ESAIM: PS 18 (2014) 365–399
DOI: 10.1051/ps/2013042

ESAIM: Probability and Statistics
www.esaim-ps.org

RANDOM COEFFICIENTS BIFURCATING AUTOREGRESSIVE PROCESSES

Benoîte de Saporta[1], Anne Gégout-Petit[2] and Laurence Marsalle[3]

Abstract. This paper presents a new model of asymmetric bifurcating autoregressive process with random coefficients. We couple this model with a Galton−Watson tree to take into account possibly missing observations. We propose least-squares estimators for the various parameters of the model and prove their consistency, with a convergence rate, and asymptotic normality. We use both the bifurcating Markov chain and martingale approaches and derive new results in both these frameworks.

Mathematics Subject Classification. 60J05, 60J80, 62M05, 62F12, 60G42, 92D25.

Received October 19, 2012. Revised May 3, 2013.

1. Introduction

In the 80's, Cowan and Staudte [7] introduced Bifurcating Autoregressive processes (BAR) as a parametric model to study cell lineage data. A quantitative characteristic of the cells (*e.g.* growth rate, age at division) is recorded over several generations descended from an initial cell, keeping track of the genealogy to study inherited effects. As a cell usually gives birth to two offspring by division, such genealogies are naturally structured as binary trees. BAR processes are thus a generalization of autoregressive processes (AR) to this binary tree structure, by modeling each line of descent as a first order AR process, allowing the environmental effects on sister cells to be correlated. Statistical inference for the parameters of BAR processes has been widely studied, either based on the observation of a single tree growing to infinity [7, 18, 20, 28] or on a large number of small independent trees [19, 21]. See also [22, 23] for processes indexed by general trees.

Various extensions of the original model have been proposed, *e.g.* non gaussian noise sequence [2, 27], higher order AR [20, 27] or moving average AR [19]. Since 2005, evidence of asymmetry in cell division has been established by biologists [25] and an asymmetric BAR model has been introduced by Guyon [13] where the coefficients of the AR processes of sister cells are allowed to be different. This model was further extended to higher order AR [3], to take missing data into account [9–11] and with parasite infection [1].

Keywords and phrases. Autoregressive process, branching process, missing data, least squares estimation, limit theorems, bifurcating Markov chain, martingale.

[1] Univ. Bordeaux, Gretha, UMR 5113, IMB, UMR 5251, 33400 Talence, France
CNRS, Gretha, UMR 5113, IMB, UMR 5251, 33400 Talence, France
INRIA Bordeaux Sud Ouest, team CQFD, 33400 Talence, France

[2] Univ. Bordeaux, IMB, UMR 5251, 33400 Talence, France
CNRS, IMB, UMR 5251, 33400 Talence, France
INRIA Bordeaux Sud Ouest, team CQFD, 33400 Talence, France

[3] Univ. Lille 1, Laboratoire Paul Painlevé, UMR 8524, 59655 Villeneuve d'Ascq, France
CNRS, Laboratoire Paul Painlevé, UMR 8524, 59655 Villeneuve d'Ascq, France. benoite.desaporta@math.u-bordeaux1.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2014

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

To the best of our knowledge, only two papers [4, 6] deal with random coefficient BAR processes. In the former by Bui and Huggins it is explained that random coefficients BAR processes can account for observations that do not fit the usual BAR model. For instance, the extra randomness can model irregularities in nutrient concentrations in the media in which the cells are grown. Other evidence for the need of richer models can be found *e.g.* in [14]. In this paper, we propose a new model for random coefficient BAR processes (R-BAR). It is more general than that of Bui and Huggins, as the random variables are not supposed to be Gaussian, they may not have moments of all order and correlation between all the sources of randomness are allowed. Moreover, we propose an asymmetric model in the continuation of [3,4,9–11,13] in the context of missing data. Indeed, experimental data are often incomplete and it is important to take this phenomenon into account for the inference. As in [9, 11] we model the structure of available data by a Galton−Watson tree, instead of a complete binary tree. Our model is close to that developed in [4], but the assumptions on the noise process are different as we allow correlation between the two sources of randomness but require higher moments because of the missing data and because we do not use a weighted estimator. The main difference is that the model in [4] is fully observed, whereas ours allows for missing observations.

Our approach for the inference of our model is also different from [4,6]. As we cannot use maximum likelihood estimation, we propose modified least squares estimators as in [24]. In [6], inference is based on an asymptotically infinite number of small replicated trees. Here, as in [4], we consider one single tree growing to infinity but our least squares estimator is not weighted. The originality of our approach is that it combines the bifurcating Markov chain and martingale approaches. Bifurcating Markov chains (BMC) were introduced in [13] on complete binary trees and further developed in [11] in the context of missing data on Galton−Watson trees. BAR models can be seen as a special case of BMC. This interpretation allows us to establish the convergence of our estimators. A by-product of our procedure is a new general result for BMC on Galton−Watson trees. Indeed, in [11,13] the driven noise sequence is assumed to have moments of all order. Here, we establish new laws of large numbers for polynomial functions of the BMC where the noise sequence only has moments up to a given order. The strong law of large numbers [12] and the central limit theorem [12,15,16] for martingales have been previously used in the context of BAR processes [2,27,28] and adapted to special cases of martingales on binary trees [3,4,9,10]. In this paper, we establish a general law of large numbers for square-integrable martingales on Galton−Watson binary trees. This result is applied to our R-BAR model to obtain sharp convergence rates and a quadratic strong law for our estimators.

The paper is organized as follows. In Section 2, we give the precise definition of our R-BAR model on a Galton−Watson tree and state our main assumptions. In Section 3, we give modified least squares estimators and state the convergence results we obtained: consistency with convergence rate and asymptotic normality. In Section 4, we recall the BMC framework, prove a new law of large numbers under limited moment conditions and apply it to our R-BAR model to derive the consistency of our estimators. In Section 5 we establish a new general law of large numbers for square-integrable martingales on Galton−Watson trees and use it to derive convergence rates and quadratic strong laws for our estimators. In Section 6 we establish the asymptotic normality by using central limit theorems for martingales. Finally in Section 7 we apply our estimation procedure to the coli data of [25].

