  $\mathbb{E}[T\mathcal{B}]$  is a positive semidefinite matrix. Denote by  $\lambda_1$  and  $\lambda_2$  the two eigenvalues of the real symmetric matrix  $\mathbb{E}[\mathcal{B}]$ . We clearly have

$$\lambda_1 + \lambda_2 = \text{tr}(\mathbb{E}[\mathcal{B}]) = \mathbb{E} \left[ \frac{T^2 + 1}{(1+T)^2} \right] > 0$$

and

$$\lambda_1 \lambda_2 = \det(\mathbb{E}[\mathcal{B}]) = \mathbb{E} \left[ \frac{T^2}{(1+T)^2} \right] \mathbb{E} \left[ \frac{1}{(1+T)^2} \right] - \mathbb{E} \left[ \frac{T}{(1+T)^2} \right]^2 \geq 0$$

thanks to the Cauchy–Schwarz inequality and  $\lambda_1 \lambda_2 = 0$  if and only if  $T$  is degenerate, which is not the case thanks to Lemma 5.1. Consequently,  $\mathbb{E}[\mathcal{B}]$  is a positive definite matrix. In the same way, we can prove that  $\mathbb{E}[T\mathcal{B}]$  is a positive semidefinite matrix. Since the Kronecker product of two positive semidefinite (respectively positive definite) matrices is a positive semidefinite (respectively positive definite) matrix, we deduce from (8.1) that  $L$  is positive definite as soon as  $\rho^2 < \sigma_c^2 \sigma_d^2$  which is the case thanks to (H.3).

### 9 Proof of Theorem 5.4

We will follow the same approach as in Bercu et al. (2009). For all  $n \geq 1$ , let  $\mathcal{V}_n = M_n^t \Sigma_{n-1}^{-1} M_n = (\hat{\theta}_n - \theta)^t \Sigma_{n-1} (\hat{\theta}_n - \theta)$ . First of all, we have

$$\begin{aligned} \mathcal{V}_{n+1} &= M_{n+1}^t \Sigma_n^{-1} M_{n+1} = (M_n + \Delta M_{n+1})^t \Sigma_n^{-1} (M_n + \Delta M_{n+1}), \\ &= M_n^t \Sigma_n^{-1} M_n + 2M_n^t \Sigma_n^{-1} \Delta M_{n+1} + \Delta M_{n+1}^t \Sigma_n^{-1} \Delta M_{n+1}, \\ &= \mathcal{V}_n - M_n^t (\Sigma_{n-1}^{-1} - \Sigma_n^{-1}) M_n + 2M_n^t \Sigma_n^{-1} \Delta M_{n+1} + \Delta M_{n+1}^t \Sigma_n^{-1} \Delta M_{n+1}. \end{aligned}$$

By summing over this identity, we obtain the main decomposition

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

where

$$\begin{aligned} \mathcal{A}_n &= \sum_{k=1}^n M_k^t (\Sigma_{k-1}^{-1} - \Sigma_k^{-1}) M_k, \\ \mathcal{B}_{n+1} &= 2 \sum_{k=1}^n M_k^t \Sigma_k^{-1} \Delta M_{k+1} \quad \text{and} \quad \mathcal{W}_{n+1} = \sum_{k=1}^n \Delta M_{k+1}^t \Sigma_k^{-1} \Delta M_{k+1}. \end{aligned}$$

**Lemma 9.1** *Assume that  $(\varepsilon_n)$  satisfies (H.1) to (H.5). Then, we have*

$$\lim_{n \rightarrow \infty} \frac{\mathcal{W}_n}{n} = \frac{1}{2} \text{tr}((I_2 \otimes A)^{-1/2} L (I_2 \otimes A)^{-1/2}) \quad \text{a.s.} \tag{9.2}$$

where  $A$  is the positive definite matrix given by (5.6). In addition, we also have

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

and

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

*Proof* First of all, we have  $\mathcal{W}_{n+1} = \mathcal{T}_{n+1} + \mathcal{R}_{n+1}$  where

$$\begin{aligned} \mathcal{T}_{n+1} &= \sum_{k=1}^n \frac{\Delta M_{k+1}^t (I_2 \otimes A)^{-1} \Delta M_{k+1}}{|\mathbb{T}_k|}, \\ \mathcal{R}_{n+1} &= \sum_{k=1}^n \frac{\Delta M_{k+1}^t (|\mathbb{T}_k| \Sigma_k^{-1} - (I_2 \otimes A)^{-1}) \Delta M_{k+1}}{|\mathbb{T}_k|}. \end{aligned}$$

One can observe that  $\mathcal{T}_{n+1} = \text{tr}((I_2 \otimes A)^{-1/2} \mathcal{H}_{n+1} (I_2 \otimes A)^{-1/2})$  where

$$\mathcal{H}_{n+1} = \sum_{k=1}^n \frac{\Delta M_{k+1} \Delta M_{k+1}^t}{|\mathbb{T}_k|}.$$

Our aim is to make use of the strong law of large numbers for martingale transforms, so we start by adding and subtracting a term involving the conditional expectation of  $\Delta \mathcal{H}_{n+1}$  given  $\mathcal{F}_n$ . We have thanks to relation (4.3) that for all  $n \geq 0$ ,  $\mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] = L_n$ . Consequently, we can split  $\mathcal{H}_{n+1}$  into two terms

$$\mathcal{H}_{n+1} = \sum_{k=1}^n \frac{L_k}{|\mathbb{T}_k|} + \mathcal{K}_{n+1},$$

where

$$\mathcal{K}_{n+1} = \sum_{k=1}^n \frac{\Delta M_{k+1} \Delta M_{k+1}^t - L_k}{|\mathbb{T}_k|}.$$