2. Model

In the sequel, all random variables are defined on the probability state space $(\Omega, \mathcal{A}, \mathbb{P})$. As in the previous literature, we use the index 1 for the original cell, and the two offspring of cell k are labelled $2k$ and $2k+1$. Consider the first-order asymmetric random coefficients bifurcating autoregressive process (R-BAR) given, for all $k \geq 1$, by

$$\begin{cases} X_{2k} = (b_{2k} + \eta_{2n})X_k + (a_{2k} + \varepsilon_{2k}), \\ X_{2k+1} = (b_{2k+1} + \eta_{2k+1})X_k + (a_{2k+1} + \varepsilon_{2k+1}), \end{cases} \tag{2.1}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

with the following notations: for all $k \geq 1$,

$$\begin{cases} a_{2k} = a, \\ b_{2k} = b, \end{cases} \quad \text{and} \quad \begin{cases} a_{2k+1} = c, \\ b_{2k+1} = d. \end{cases}$$

The initial state X_1 is the characteristic of the original ancestor while the sequence $(\varepsilon_{2k}, \eta_{2k}, \varepsilon_{2k+1}, \eta_{2k+1})_{n\geq 1}$ is the driving noise of the process, and the parameter (a, b, c, d) belongs to $\mathbb{R}^4$. One can see this R-BAR process as a random-coefficient first-order autoregressive process on a binary tree, where each vertex represents an individual or cell, vertex 1 being the original ancestor. For all $n \geq 1$, denote the nth generation by

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

In particular, $\mathbb{G}_0 = \{1\}$ is the initial generation and $\mathbb{G}_1 = \{2, 3\}$ is the first generation of offspring from the original ancestor. Finally, denote by

$$\mathbb{T}_n = \bigcup_{\ell=0}^{n} \mathbb{G}_\ell,$$

the sub-tree of all individuals from the original individual up to the nth generation and $\mathbb{T}$ the complete tree. Note that the cardinality $|\mathbb{G}_n|$ of $\mathbb{G}_n$ is 2^n while that of $\mathbb{T}_n$ is $|\mathbb{T}_n| = 2^{n+1} - 1$. In all the sequel, we shall use the following hypotheses.

(H.1) The sequence $(\varepsilon_{2k}, \eta_{2k}, \varepsilon_{2k+1}, \eta_{2k+1})_{k\geq 1}$ is independent and identically distributed. It is also independent from X_1.

(H.2) The random variables ε_2, η_2, ε_3, η_3 and X_1 have moments of all order up to 4γ, for some $\gamma \geq 1$. The following hypotheses will be used

$$\begin{gathered} \mathbb{E}[\varepsilon_2] = \mathbb{E}[\varepsilon_3] = 0, \mathbb{E}[\varepsilon_2^2] = \mathbb{E}[\varepsilon_3^2] = \sigma_\varepsilon^2 > 0 \quad \text{and} \quad \mathbb{E}[\varepsilon_2\varepsilon_3] = \rho_\varepsilon, \\ \mathbb{E}[\eta_2] = \mathbb{E}[\eta_3] = 0, \mathbb{E}[\eta_2^2] = \mathbb{E}[\eta_3^2] = \sigma_\eta^2 > 0 \quad \text{and} \quad \mathbb{E}[\eta_2\eta_3] = \rho_\eta, \\ \mathbb{E}[\varepsilon_{2+i}\eta_{2+j}] = \rho_{ij}, \text{ for } (i,j) \in \{0,1\}, \text{ and } \rho = \tfrac{1}{2}(\rho_{01} + \rho_{10}). \end{gathered}$$

When dealing with the biological issue of cell lineages, it may happen that a lineage is incomplete. Indeed, cells may die or measurements may be impossible or faulty on some cells. Taking into account such a phenomenon, we introduce the observation process, $(\delta_k)_{k\in\mathbb{T}}$. We use the same framework as in [11], and not the more general one introduced in [9]. Basically, $\delta_k = 1$ if cell k is observed, $\delta_k = 0$ otherwise. We set $\delta_1 = 1$ and define the whole sequence through the following equalities:

$$\delta_{2k} = \delta_k \xi_k^0 \quad \text{and} \quad \delta_{2k+1} = \delta_k \xi_k^1, \tag{2.2}$$

where the sequence $\left(\boldsymbol{\xi}_k = (\xi_k^0, \xi_k^1)\right)_{k\in\mathbb{T}}$ is a sequence of independent identically distributed random vectors of $\{0,1\}^2$ whose common distribution is specified by the following generating function

$$\mathbb{E}\left[s_0^{\xi_1^0} s_1^{\xi_1^1}\right] = (1 - p_0 - p_1 - p_{01}) + p_0 s_0 + p_1 s_1 + p_{01} s_0 s_1.$$

We also suppose that the observation process is independent from the state process (X_n).

(H.3) The sequence $(\boldsymbol{\xi}_k)_{k\in\mathbb{T}}$ is independent from $(\varepsilon_{2k}, \eta_{2k}, \varepsilon_{2k+1}, \eta_{2k+1})_{k\in\mathbb{T}}$ and from X_1.

Notice that the process $(\delta_k)_{k\in\mathbb{T}}$ takes its values in $\{0,1\}$, and that if $k \in \mathbb{T}$ is such that $\delta_k = 0$, then $\delta_{2^n k+i} = 0$, for all $i \in \{0, \ldots, 2^n - 1\}$ and all $n \geq 1$. So to speak, if individual k is not observed, all its descendants are also missing. We now define the sets of observed data

$$\mathbb{G}_n^* = \{k \in \mathbb{G}_n : \delta_k = 1\} \quad \text{and} \quad \mathbb{T}_n^* = \{k \in \mathbb{T}_n : \delta_k = 1\} = \cup_{\ell=0}^n \mathbb{G}_\ell^*.$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Thanks to the i.i.d. property of the sequence $(\boldsymbol{\xi}_k)$, the sequence of cardinalities $(|\mathbb{G}_n^*|)_{n\geq 0}$ is a Galton−Watson (GW) process with reproduction generating function

$$z \mapsto (1-p_0-p_1-p_{01}) + (p_0+p_1)z + p_{01}z^2,$$

and mean $m = 2p_{01} + p_0 + p_1$. The following equalities thus hold (see *e.g.* [17])

$$\mathbb{E}[|\mathbb{G}_n^*|] = m^n \quad \text{and} \quad \mathbb{E}[|\mathbb{T}_n^*|] = \sum_{\ell=0}^{n} \mathbb{E}[|\mathbb{G}_\ell^*|] = \frac{m^{n+1}-1}{m-1}\cdot$$

According to the position of the mean m of the reproduction law with respect to 1, it is well known that the population becomes extinct or not. More precisely, if $m \leq 1$ then we have extinction almost surely, in the sense that $\mathbb{P}(\cup_{n\geq 0}\{|\mathbb{G}_n^*| = 0\}) = 1$. But if $m > 1$, there is a positive probability of survival of the population: $\mathbb{P}(\cap_{n\geq 0}\{|\mathbb{G}_n^*| > 0\}) > 0$. This latter case is called the super-critical case, and we assume that we are in that case.

(H.4) The mean of the reproduction law is greater than 1: $m > 1$.

On the set of non-extinction, the growth of the population is exponential, more precisely there exists some non-negative square-integrable random variable W such that

$$\lim_{n\to\infty} \frac{|\mathbb{G}_n^*|}{m^n} = W \text{ a.s.}, \qquad \text{and} \qquad \{W > 0\} = \cap_{n\geq 0}\{|\mathbb{G}_n^*| > 0\} \text{ a.s.} \tag{2.3}$$

This immediately entails that

$$\lim_{n\to\infty} \frac{|\mathbb{T}_n^*|}{m^n} = W \times \frac{m}{m-1} \quad \text{a.s.} \tag{2.4}$$

We will denote by $\mathcal{E}$ the extinction set $\mathcal{E} = \cup_{n\geq 0}\{|\mathbb{G}_n^*| = 0\}$ and $\overline{\mathcal{E}}$ its complementary set. Note that under assumption **(H.4)**, $\overline{\mathcal{E}}$ has a positive probability: $\mathbb{P}(\overline{\mathcal{E}}) > 0$. We need one more assumption combining the R-BAR and GW processes.

(H.5) There exist $1 \leq \kappa \leq \gamma$ such that

$$\frac{p_0+p_{01}}{m}\mathbb{E}[(b+\eta_2)^{4\kappa}] + \frac{p_1+p_{01}}{m}\mathbb{E}[(d+\eta_3)^{4\kappa}] < 1.$$

This is the analogous of the usual stability assumption for the autoregression expressed by $\max\{|b|,|d|\} < 1$ in the case of deterministic coefficients. It ensures that the values of $|X_k|$ do not tend to infinity. Note that the assumption above is slightly weaker than the one for deterministic coefficients. Indeed, in the fully observed case and when $\eta_2 = \eta_3 = 0$, this equation reduces to $(b^{4\kappa} + d^{4\kappa})/2 < 1$. The special form of this assumption is explained in Section 4.2 and is closely linked to the properties of the R-BAR process as a bifurcating Markov chain.

Finally, denote by $\mathbb{F} = (\mathcal{F}_n)$ the natural filtration of the R-BAR process $(X_k)_{k\in\mathbb{T}}$, which means that $\mathcal{F}_n$ is the σ-algebra generated by all individuals up to the nth generation, $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$. We also introduce the sigma field $\mathcal{O} = \sigma\{\delta_k, k \in \mathbb{T}\}$ generated by the observation process. We shall assume that all the history of the observation process (δ_k) is known at time 0 and use the filtration $\mathbb{F}^{\mathcal{O}} = (\mathcal{F}_n^{\mathcal{O}})$ defined for all n by

$$\mathcal{F}_n^{\mathcal{O}} = \mathcal{O} \vee \sigma\{\delta_k X_k, k \in \mathbb{T}_n\} = \mathcal{O} \vee \sigma\{X_k, k \in \mathbb{T}_n^*\}.$$

Note that $\mathcal{F}_n^{\mathcal{O}}$ is a sub-σ-field of $\mathcal{O} \vee \mathcal{F}_n$.

3. Estimation

We now give some least-squares estimators of our parameters and state our main results on their asymptotic behavior.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

3.1. Estimators

We propose to use the standard least-squares (LS) estimator $\widehat{\boldsymbol{\theta}}_n = (\widehat{a}_n, \widehat{b}_n, \widehat{c}_n, \widehat{d}_n)^t$ of $\boldsymbol{\theta} = (a, b, c, d)^t$ which minimizes the following expression

$$\Delta_n(\boldsymbol{\theta}) = \frac{1}{2} \sum_{k \in \mathbb{T}_{n-1}} \delta_{2k}(X_{2k} - a - bX_k)^2 + \delta_{2k+1}(X_{2k+1} - c - dX_k)^2.$$

Consequently, for all $n \geq 1$ the following equality holds

$$\widehat{\boldsymbol{\theta}}_n \;=\; \boldsymbol{S}_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} \delta_{2k} X_{2k} \\ \delta_{2k} X_k X_{2k} \\ \delta_{2k+1} X_{2k+1} \\ \delta_{2k+1} X_k X_{2k+1} \end{pmatrix}, \text{ with } \quad \boldsymbol{S}_{n-1} = \begin{pmatrix} \boldsymbol{S}_{n-1}^0 & 0 \\ 0 & \boldsymbol{S}_{n-1}^1 \end{pmatrix},$$

and $\boldsymbol{S}_{n-1}^0 = \sum_{k \in \mathbb{T}_{n-1}} \delta_{2k} \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}$, $\boldsymbol{S}_{n-1}^1 = \sum_{k \in \mathbb{T}_{n-1}} \delta_{2k+1} \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}$.