It clearly follows from convergence (5.3) that

$$\lim_{n \rightarrow \infty} \frac{L_n}{|\mathbb{T}_n|} = \frac{1}{2} L \quad \text{a.s.}$$

Hence, Cesaro convergence immediately implies that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \frac{L_k}{|\mathbb{T}_k|} = \frac{1}{2} L \quad \text{a.s.} \tag{9.5}$$

On the other hand, the sequence  $(\mathcal{K}_n)_{n \geq 2}$  is obviously a square integrable martingale. Moreover, we have

$$\Delta \mathcal{K}_{n+1} = \mathcal{K}_{n+1} - \mathcal{K}_n = \frac{1}{|\mathbb{T}_n|} (\Delta M_{n+1} \Delta M_{n+1}^t - L_n).$$

For all  $u \in \mathbb{R}^4$ , denote  $\mathcal{K}_n(u) = u^t \mathcal{K}_n u$ . It follows from tedious but straightforward calculations, together with Lemma 5.2, that the increasing process of the martingale  $(\mathcal{K}_n(u))_{n \geq 2}$  satisfies  $\langle \mathcal{K}(u) \rangle_n = \mathcal{O}(n)$  a.s. Therefore, we deduce from the strong law of large numbers for martingales that for all  $u \in \mathbb{R}^4$ ,  $\mathcal{K}_n(u) = o(n)$  a.s. leading to  $\mathcal{K}_n = o(n)$  a.s. Hence, we infer from (9.5) that

$$\lim_{n \rightarrow \infty} \frac{\mathcal{H}_{n+1}}{n} = \frac{1}{2} L \quad \text{a.s.} \tag{9.6}$$

Via the same arguments as in the proof of convergence (5.3), we find that

$$\lim_{n \rightarrow \infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes A \quad \text{a.s.} \tag{9.7}$$

where  $A$  is the positive definite matrix given by (5.6). Then, we obtain from (9.6) that

$$\lim_{n \rightarrow \infty} \frac{\mathcal{T}_n}{n} = \frac{1}{2} \text{tr}((I_2 \otimes A)^{-1/2} L (I_2 \otimes A)^{-1/2}) \quad \text{a.s.}$$

which allows us to say that  $\mathcal{R}_n = o(n)$  a.s. leading to (9.2) We are now in position to prove (9.3). Let us recall that

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

Hence,  $(\mathcal{B}_n)_{n \geq 2}$  is a square integrable martingale. In addition, we have

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

Thus

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

We can observe that

$$L_n = \sum_{k \in \mathbb{G}_n} \frac{1}{c_k^2} \begin{pmatrix} \sigma_a^2 X_k + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_k + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}$$

and

$$\Psi_n \Psi_n^t = \sum_{k \in \mathbb{G}_n} \frac{1}{c_k} I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

For  $\alpha = \max(\sigma_a^2 + \sigma_b^2, \sigma_c^2 + \sigma_d^2)$ , denote

$$\Delta_n = \alpha c_n I_2 - \begin{pmatrix} \sigma_a^2 X_n + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_n + \sigma_d^2 \end{pmatrix}.$$

It is not hard to see that  $\Delta_n$  is a positive definite matrix. As a matter of fact, we deduce from the elementary inequality

$$(\sigma_a^2 + \sigma_b^2) X_n + \sigma_c^2 + \sigma_d^2 \leq \alpha c_n \tag{9.8}$$

that

$$\text{tr}(\Delta_n) = 2\alpha c_n - ((\sigma_a^2 + \sigma_b^2)X_n + \sigma_c^2 + \sigma_d^2) \geq \alpha c_n > 0.$$

In addition, we also have from (9.8) that

$$\begin{aligned} \det(\Delta_n) &= (\alpha c_n - (\sigma_a^2 X_n + \sigma_c^2)) (\alpha c_n - (\sigma_b^2 X_n + \sigma_d^2)) - \rho^2, \\ &= \alpha^2 c_n^2 - \alpha c_n ((\sigma_a^2 + \sigma_b^2)X_n + \sigma_c^2 + \sigma_d^2) \\ &\quad + (\sigma_a^2 X_n + \sigma_c^2)(\sigma_b^2 X_n + \sigma_d^2) - \rho^2, \\ &\geq \sigma_a^2 \sigma_b^2 X_n^2 + (\sigma_a^2 \sigma_d^2 + \sigma_b^2 \sigma_c^2)X_n + \sigma_c^2 \sigma_d^2 - \rho^2, \\ &\geq \sigma_c^2 \sigma_d^2 - \rho^2 > 0 \end{aligned}$$

thanks to (H.3). Consequently,

$$\begin{pmatrix} \sigma_a^2 X_n + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_n + \sigma_d^2 \end{pmatrix} \leq \alpha c_n I_2$$

which immediately implies that  $L_n \leq \alpha \Psi_n \Psi_n^t$ . Moreover, we can use Lemma B.1 of [Bercu et al. \(2009\)](#) to say that

$$\Sigma_n^{-1} \Psi_n \Psi_n^t \Sigma_n^{-1} \leq \Sigma_{n-1}^{-1} - \Sigma_n^{-1}.$$

Hence

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

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  $\mathcal{B}_n = o(\mathcal{A}_n)$ . Finally, we deduce from decomposition (9.1) that

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

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  $(\varepsilon_n)$  satisfies (H.1) to (H.5). For all  $\delta > 1/2$ , we have*

$$\|M_n\|^2 = o(|\mathbb{T}_n|n^\delta) \quad \text{a.s.} \tag{9.9}$$

*Proof* Let us recall that

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

Denote

$$P_n = \sum_{k \in \mathbb{T}_{n-1}} \frac{X_k V_{2k}}{c_k} \quad \text{and} \quad Q_n = \sum_{i \in \mathbb{T}_{n-1}} \frac{V_{2k}}{c_k}.$$