We now turn to the estimation of the parameters of the conditional covariance of $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$. Following [24], we obtain a modified least squares estimator of $\boldsymbol{\sigma} = (\sigma_\varepsilon^2, \rho_{00}, \rho_{11}, \sigma_\eta^2)^t$ by minimizing

$$\Delta_n'(\boldsymbol{\sigma}) = \frac{1}{2} \sum_{\ell=1}^{n-1} \sum_{k \in \mathbb{G}_\ell} (\widehat{\epsilon}_{2k}^2 - \mathbb{E}[\epsilon_{2k}^2 | \mathcal{F}_\ell^{\mathcal{O}}])^2 + (\widehat{\epsilon}_{2k+1}^2 - \mathbb{E}[\epsilon_{2k+1}^2 | \mathcal{F}_\ell^{\mathcal{O}}])^2,$$

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

$$\begin{cases} \epsilon_{2k} \;\; = \delta_{2k}(\varepsilon_{2k} + \eta_{2k} X_k), \\ \epsilon_{2k+1} = \delta_{2k+1}(\varepsilon_{2k+1} + \eta_{2k+1} X_k), \end{cases} \quad \begin{cases} \widehat{\epsilon}_{2k} \;\; = \delta_{2k}(X_{2k} - \widehat{a}_n - \widehat{b}_n X_k), \\ \widehat{\epsilon}_{2k+1} = \delta_{2k}(X_{2k+1} - \widehat{c}_n - \widehat{d}_n X_k). \end{cases}$$

Under assumptions **(H.2)** and **(H.3)**, one obtains the following estimator

$$\widehat{\boldsymbol{\sigma}}_n \;=\; \boldsymbol{U}_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \left(\widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2, 2X_k \widehat{\epsilon}_{2k}^2, 2X_k \widehat{\epsilon}_{2k+1}^2, X_k^2 (\widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2)\right)^t, \tag{3.1}$$

where

$$\boldsymbol{U}_n = \sum_{k \in \mathbb{T}_n} \begin{pmatrix} \delta_{2k} + \delta_{2k+1} & 2\delta_{2k} X_k & 2\delta_{2k+1} X_k & (\delta_{2k} + \delta_{2k+1}) X_k^2 \\ 2\delta_{2k} X_k & 4\delta_{2k} X_k^2 & 0 & 2\delta_{2k} X_k^3 \\ 2\delta_{2k+1} X_k & 0 & 4\delta_{2k+1} X_k^2 & 2\delta_{2k+1} X_k^3 \\ (\delta_{2k} + \delta_{2k+1}) X_k^2 & 2\delta_{2k} X_k^3 & 2\delta_{2k+1} X_k^3 & (\delta_{2k} + \delta_{2k+1}) X_k^4 \end{pmatrix}.$$

Note that if $\sigma_\eta^2 = 0$ the estimator of σ_ε^2 above corresponds to the empirical estimator already used in [9]. Similarly, the least-squares estimator of $\boldsymbol{\rho} = (\rho_\varepsilon, \rho, \rho_\eta)^t$ minimizes

$$\Delta_n''(\boldsymbol{\rho}) = \frac{1}{2} \sum_{\ell=1}^{n-1} \sum_{k \in \mathbb{G}_\ell} (\widehat{\epsilon}_{2k} \widehat{\epsilon}_{2k+1} - \mathbb{E}[\epsilon_{2k} \epsilon_{2k+1} | \mathcal{F}_\ell^{\mathcal{O}}])^2,$$

and one obtains

$$\widehat{\boldsymbol{\rho}}_n = \boldsymbol{V}_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \left(\widehat{\epsilon}_{2k} \widehat{\epsilon}_{2k+1}, 2X_k \widehat{\epsilon}_{2k} \widehat{\epsilon}_{2k+1}, X_k^2 \widehat{\epsilon}_{2k} \widehat{\epsilon}_{2k+1}\right)^t, \tag{3.2}$$

where

$$\boldsymbol{V}_n = \sum_{k \in \mathbb{T}_n} \delta_{2k} \delta_{2k+1} \begin{pmatrix} 1 & 2X_k & X_k^2 \\ 2X_k & 4X_k^2 & 2X_k^3 \\ X_k^2 & 2X_k^3 & X_k^4 \end{pmatrix}.$$

Note that one cannot identify ρ_{01} from ρ_{10}, hence the use of $\rho = (\rho_{01} + \rho_{10})/2$. Again if $\sigma_\eta^2 = 0$, we retrieve the empirical estimator of ρ_ε used in [9].

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

3.2. Main results

We now state our main results. The first one establishes the consistency of our estimators on the non-extinction set.

Theorem 3.1. *Under assumptions* **(H.1−5)**, *and if* $\kappa \geq 2$, *the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \widehat{\boldsymbol{\theta}}_n = \boldsymbol{\theta} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

and if in addition $\kappa \geq 4$ *then the following convergences also hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \widehat{\boldsymbol{\sigma}}_n = \boldsymbol{\sigma} \mathbb{1}_{\overline{\mathcal{E}}} \ a.s., \qquad \lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \widehat{\boldsymbol{\rho}}_n = \boldsymbol{\rho} \mathbb{1}_{\overline{\mathcal{E}}} \ a.s.$$

The next results give convergence rates for the estimators.

Theorem 3.2. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 4$, *for all* $\delta > 1/2$, *the following convergence rate holds*

$$\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\|^2 = o(n^\delta m^{-n}) \quad a.s.$$

with the quadratic strong law

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{\ell=1}^{n} |\mathbb{T}_{\ell-1}^*|^{-1} (\widehat{\boldsymbol{\theta}}_\ell - \boldsymbol{\theta})^t \boldsymbol{S} \boldsymbol{\Sigma}^{-1} \boldsymbol{S} (\widehat{\boldsymbol{\theta}}_\ell - \boldsymbol{\theta}) = \mathrm{tr}(\boldsymbol{\Gamma} \boldsymbol{\Sigma}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

where $\boldsymbol{S}$, $\boldsymbol{\Gamma}$ *and* $\boldsymbol{\Sigma}$ *are* 4×4 *matrices defined respectively in Proposition 4.14, Lemmas 5.4 and 5.5.*

For all n, set

$$\boldsymbol{\sigma}_n = \boldsymbol{U}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \begin{pmatrix} \epsilon_{2k}^2 + \epsilon_{2k+1}^2 \\ 2X_k\epsilon_{2k}^2 \\ 2X_k\epsilon_{2k+1}^2 \\ X_k^2(\epsilon_{2k}^2 + \epsilon_{2k+1}^2) \end{pmatrix}, \ \boldsymbol{\rho}_n = \boldsymbol{V}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \begin{pmatrix} \epsilon_{2k}\epsilon_{2k+1} \\ 2X_k\epsilon_{2k}\epsilon_{2k+1} \\ X_k^2\epsilon_{2k}\epsilon_{2k+1} \end{pmatrix}.$$

Theorem 3.3. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 8$, *the following convergences hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{\sigma}_n = \boldsymbol{\sigma} \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

and

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{|\mathbb{T}_{n-1}^*|}{n} (\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n) = \boldsymbol{U}^{-1} \left(q_0(0) + q_1(0),\ 2q_0(1),\ 2q_1(1),\ q_0(2) + q_1(2)\right)^t \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

where $\boldsymbol{U}$ *is a* 4×4 *matrix defined in Proposition 4.14 and the* $q_i(r)$ *are scalars defined in Lemmas 5.10, 5.11, 5.12 and 5.13.*

Theorem 3.4. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 8$, *the following convergences hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{\rho}_n = \boldsymbol{\rho} \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

and

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{|\mathbb{T}_{n-1}^*|}{n} (\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n) = \boldsymbol{V}^{-1} \big(q_{01}(0), 2q_{01}(1), q_{01}(2)\big)^t \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

where $\boldsymbol{V}$ *is a* 3×3 *matrix defined in Proposition 4.14 and the* $q_{01}(r)$ *are scalars defined in Lemmas 5.18, 5.19 and 5.20.*

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

We now turn to the asymptotic normality for all our estimators $\widehat{\boldsymbol{\theta}}_n$, $\widehat{\boldsymbol{\sigma}}_n$ and $\widehat{\boldsymbol{\rho}}_n$ given the non-extinction of the underlying Galton−Watson process. For this, using the fact that $\mathbb{P}(\overline{\mathcal{E}}) \neq 0$ thanks to the super-criticality assumption **(H.4)**, we define the probability $\mathbb{P}_{\overline{\mathcal{E}}}$ on $(\Omega, \mathcal{A})$ by

$$\mathbb{P}_{\overline{\mathcal{E}}}(A) = \frac{\mathbb{P}(A \cap \overline{\mathcal{E}})}{\mathbb{P}(\overline{\mathcal{E}})} \qquad \text{for all } A \in \mathcal{A}.$$

Theorem 3.5. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 4$*, the following central limit theorem holds*

$$|\mathbb{T}_{n-1}^*|^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{S}^{-1}\boldsymbol{\Gamma}\boldsymbol{S}^{-1}) \quad \textit{on} \quad (\overline{\mathcal{E}}, \mathbb{P}_{\overline{\mathcal{E}}}) \tag{3.3}$$

with $\boldsymbol{S}$ *defined in Proposition 4.14 and* $\boldsymbol{\Gamma}$ *in Lemma 5.4. If moreover* $\kappa \geq 8$*,*

$$|\mathbb{T}_{n-1}^*|^{1/2}(\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \boldsymbol{U}^{-1}\boldsymbol{\Gamma}^{\boldsymbol{\sigma}}\boldsymbol{U}^{-1}\Big) \quad \textit{on} \quad (\overline{\mathcal{E}}, \mathbb{P}_{\overline{\mathcal{E}}}), \tag{3.4}$$

and

$$|\mathbb{T}_{n-1}^*|^{1/2}(\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{V}^{-1}\boldsymbol{\Gamma}^{\boldsymbol{\rho}}\boldsymbol{V}^{-1}) \quad \textit{on} \quad (\overline{\mathcal{E}}, \mathbb{P}_{\overline{\mathcal{E}}}), \tag{3.5}$$

with $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *defined in Proposition 4.14 and* $\boldsymbol{\Gamma}^{\boldsymbol{\sigma}}$ *and* $\boldsymbol{\Gamma}^{\boldsymbol{\rho}}$ *defined in equation (6.1) and (6.2).*

The proofs of these theorems are detailed in the next sections.

4. Bifurcating Markov chains and consistency

In order to investigate the convergence of our estimators, we need laws of large numbers for quantities such as $(\delta_{2k+i} X_k^q X_{2k}^r X_{2k+1}^s)_{k \in \mathbb{T}}$. To obtain them, we use the bifurcating Markov chain framework introduced by Guyon in [13] and adapted to Galton−Watson trees by Delmas and Marsalle in [11]. We first recall the general framework, then prove the ergodicity of the induced Markov chain and finally derive strong laws of large numbers. We conclude this section by establishing the strong consistency of our estimators. Note that we cannot directly use the results in [11] because our noise sequences do not have moments of all order. Therefore, our first step is to provide a general result for bifurcating Markov chains on GW trees with only a finite number of moments.