On the one hand,  $P_n$  can be rewritten as

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

We already saw in Sect. 3 that for all  $k \in \mathbb{G}_n$ ,

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

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

$$\mathbb{E}[V_{2k}^4 | \mathcal{F}_n] = 3\sigma_a^4 X_k^2 + X_k(\mu_a^4 - 3\sigma_a^4 + 6\sigma_a^2 \sigma_c^2) + \mu_c^4 \quad \text{a.s.}$$

which implies that

$$\mathbb{E}[V_{2k}^4 | \mathcal{F}_n] \leq \mu_{ac}^4 c_k^2 \quad \text{a.s.} \tag{9.10}$$

where  $\mu_{ac}^4 = \mu_a^4 + \mu_c^4 + 6\sigma_a^2 \sigma_c^2$ . Consequently,  $\mathbb{E}[f_{n+1} | \mathcal{F}_n] = 0$  a.s. and we deduce from (9.10) together with the Cauchy–Schwarz inequality that

$$\begin{aligned} \mathbb{E}[f_{n+1}^4 | \mathcal{F}_n] &= \frac{1}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} \left( \frac{X_k}{c_k} \right)^4 \mathbb{E}[V_{2k}^4 | \mathcal{F}_n] \\ &\quad + \frac{3}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} \sum_{\substack{l \in \mathbb{G}_n \\ l \neq k}} \left( \frac{X_k}{c_k} \right)^2 \left( \frac{X_l}{c_l} \right)^2 \mathbb{E}[V_{2k}^2 | \mathcal{F}_n] \mathbb{E}[V_{2l}^2 | \mathcal{F}_n] \quad \text{a.s.} \\ &\leq \frac{\mu_{ac}^4}{|\mathbb{G}_n|^2} (1 + 3\sqrt{|\mathbb{G}_n|(|\mathbb{G}_n| - 1)}) \sum_{k \in \mathbb{G}_n} c_k^2 \quad \text{a.s.} \\ &\leq \frac{3\mu_{ac}^4}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} c_k^2 \quad \text{a.s.} \end{aligned} \tag{9.11}$$

However, it follows from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} c_k^2 = \mathbb{E}[(1 + T)^2] \quad \text{a.s.}$$

which is equivalent to say that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} c_k^2 = \mathbb{E}[(1 + T)^2] \quad \text{a.s.} \tag{9.12}$$

Therefore, we infer from (9.11) and (9.12) that

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

Hence, we obtain from Wei’s Lemma given in Wei (1987) page 1672 that for all  $\delta > 1/2$ ,

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

On the other hand,  $Q_n$  can be rewritten as

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

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

$$\mathbb{E}[g_{n+1}^4|\mathcal{F}_n] \leq \frac{3\mu_{bd}^4}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} \frac{1}{c_k^2} \leq 3\mu_{bd}^4 \quad \text{a.s.}$$

Hence, we deduce once again from Wei's Lemma that for all  $\delta > 1/2$ ,

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

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.  $\square$

*Proof of Theorem 5.4* We recall from (4.1) that  $\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1}M_n$  which implies

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

where  $\mathcal{V}_n = M_n^t \Sigma_{n-1}^{-1} M_n$ . 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 (9.7) that

$$\lim_{n \rightarrow \infty} \frac{\lambda_{\min}(\Sigma_n)}{|\mathbb{T}_n|} = \lambda_{\min}(A) > 0 \quad \text{a.s.}$$

Consequently, we find that

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

We are now in position to prove the quadratic strong law (5.5). 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

$$\lim_{n \rightarrow \infty} \frac{\mathcal{A}_n}{n} = \frac{1}{2} \text{tr}((I_2 \otimes A)^{-1/2} L(I_2 \otimes A)^{-1/2}) \quad \text{a.s.} \quad (9.13)$$

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

$$\mathcal{A}_n = \sum_{k=1}^n M_k^t \left( \Sigma_{k-1}^{-1} - \Sigma_k^{-1} \right) M_k = \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1/2} \Delta_k \Sigma_{k-1}^{-1/2} M_k$$

where  $\Delta_k = I_4 - \Sigma_{k-1}^{1/2} \Sigma_k^{-1} \Sigma_{k-1}^{1/2}$ . We already saw from (9.7) that

$$\lim_{n \rightarrow \infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes A \quad \text{a.s.}$$

which ensures that

$$\lim_{n \rightarrow \infty} \Delta_n = \frac{1}{2} I_4 \quad \text{a.s.}$$

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

$$\frac{\mathcal{A}_n}{n} = \left( \frac{1}{2n} \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1} M_k \right) + o(1) \quad \text{a.s.} \quad (9.14)$$

Moreover we have

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1} M_k &= \frac{1}{n} \sum_{k=1}^n (\hat{\theta}_k - \theta)^t \Sigma_{k-1} (\hat{\theta}_k - \theta), \\ &= \frac{1}{n} \sum_{k=1}^n |\mathbb{T}_{k-1}| (\hat{\theta}_k - \theta)^t \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} (\hat{\theta}_k - \theta), \\ &= \frac{1}{n} \sum_{k=1}^n |\mathbb{T}_{k-1}| (\hat{\theta}_k - \theta)^t (I_2 \otimes A) (\hat{\theta}_k - \theta) + o(1) \quad \text{a.s.} \end{aligned} \tag{9.15}$$

Therefore, (9.13) together with (9.14) and (9.15) lead to (5.5). □

### 10 Proof of Theorem 5.5

First of all, we shall only prove (5.7) since the proof of (5.8) follows exactly the same lines. We clearly have from (3.7) that

$$\begin{aligned} Q_{n-1}(\hat{\eta}_n - \eta_n) &= \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} (\hat{V}_{2k}^2 - V_{2k}^2) \Phi_k, \\ &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{1}{d_k} (\hat{V}_{2k}^2 - V_{2k}^2) \Phi_k, \\ &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{1}{d_k} ((\hat{V}_{2k} - V_{2k})^2 + 2(\hat{V}_{2k} - V_{2k})V_{2k}) \Phi_k. \end{aligned} \tag{10.1}$$

In addition, we already saw in Sect. 3 that for all  $l \geq 0$  and  $k \in \mathbb{G}_l$ ,

$$\hat{V}_{2k} - V_{2k} = - \begin{pmatrix} \hat{a}_l - a \\ \hat{c}_l - c \end{pmatrix}^t \Phi_k.$$

Consequently,

$$(\hat{V}_{2k} - V_{2k})^2 \leq \|\Phi_k\|^2 ((\hat{a}_l - a)^2 + (\hat{c}_l - c)^2).$$

Hence, we obtain that

$$\begin{aligned} \left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{(\hat{V}_{2k} - V_{2k})^2}{d_k} \Phi_k \right\| &\leq \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{\|\Phi_k\|^3}{d_k} ((\hat{a}_l - a)^2 + (\hat{c}_l - c)^2), \\ &\leq \sum_{l=0}^{n-1} ((\hat{a}_l - a)^2 + (\hat{c}_l - c)^2) \sum_{k \in \mathbb{G}_l} c_k, \\ &\leq \sum_{l=0}^{n-1} ((\hat{a}_l - a)^2 + (\hat{c}_l - c)^2) |\mathbb{T}_{l-1}| \frac{1}{|\mathbb{T}_{l-1}|} \sum_{k \in \mathbb{G}_l} c_k. \end{aligned} \tag{10.2}$$

Moreover, we can deduce from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{G}_n} c_k = \mathbb{E}[1 + T] \quad \text{a.s.} \tag{10.3}$$

Then, we find from (10.2) and (10.3) that

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k} \Phi_k \right\| = \mathcal{O} \left( \sum_{l=0}^{n-1} |\mathbb{T}_{l-1}| ((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2) \right) \quad \text{a.s.}$$

However, as  $\Lambda$  is positive definite, we obtain from (5.5) that

$$\sum_{l=0}^{n-1} |\mathbb{T}_{l-1}| ((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2) = \mathcal{O}(n) \quad \text{a.s.}$$

which implies that

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k} \Phi_k \right\| = \mathcal{O}(n) \quad \text{a.s.} \tag{10.4}$$

Furthermore, denote

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

We clearly have

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

In addition, for all  $k \in \mathbb{G}_n$ ,  $\mathbb{E}[V_{2k} | \mathcal{F}_n] = 0$  a.s. and  $\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = \sigma_a^2 X_k + \sigma_c^2 \leq \alpha c_k$  a.s. where  $\alpha = \max(\sigma_a^2, \sigma_c^2)$ . Consequently,  $\mathbb{E}[\Delta P_{n+1} | \mathcal{F}_n] = 0$  a.s. and

$$\begin{aligned} \mathbb{E}[\Delta P_{n+1} \Delta P_{n+1}^t | \mathcal{F}_n] &= \sum_{k \in \mathbb{G}_n} \frac{1}{d_k^2} \mathbb{E}[V_{2k}^2 | \mathcal{F}_n] \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \quad \text{a.s.} \\ &= \sum_{k \in \mathbb{G}_n} \frac{\sigma_a^2 X_k + \sigma_c^2}{d_k^2} \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \quad \text{a.s.} \end{aligned}$$

Therefore,  $(P_n)$  is a square integrable vector martingale with increasing process  $\langle P \rangle_n$  given by

$$\begin{aligned} \langle P \rangle_n &= \sum_{l=1}^{n-1} \mathbb{E}[\Delta P_{l+1} \Delta P_{l+1}^t | \mathcal{F}_l] \quad \text{a.s.} \\ &= \sum_{l=1}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{\sigma_a^2 X_k + \sigma_c^2}{d_k^2} \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \quad \text{a.s.} \end{aligned}$$

It immediately follows from the previous calculation that

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

leading to

$$\|\langle P \rangle_n\| = \mathcal{O}(n) \quad \text{a.s.}$$

Then, we deduce from the strong law of large numbers for martingale given e.g. in Theorem 1.3.15 of [Duflo \(1997\)](#) that

$$\|P_n\| = o(n) \quad \text{a.s.} \tag{10.5}$$

Hence, we find from (10.1), (10.4) and (10.5) that

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

Moreover, we infer once again from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_n|} Q_n = \mathbb{E} \left[ \begin{pmatrix} \frac{T^2}{(1+T)^2} & \frac{T}{(1+T)^2} \\ \frac{T}{(1+T)^2} & \frac{1}{(1+T)^2} \end{pmatrix} \right] \quad \text{a.s.} \tag{10.6}$$

which ensures that

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

It remains to establish (5.9). Denote

$$\widehat{W}_n = \begin{pmatrix} \widehat{V}_{2n} \\ \widehat{V}_{2n+1} \end{pmatrix} \quad \text{and} \quad R_n = \sum_{k \in \mathbb{T}_{n-1}} (\widehat{W}_k - W_k)' J W_k$$

where

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then, we have

$$|\mathbb{T}_{n-1}|(\widehat{\rho}_n - \rho_n) = \sum_{k \in \mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k}) (\widehat{V}_{2k+1} - V_{2k+1}) + R_n.$$

It is not hard to see that  $(R_n)$  is a square integrable real martingale with increasing process given by

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

$$\begin{aligned}
 &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{W}_k - W_k)^t J \begin{pmatrix} \sigma_a^2 X_k + \sigma_c^2 & \\ \rho & \sigma_b^2 X_k + \sigma_d^2 \end{pmatrix} J (\widehat{W}_k - W_k) \quad \text{a.s.} \\
 &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{W}_k - W_k)^t \begin{pmatrix} \sigma_b^2 X_k + \sigma_d^2 & \\ \rho & \sigma_a^2 X_k + \sigma_c^2 \end{pmatrix} (\widehat{W}_k - W_k) \quad \text{a.s.}
 \end{aligned}$$