4.1. Bifurcating Markov chain

Let $\mathbb{B}$ be the Borel σ-field of $\mathbb{R}$, and $\mathbb{B}^p$ be the Borel σ-field of $\mathbb{R}^p$, for $p > 1$. We add a cemetery point ∂ to $\mathbb{R}$, denote by $\overline{\mathbb{R}}$ the set $\mathbb{R} \cup \{\partial\}$, and by $\overline{\mathbb{B}}$ the σ-field generated by $\mathbb{B}$ and $\{\partial\}$. This cemetery point models the state of a non-observed cell. We recall the following definitions from [11].

Definition 4.1. We call $\mathbb{T}^*$-transition probability any mapping P from $\overline{\mathbb{R}} \times \overline{\mathbb{B}}^2$ onto $[0, 1]$ such that

- $P(\cdot, A)$ is measurable for all A in $\overline{\mathbb{B}}^2$;
- $P(x, \cdot)$ is probability measure on $(\overline{\mathbb{R}}^2, \overline{\mathbb{B}}^2)$ for all x in $\overline{\mathbb{R}}$;
- $P(\partial, \{(\partial, \partial)\}) = 1$.

For any measurable function f from $\overline{\mathbb{R}}^3$ onto $\mathbb{R}$, one defines the measurable function Pf from $\overline{\mathbb{R}}$ onto $\mathbb{R}$ by

$$Pf(x) = \int f(x, y, z) P(x, \mathrm{d}y, \mathrm{d}z),$$

provided the integral is well defined. Let ν be a probability measure on $\mathbb{R}$. In the sequel, ν will denote the distribution of X_1.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Definition 4.2. We say that $(Z_n)_{n\in\mathbb{T}}$ is a bifurcating Markov chain with initial distribution ν and $\mathbb{T}^*$-transition probability P, a P-BMC in short, if

- Z_1 has distribution ν;
- for all n in $\mathbb{N}$, and for all family of measurable bounded functions $(f_k)_{k\in\mathbb{G}_n}$ on $\mathbb{R}^2$;

$$\mathbb{E}\left[\prod_{k\in\mathbb{G}_n} f_k(Z_{2k}, Z_{2k+1}) \;\middle|\; \sigma(Z_j, j\in\mathbb{T}_n)\right] = \prod_{k\in\mathbb{G}_n} Pf_k(Z_k).$$

As explained in [13], this means that given the first n generations $\mathbb{T}_n$, one builds generation $\mathbb{G}_{n+1}$ by drawing 2^n independent couples (Z_{2k}, Z_{2k+1}) according to $P(Z_k,\cdot)$, $k\in\mathbb{G}_n$. This also means that any couple (Z_{2k}, Z_{2k+1}) depends on past generations only through its mother Z_k. The assumption $P(\partial,\{(\partial,\partial)\}) = 1$ means that ∂ is an absorbing state, and this hypothesis corresponds to the fact that a cell that is not observed cannot give birth to an observed one. We also assume that $P(x,\mathbb{R}^2)$, $P(x,\mathbb{R}\times\{\partial\})$ and $P(x,\{\partial\}\times\mathbb{R})$ do not depend on $x\in\mathbb{R}$. The P-BMC is thus said to be spatially homogeneous. Such a spatially homogeneous P-BMC with an absorbing cemetery state is called a bifurcating Markov chain on a Galton−Watson tree, see [11] for details.

Now let us turn back to our observed R-BAR process. In order to use the framework of P-BMC's, we define the auxilliary process $(X_n^*)_{n\in\mathbb{T}}$ by

$$X_n^* = X_n \mathbb{1}_{\{\delta_n=1\}} + \partial \mathbb{1}_{\{\delta_n=0\}}, \tag{4.1}$$

which means that $X_n^* = X_n$ if cell n is observed, $X_n^* = \partial$ the cemetery state otherwise. It is clear from assumptions **(H.1)** and **(H.3)** that the process $(X_n^*)_{n\in\mathbb{T}}$ is a P-BMC on a GW tree with $\mathbb{T}^*$-transition probability given for all x in $\mathbb{R}$ and all measurable non-negative functions f on $\overline{\mathbb{R}}^3$ by

$$\begin{aligned} Pf(x) = {} & p_{01}\mathbb{E}\left[f\big(x,(b+\eta_2)x+a+\varepsilon_2,(d+\eta_3)x+c+\varepsilon_3\big)\right] + p_0\mathbb{E}\left[f\big(x,(b+\eta_2)x+a+\varepsilon_2,\partial\big)\right] \\ & + p_1\mathbb{E}\left[f\big(x,\partial,(d+\eta_3)x+c+\varepsilon_3\big)\right] + (1-p_{01}-p_0-p_1)f(x,\partial,\partial), \end{aligned} \tag{4.2}$$

if $x\neq\partial$ and $Pf(\partial) = f(\partial,\partial,\partial)$. As explained in [13], the asymptotic behavior of the P-BMC is driven by that of the *induced* Markov chain (Y_n) defined on $\mathbb{R}$ as follows.

- For all $n\geq 1$, define the sequence $(A_n, B_n)_{n\geq 1}$ to be i.i.d. random variables with the same distribution as $(a_{2+\zeta}+\varepsilon_{2+\zeta}, b_{2+\zeta}+\eta_{2+\zeta})$, where ζ is a Bernoulli random variable with mean $(p_{01}+p_1)/m$ independent from $(\varepsilon_2,\eta_2,\varepsilon_3,\eta_3)$.
- Then, set $Y_0 = X_1^* = X_1$ and define Y_{n+1} recursively by

$$Y_{n+1} = A_{n+1} + B_{n+1}Y_n. \tag{4.3}$$

The sequence $(Y_n)_{n\in\mathbb{N}}$ is clearly an $\mathbb{R}$-valued Markov chain with transition kernel given for all x in $\mathbb{R}$ and A in $\mathbb{B}$ by

$$Q(x,A) = \frac{P_0(x,A)+P_1(x,A)}{m}, \tag{4.4}$$

with $P_i(x,A) = (p_{01}+p_i)\mathbb{E}\left[\mathbb{1}_A\big((b_{2+i}+\eta_{2+i})x + a_{2+i}+\varepsilon_{2+i}\big)\right]$. Note that the P_0 and P_1 are sub-probability kernels on $(\mathbb{R},\mathbb{B})$, whereas Q is a proper probability kernel on $(\mathbb{R},\mathbb{B})$.

4.2. Ergodicity of the induced Markov chain

We now turn to the investigation of ergodicity for the induced Markov chain $(Y_n)_{n\in\mathbb{N}}$. We start with some preliminary results on the random variables A_1 and B_1.

Lemma 4.3. *Under assumptions* **(H.2)** *and* **(H.5)**, *the random variables* A_1 *and* B_1 *have moments of all order up to* 4γ. *In addition,* $\mathbb{E}[\log|B_1|] < 0$ *and for all* $0 < s \leq 4\kappa$, *the inequality* $\mathbb{E}[|B_1|^s] < 1$ *holds.*

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proof. First, for all $0 \leq s \leq 4\gamma$, the following equalities clearly hold

$$\mathbb{E}[|A_1|^s] = \frac{p_0+p_{01}}{m}\mathbb{E}[|a+\varepsilon_2|^s] + \frac{p_1+p_{01}}{m}\mathbb{E}[|c+\varepsilon_3|^s], \quad \mathbb{E}[|B_1|^s] = \frac{p_0+p_{01}}{m}\mathbb{E}[|b+\eta_2|^s] + \frac{p_1+p_{01}}{m}\mathbb{E}[|d+\eta_3|^s].$$

Hence, under assumption **(H.2)**, it is clear that $\mathbb{E}[|A_1|^s]$ and $\mathbb{E}[|B_1|^s]$ are finite. Next, notice that the function $s \longmapsto \mathbb{E}[|B_1|^s]$ is convex, that $\mathbb{E}[|B_1|^0] = 1$ and $\mathbb{E}[|B_1|^{4\kappa}] < 1$ by assumption **(H.5)**. This implies that $\mathbb{E}[|B_1|^s] < 1$ for all $0 < s \leq 4\kappa$. Last, consider $\mathbb{E}[|\log|B_1||]$: if it is finite, $\mathbb{E}[\log|B_1|]$ is the right-derivative at 0 of $s \longmapsto \mathbb{E}[|B_1|^s]$, and convexity arguments with assumption **(H.5)** yield that $\mathbb{E}[\log|B_1|] < 0$; if it is infinite, the moment assumption on B_1 gives that necessarily $\mathbb{E}[(\log|B_1|)^+] < \infty$ and $\mathbb{E}[(\log|B_1|)^-] = \infty$, so that finally $\mathbb{E}[\log|B_1|] = -\infty < 0$, as expected. □

The next result states the existence of an invariant distribution for the Markov chain $(Y_n)_{n\in\mathbb{N}}$. It is well known as (Y_n) is a real-valued auto-regressive process with random i.i.d. coefficients satisfying Lemma 4.3, see *e.g.* [5,8].

Lemma 4.4. *Under assumptions* **(H.2)** *and* **(H.5)**, *there exists a probability distribution* μ *on* $(\mathbb{R},\mathbb{B})$ *(which is the distribution of the convergent series* $Y_\infty = \sum_{\ell=1}^{\infty} B_1B_2\dots B_{\ell-1}A_\ell$*) such that for all continuous bounded functions* f *on* $\mathbb{R}$ *and all* x *in* $\mathbb{R}$, *the following equality holds*

$$\mathbb{E}_x[f(Y_n)] \xrightarrow[n\to\infty]{} \int f\mathrm{d}\mu = \langle \mu, f\rangle.$$

We investigate the moments of the invariant distribution μ to extend the above result to polynomial functions. For all $s \geq 1$, set $\|X\|_s = (\mathbb{E}[|X|^s])^{1/s}$.

Lemma 4.5. *Under assumptions* **(H.2)** *and* **(H.5)**, μ *has moments of all order up to* 4κ. *In addition, for all* $1 \leq s \leq 4\kappa$, *all* $x \in \mathbb{R}$ *and all* $n \in \mathbb{N}$, $(\mathbb{E}_x[|Y_n|^s])^{1/s} \leq |x| + \|A_1\|_s/(1-\|B_1\|_s) < \infty$.

Proof. Set $1 \leq s \leq 4\kappa$. As the sequence (A_n, B_n) is i.i.d., the following inequality holds

$$\mathbb{E}[|Y_\infty|^s]^{1/s} = \mathbb{E}\left[\left|\sum_{\ell=1}^{\infty} B_1\dots B_{\ell-1}A_\ell\right|^s\right]^{1/s} \leq \sum_{\ell=1}^{\infty}\|B_1\|_s^{\ell-1}\|A_1\|_s.$$