Consequently,

$$\begin{aligned}
 \langle R \rangle_n &\leq \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} ((\sigma_a^2 + \sigma_b^2) X_k + \sigma_c^2 + \sigma_d^2) \|\widehat{W}_k - W_k\|^2 \quad \text{a.s.} \\
 &\leq 2\beta \sum_{l=0}^{n-1} ((\widehat{a}_l - a)^2 + (\widehat{b}_l - b)^2) \sum_{k \in \mathbb{G}_l} X_k^2 c_k \\
 &\quad + 2\beta \sum_{l=0}^{n-1} ((\widehat{c}_l - c)^2 + (\widehat{d}_l - d)^2) \sum_{k \in \mathbb{G}_l} c_k \quad \text{a.s.}
 \end{aligned}$$

where  $\beta = \max(\sigma_a^2 + \sigma_b^2, \sigma_c^2 + \sigma_d^2)$ . As previously, we obtain through Lemma 5.2 together with (5.5) that  $\langle R \rangle_n = \mathcal{O}(n)$  a.s. which ensures that  $R_n = o(n)$  a.s. Moreover,

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

which implies via Lemma 5.2 and (5.5) that

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

Therefore, we obtain that

$$|\mathbb{T}_{n-1}| (\widehat{\rho}_n - \rho_n) = \mathcal{O}(n) \quad \text{a.s.}$$

which leads to (5.9). Finally, it only remains to prove the a.s. convergence of  $\eta_n, \zeta_n$  and  $\rho_n$  to  $\eta, \zeta$  and  $\rho$  which will immediately lead to the a.s. convergence of  $\widehat{\eta}_n, \widehat{\zeta}_n$  and  $\widehat{\rho}_n$  through (5.7), (5.8) and (5.9), respectively. On the one hand,

$$Q_{n-1}(\eta_n - \eta) = N_n = \sum_{k \in \mathbb{T}_n} \frac{1}{d_k} \Phi_k v_{2k} \tag{10.7}$$

where we recall that  $v_{2n} = V_{2n}^2 - \eta^t \Phi_n$ . It is clear that  $(N_n)$  is a square integrable vector martingale with increasing process  $\langle N \rangle_n$  given by

$$\langle N \rangle_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{1}{d_k^2} \Phi_k \Phi_k^t (2\sigma_a^4 X_k^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2) X_k + \mu_c^4 - \sigma_c^4) \quad \text{a.s.}$$

Hence,

$$\langle N \rangle_n \leq \gamma \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{1}{d_k} \Phi_k \Phi_k^l \quad \text{a.s.}$$

where  $\gamma = \mu_a^4 - \sigma_a^4 + 4\sigma_a^2\sigma_c^2 + \mu_c^4 - \sigma_c^4$ , which implies that

$$\|\langle N \rangle_n\| = \mathcal{O}(|\mathbb{T}_{n-1}|) \quad \text{a.s.}$$

Consequently,

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

which leads via (10.6) and (10.7) to the a.s. convergence of  $\eta_n$  to  $\eta$  and to the rate of convergence of Remark 5.6. The proof of the a.s. convergence of  $\zeta_n$  to  $\zeta$  follows exactly the same lines. On the other hand

$$|\mathbb{T}_{n-1}|(\rho_n - \rho) = H_n = \sum_{k \in \mathbb{T}_{n-1}} (V_{2k}V_{2k+1} - \rho) \tag{10.8}$$

It is obvious to see that  $(H_n)$  is a square integrable real martingale with increasing process  $\langle H \rangle_n$  such that  $\langle H \rangle_n = \mathcal{O}(|\mathbb{T}_{n-1}|)$  a.s. Finally, as  $H_n^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|)$  a.s., we deduce from (10.8) that  $\rho_n$  goes a.s. to  $\rho$  and that the rate of convergence of Remark 5.6 is verified, which completes the proof of Theorem 5.5.

### 11 Proof of Theorem 5.7

In order to establish the asymptotic normality of our estimators, we will extensively make use of the central limit theorem for triangular arrays of vector martingales given e.g. by Theorem 2.1.9 of Duffo (1997). First of all, instead of using the generation-wise filtration  $(\mathcal{F}_n)$ , we will use the sister pair-wise filtration  $(\mathcal{G}_n)$  given by

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

*Proof of Theorem 5.7, first part* We focus our attention to the proof of the asymptotic normality (5.10). Let  $M^{(n)} = (M_k^{(n)})$  be the square integrable vector martingale defined as

$$M_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^k D_i \tag{11.1}$$

where

$$D_i = \frac{1}{c_i} \begin{pmatrix} X_i V_{2i} \\ V_{2i} \\ X_i V_{2i+1} \\ V_{2i+1} \end{pmatrix}.$$

We clearly have

$$M_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{t_n} D_i = \frac{1}{\sqrt{|\mathbb{T}_n|}} M_{n+1} \tag{11.2}$$

where  $t_n = |\mathbb{T}_n|$ . Moreover, the increasing process associated to  $(M_k^{(n)})$  is given by

$$\begin{aligned} \langle M^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^k \mathbb{E} [D_i D_i^t | \mathcal{G}_{i-1}], \\ &= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^k \frac{1}{c_i^2} \begin{pmatrix} \sigma_a^2 X_i + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_i + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1 \end{pmatrix} \quad \text{a.s.} \end{aligned}$$

Consequently, it follows from convergence (5.3) that

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

It is now necessary to verify Lindeberg's condition by use of Lyapunov's condition. Denote

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

We obtain from (11.1) that

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

In addition, we already saw in Section 9 that

$$\mathbb{E}[V_{2n}^4 | \mathcal{G}_{n-1}] \leq \mu_{ac}^4 c_n^2, \quad \mathbb{E}[V_{2n+1}^4 | \mathcal{G}_{n-1}] \leq \mu_{bd}^4 c_n^2 \quad \text{a.s.}$$

where  $\mu_{ac}^4 = \mu_a^4 + \mu_c^4 + 6\sigma_a^2 \sigma_c^2$  and  $\mu_{bd}^4 = \mu_b^4 + \mu_d^4 + 6\sigma_b^2 \sigma_d^2$ . Hence,

$$\phi_n \leq \frac{2\mu^4}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} c_k^2 \quad \text{a.s.}$$

where  $\mu^4 = \mu_{ac}^4 + \mu_{bd}^4$ . We can deduce from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} c_k^2 = \mathbb{E}[(1 + T)^2] \quad \text{a.s.}$$

which immediately implies that

$$\lim_{n \rightarrow \infty} \phi_n = 0 \quad \text{a.s.}$$

Therefore, Lyapunov's condition is satisfied and Theorem 2.1.9 of Duflo (1997) allows us to say via (11.2) that

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

Finally, we infer from (4.1) together with (9.7) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|} (\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (I_2 \otimes A^{-1})L(I_2 \otimes A^{-1})).$$

□



*Proof of Theorem 5.7, second part* We shall now establish the asymptotic normality given by (5.11). Denote by  $N^{(n)} = (N_k^{(n)})$  the square integrable vector martingale defined as

$$N_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^k \frac{v_{2i}}{d_i} \Phi_i.$$

We immediately see from (10.7) that

$$N_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} Q_n(\eta_{n+1} - \eta) = \frac{1}{\sqrt{|\mathbb{T}_n|}} N_{n+1}. \tag{11.3}$$

In addition, the increasing process associated to  $(N_k^{(n)})$  is given by

$$\begin{aligned} \langle N^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^k \mathbb{E} \left[ \frac{v_{2i}^2}{d_i^2} \Phi_i \Phi_i^t \middle| \mathcal{G}_{i-1} \right], \\ &= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^k \frac{1}{d_i^2} \Phi_i \Phi_i^t (2\sigma_a^4 X_i^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2) X_i + \mu_c^4 - \sigma_c^4) \quad \text{a.s.} \end{aligned}$$

Consequently, we obtain from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \langle N^{(n)} \rangle_{t_n} = \mathbb{E} \left[ \frac{2\sigma_a^4 T^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2) T + (\mu_c^4 - \sigma_c^4)}{(1 + T)^4} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right] = M_{ac} \quad \text{a.s.}$$

In order to verify Lyapunov’s condition, let

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E} \left[ \|N_k^{(n)} - N_{k-1}^{(n)}\|^3 \middle| \mathcal{G}_{k-1} \right].$$

We clearly have

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^2 = \frac{1}{|\mathbb{T}_n|} \frac{(1 + X_k^2) v_{2k}^2}{d_k^2} \leq \frac{1}{|\mathbb{T}_n|} \frac{v_{2k}^2}{d_k},$$

which implies that

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^3 \leq \frac{1}{|\mathbb{T}_n|^{3/2}} \frac{|v_{2k}|^3}{d_k^{3/2}}.$$

However,

$$\begin{aligned} |v_{2k}|^3 &= |V_{2k}^2 - \sigma_a^2 X_k - \sigma_c^2|^3 \leq (V_{2k}^2 + \sigma_a^2 X_k + \sigma_c^2)^3 \\ &\leq V_{2k}^6 + 3V_{2k}^4 (\sigma_a^2 X_k + \sigma_c^2) + 3V_{2k}^2 (\sigma_a^2 X_k + \sigma_c^2)^2 + (\sigma_a^2 X_k + \sigma_c^2)^3. \end{aligned} \tag{11.4}$$

We already saw that  $\mathbb{E}[V_{2k}^2 | \mathcal{G}_{k-1}] = \sigma_a^2 X_k + \sigma_c^2$  a.s. and it follows from (9.10) that

$$\mathbb{E}[V_{2k}^4 | \mathcal{G}_{k-1}] \leq \mu_{ac} c_k^2 \quad \text{a.s.}$$

It only remains to study  $\mathbb{E}[V_{2k}^6 | \mathcal{G}_{k-1}]$ . Denote

$$A_k = \sum_{i=1}^{X_k} (Y_{k,i} - a) \quad \text{and} \quad B_k = \varepsilon_{2k} - c.$$

We clearly have from the identity  $V_{2k} = A_k + B_k$  that

$$\begin{aligned} \mathbb{E}[V_{2k}^6 | \mathcal{G}_{k-1}] &= \mathbb{E}[A_k^6 | \mathcal{G}_{k-1}] + 15\mathbb{E}[A_k^4 | \mathcal{G}_{k-1}]\mathbb{E}[B_k^2 | \mathcal{G}_{k-1}] \\ &\quad + 20\mathbb{E}[A_k^3 | \mathcal{G}_{k-1}]\mathbb{E}[B_k^3 | \mathcal{G}_{k-1}] + \mathbb{E}[A_k^2 | \mathcal{G}_{k-1}]\mathbb{E}[B_k^4 | \mathcal{G}_{k-1}] \\ &\quad + \mathbb{E}[B_k^6 | \mathcal{G}_{k-1}]. \end{aligned} \tag{11.5}$$

On the one hand,  $\mathbb{E}[A_k^2 | \mathcal{G}_{k-1}] = \sigma_a^2 X_k$  a.s. and

$$\mathbb{E}[A_k^4 | \mathcal{G}_{k-1}] = \mu_a^4 X_k + 3X_k(X_k - 1)\sigma_a^4 \quad \text{a.s.}$$

Moreover, we have from Cauchy–Schwarz inequality that

$$|\mathbb{E}[A_k^3 | \mathcal{G}_{k-1}]| \leq 2\mu_a^2 \sigma_a X_k^{3/2} \quad \text{a.s.}$$

Furthermore, it follows from tedious but straightforward calculations that

$$\begin{aligned} \mathbb{E}[A_k^6 | \mathcal{G}_{k-1}] &\leq \tau_a^6 X_k + 15X_k(X_k - 1)\mu_a^4 \sigma_a^2 + 15\sigma_a^6 X_k(X_k - 1)(X_k - 2) \\ &\quad + 10\mu_a^6 X_k(X_k - 1) \quad \text{a.s.} \end{aligned}$$

Then, it exists some constant  $\alpha > 0$  such that

$$\mathbb{E}[A_k^6 | \mathcal{G}_{k-1}] \leq \alpha c_k^3 \quad \text{a.s.}$$

On the other hand,  $\mathbb{E}[B_k^2 | \mathcal{G}_{k-1}] = \sigma_c^2$  a.s. and  $\mathbb{E}[B_k^4 | \mathcal{G}_{k-1}] = \mu_c^4$  a.s. In addition

$$|\mathbb{E}[B_k^3 | \mathcal{G}_{k-1}]| \leq \mu_c^2 \sigma_c \quad \text{and} \quad \mathbb{E}[B_k^6 | \mathcal{G}_{k-1}] \leq \tau_c^6 \quad \text{a.s.}$$

Consequently, we deduce from (11.5) that it exists some constant  $\beta > 0$  such that

$$\mathbb{E}[V_{2k}^6 | \mathcal{G}_{k-1}] \leq \beta c_k^3 \quad \text{a.s.}$$

which implies from (11.4) that for some constant  $\gamma > 0$ ,

$$\mathbb{E}[|v_{2k}|^3 | \mathcal{G}_{k-1}] \leq \gamma c_k^3 \quad \text{a.s.}$$

Then, as  $c_k^2 = d_k$ , we can conclude that

$$\phi_n \leq \frac{\gamma}{\sqrt{|\mathbb{T}_n|}} \quad \text{a.s.}$$

which immediately leads to

$$\lim_{n \rightarrow \infty} \phi_n = 0 \quad \text{a.s.}$$

Therefore, Lyapunov’s condition is satisfied and we find from Theorem 2.1.9 of [Duflo \(1997\)](#) and (11.3) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} N_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{ac}). \tag{11.6}$$

Hence, we obtain from (10.6), (11.6) and Slutsky’s lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\eta_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1}M_{ac}B^{-1}).$$

Finally, (5.7) ensures that

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

The proof of (5.12) follows exactly the same lines. □

*Proof of Theorem 5.7, third part* It remains to establish the asymptotic normality given by (5.13). Denote by  $H^{(n)} = (H_k^{(n)})$  the square integrable martingale defined as

$$H_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^k (V_{2i} V_{2i+1} - \rho). \tag{11.7}$$

We clearly have from (10.8) that

$$H_{t_n}^{(n)} = \sqrt{|\mathbb{T}_n|}(\rho_{n+1} - \rho) = \frac{1}{\sqrt{|\mathbb{T}_n|}} H_{n+1}.$$

Moreover, the increasing process of  $(H_k^{(n)})$  is given by

$$\langle H^{(n)} \rangle_k = \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^k (\mathbb{E}[V_{2i}^2 V_{2i+1}^2 | \mathcal{G}_{n-1}] - \rho^2).$$

As before, let

$$C_k = \sum_{i=1}^{X_k} (Z_{k,i} - b) \quad \text{and} \quad B_k = \varepsilon_{2k+1} - d.$$

As  $V_{2k} = A_k + B_k$  and  $V_{2k+1} = C_k + D_k$ , we clearly have

$$\begin{aligned} \mathbb{E}[V_{2k}^2 V_{2k+1}^2 | \mathcal{G}_{k-1}] &= \mathbb{E}[A_k^2 | \mathcal{G}_{k-1}] (\mathbb{E}[C_k^2 | \mathcal{G}_{k-1}] + \mathbb{E}[D_k^2 | \mathcal{G}_{k-1}]) \\ &\quad + \mathbb{E}[B_k^2 | \mathcal{G}_{k-1}] \mathbb{E}[C_k^2 | \mathcal{G}_{k-1}] + \mathbb{E}[B_k^2 D_k^2 | \mathcal{G}_{k-1}] \quad \text{a.s.} \end{aligned}$$

Consequently,

$$\mathbb{E}[V_{2k}^2 V_{2k+1}^2 | \mathcal{G}_{k-1}] = \sigma_a^2 \sigma_b^2 X_k^2 + (\sigma_a^2 \sigma_d^2 + \sigma_b^2 \sigma_c^2) X_k + v^2 \quad \text{a.s.} \tag{11.8}$$

Then, we deduce once again from Lemma 5.2 that

$$\lim_{n \rightarrow \infty} \langle H^{(n)} \rangle_{t_n} = \sigma_\rho^2 \quad \text{a.s.}$$

where  $\sigma_\rho^2$  is given by (5.14). In order to verify Lyapunov's condition, denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}[|H_k^{(n)} - H_{k-1}^{(n)}|^3 | \mathcal{G}_{k-1}].$$

We obtain from (11.7) that

$$\begin{aligned} \phi_n &= \frac{1}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} \mathbb{E}[|V_{2k} V_{2k+1} - \rho|^3 | \mathcal{G}_{k-1}], \\ &\leq \frac{1}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} (\mathbb{E}[|V_{2k}|^3 |V_{2k+1}|^3 | \mathcal{G}_{k-1}] + 3|\rho| \mathbb{E}[V_{2k}^2 V_{2k+1}^2 | \mathcal{G}_{k-1}] \\ &\quad + 3\rho^2 \mathbb{E}[|V_{2k}| |V_{2k+1}| | \mathcal{G}_{k-1}] + |\rho|^3). \end{aligned} \tag{11.9}$$

It follows from Cauchy–Schwarz inequality together with the previous calculations that it exists two constants  $\alpha, \beta > 0$  such that

$$\mathbb{E}[|V_{2k}| |V_{2k+1}| | \mathcal{G}_{k-1}] \leq \alpha c_k \quad \text{a.s.}$$

and

$$\mathbb{E} [ |V_{2k}|^3 |V_{2k+1}|^3 | \mathcal{G}_{k-1} ] \leq \beta c_k^3 \quad \text{a.s.}$$

In addition, we already saw from (11.8) that for some constant  $\gamma > 0$

$$\mathbb{E} [ V_{2k}^2 V_{2k+1}^2 | \mathcal{G}_{k-1} ] \leq \gamma c_k^2 \quad \text{a.s.}$$

Consequently, we obtain from (11.9) that for some constant  $\delta > 0$

$$\phi_n \leq \frac{\delta}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} c_k^3 \quad \text{a.s.}$$

which, via Lemma (5.2), leads to

$$\lim_{n \rightarrow \infty} \phi_n = 0 \quad \text{a.s.}$$

Hence, we can conclude that

$$H_{t_n}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2).$$

In other words

$$\sqrt{|\mathbb{T}_{n-1}|}(\rho_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2).$$

Finally, we find via (5.9) that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2)$$

which achieves the proof of Theorem 5.7. □

**Acknowledgments** The authors are thankful to the anonymous reviewer for his very careful reading of the manuscript.

## References

- Al-Osh MA, Alzaid AA (1987) First-order integer-valued autoregressive (INAR(1)) process. *J Time Ser Anal* 8(3):261–275
- Alzaid AA, Al-Osh M (1990) An integer-valued  $p$ th-order autoregressive structure (INAR( $p$ )) process. *J Appl Probab* 27(2):314–324
- Bansaye V (2008) Proliferating parasites in dividing cells: Kimmel’s branching model revisited. *Ann. Appl. Probab.* 18(3):967–996
- Bansaye V, Tran VC (2011) Branching Feller diffusion for cell division with parasite infection. *ALEA Lat Am J Probab Math Stat* 8:95–127
- Basawa IV, Zhou J (2004) Non-Gaussian bifurcating models and quasi-likelihood estimation. *J Appl Probab* 41A(2004):55–64 *Stochastic methods and their applications*
- Bercu B, Blandin V (2014) A Rademacher–Menchov approach for random coefficient bifurcating autoregressive processes. [arXivmath.PR/1210.5835](https://arxiv.org/abs/1210.5835)
- Bercu B, De Saporta B, Gégout-Petit A (2009) Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron J Probab* 14(87):2492–2526
- Blandin V (2014) Asymptotic results for bifurcating random coefficient autoregressive processes. *Statistics* 108:48
- Cowan R, Staudte RG (1986) The bifurcating autoregressive model in cell lineage studies. *Biometrics* 42:769–783
- De Saporta B, Gégout-Petit A, Marsalle L (2011) Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electron J Stat* 5:1313–1353

- De Saporta B, Gégout-Petit A, Marsalle L (2012) Asymmetry tests for bifurcating auto-regressive processes with missing data. *Stat Probab Lett* 82(7):1439–1444
- De Saporta B, Gégout-Petit A, Marsalle L (2014) Statistical study of asymmetry in cell lineage data. *Comput Stat Data Anal* 69:15–39
- Delmas J-F, Marsalle L (2010) Detection of cellular aging in a Galton–Watson process. *Stoch Process Appl* 120(12):2495–2519
- Dufflo M (1997) *Random iterative models*, vol 34. Springer, Berlin
- Guyon J (2007) Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann Appl Probab* 17(5–6):1538–1569
- Hall P, Heyde CC (1980) *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics
- Heyde CC, Seneta E (1972) Estimation theory for growth and immigration rates in a multiplicative process. *J Appl Probab* 9:235–256
- Huggins RM, Basawa IV (1999) Extensions of the bifurcating autoregressive model for cell lineage studies. *J Appl Probab* 36(4):1225–1233
- Huggins RM, Basawa IV (2000) Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust N Z J Stat* 42(4):423–432
- Huggins RM, Staudte RG (1994) Variance components models for dependent cell populations. *J Am Stat Assoc* 89(425):19–29
- Kachour M, Yao JF (2009) First-order rounded integer-valued autoregressive (RINAR(1)) process. *J Time Ser Anal* 30(4):417–448
- McKenzie E (1985) Some simple models for discrete variate time series. *J Am Water Res Assoc* 21:645–650
- Nicholls DF, Quinn BG (1980) The estimation of random coefficient autoregressive models. I. *J Time Ser Anal* 1(1):37–46
- Nicholls DF, Quinn BG (1982) *Random coefficient autoregressive models: an introduction*, vol. 11 of *Lecture Notes in Statistics*. Springer, New York. *Lecture Notes in Physics*, 151
- Quinn BG, Nicholls DF (1981) The estimation of random coefficient autoregressive models. II. *J Time Ser Anal* 2(3):185–203
- Wei CZ (1987) Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. *Ann Stat* 15(4):1667–1682
- Wei CZ, Winnicki J (1990) Estimation of the means in the branching process with immigration. *Ann Stat* 18(4):1757–1773
- Winnicki J (1991) Estimation of the variances in the branching process with immigration. *Probab Theory Relat Fields* 88(1):77–106
- Zhou J, Basawa IV (2005) Least-squares estimation for bifurcating autoregressive processes. *Stat Probab Lett* 74(1):77–88
- Zhou J, Basawa IV (2005) Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J Time Ser Anal* 26(6):825–842