Since $\mathbb{E}[|B_1|^s] < 1$ and $\mathbb{E}[|A_1|^s] < \infty$ thanks to Lemma 4.3, the series converges. Now let us turn to Y_n. The recursive equation (4.3) yields

$$Y_n = Y_0B_1\dots B_n + \sum_{\ell=1}^{n} B_n\dots B_{\ell+1}A_\ell,$$

with the usual convention that an empty product equals 1. As the sequence (A_n, B_n) is i.i.d., Y_n also has the same distribution (under $\mathbb{P}_x$) as

$$xB_1\dots B_n + \sum_{\ell=1}^{n} B_1\dots B_{\ell-1}A_\ell, \tag{4.5}$$

so that, for $1 \leq s \leq 4\kappa$, the following inequality holds

$$\mathbb{E}_x[|Y_n|^s]^{1/s} \leq |x|\|B_1\|_s^n + \sum_{\ell=1}^{n}\|B_1\|_s^{\ell-1}\|A_1\|_s \leq |x| + \frac{\|A_1\|_s}{1-\|B_1\|_s},$$

hence the result. □

The next result is a direct consequence of Lemma 4.5.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Corollary 4.6. *Under assumptions* **(H.2)** *and* **(H.5)**, *all polynomial functions f of degree less than 4κ are in $L_1(\mu)$: $\langle \mu, |f| \rangle = \mathbb{E}[|f(Y_\infty)|] < \infty$.*

We state a technical domination result that will be useful in the next section.

Lemma 4.7. *Under assumptions* **(H.2)** *and* **(H.5)**, *for all polynomial functions f of degree less than $2q$ with $q \leq 2\kappa$, there exists a nonnegative polynomial function g of degree less than $2q$ such that for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}$,*

$$\left|\mathbb{E}_x[f(Y_n)]\right| \leq g(x).$$

Proof. It is sufficient to prove the result for $f(x) = x^p$ with $p \leq 2q$. For $p \geq 1$, Lemma 4.5 yields

$$\left|\mathbb{E}_x[Y_n^p]\right| \leq \left(|x| + \frac{\|A_1\|_p}{1 - \|B_1\|_p}\right)^p \leq 2^{p-1}\left(|x|^p + \frac{\|A_1\|_p^p}{(1 - \|B_1\|_p)^p}\right).$$

If p is even, we set $g(x) = 2^{p-1}\Big(x^p + \|A_1\|_p^p/(1 - \|B_1\|_p)^p\Big)$, and if p is odd, we set $g(x) = 2^{p-1}\Big(x^{p+1} + 1 + \|A_1\|_p^p/(1-\|B_1\|_p)^p\Big)$, as for all $x \in \mathbb{R}$, $|x|^p \leq x^{p+1}+1$. Notice that if p is odd and $p \leq 2q$, one also has $p+1 \leq 2q$, hence the result. □

Finally, we prove the geometric ergodicity of the induced Markov chain for polynomial functions.

Lemma 4.8. *Under assumptions* **(H.1−2)** *and* **(H.5)**, *for all polynomial functions f of degree less than $2q$ with $q \leq 2\kappa$, there exists a nonnegative polynomial function g of degree less than $2q$ and a positive constant c such that for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}$, the following inequalities hold*

$$\left|\mathbb{E}_x[f(Y_n)] - \langle \mu, f \rangle\right| \leq g(x)\|B_1\|_{4\kappa}^n, \quad \left|\mathbb{E}_\nu[f(Y_n)] - \langle \mu, f \rangle\right| \leq c\|B_1\|_{4\kappa}^n.$$

Proof. Without loss of generality, it is sufficient to prove the result for polynomials f of the form x^p with $1 \leq p \leq 2q$. Hölder's inequality yields

$$\begin{aligned}\left|\mathbb{E}_x[f(Y_n)] - \langle \mu, f \rangle\right| = \left|\mathbb{E}_x[Y_n^p - Y_\infty^p]\right| &= \left|\mathbb{E}_x[(Y_n - Y_\infty)\sum_{s=0}^{p-1} Y_n^s Y_\infty^{p-1-s}]\right| \\ &\leq \left(\mathbb{E}_x[|Y_n - Y_\infty|^p]\right)^{\frac{1}{p}} \sum_{s=0}^{p-1}\left(\mathbb{E}_x[|Y_n^s Y_\infty^{p-1-s}|^{\frac{p}{p-1}}]\right)^{\frac{p-1}{p}}.\end{aligned}$$

We are going to study the two terms above separately. For the first term, equation (4.5) and the definition of Y_∞ yield

$$\begin{aligned}(\mathbb{E}_x[|Y_n - Y_\infty|^p])^{1/p} &= \left(\mathbb{E}[|xB_1 \ldots B_n - \sum_{\ell=n+1}^{\infty} B_1 \ldots B_{\ell-1}A_\ell|^p]\right)^{1/p} \\ &\leq |x|\|B_1\|_p^n + \|A_1\|_p \frac{\|B_1\|_p^n}{1 - \|B_1\|_p} \\ &\leq \left(|x| + \frac{\|A_1\|_p}{1 - \|B_1\|_p}\right)\|B_1\|_{4\kappa}^n,\end{aligned}$$

by Lemma 4.3 as $p \leq 4\kappa$ by assumption. We now turn to the second term. Hölder's inequality with $\alpha = (p-1)/s$ and $\beta = (p-1)/(p-1-s)$ yields

$$\begin{aligned}\left(\mathbb{E}_x[|Y_n^s Y_\infty^{p-1-s}|^{\frac{p}{p-1}}]\right) &\leq \left(\mathbb{E}_x[|Y_n|^p]\right)^{s/(p-1)}\left(\mathbb{E}[|Y_\infty|^p]\right)^{(p-1-s)/(p-1)} \\ &\leq \left(|x| + \frac{\|A_1\|_p}{1 - \|B_1\|_p}\right)^s \|Y_\infty\|_p^{p-1-s},\end{aligned}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

this last upper bound coming from Lemma 4.5. Finally, one obtains

$$\left|\mathbb{E}_x[Y_n^p - Y_\infty^p]\right| \leq \|B_1\|_{4\kappa}^n \sum_{s=0}^{p-1} \left(|x| + \frac{\|A_1\|_p}{1-\|B_1\|_p}\right)^{s+1} \|Y_\infty\|_p^{p-1-s} \;\leq\; \|B_1\|_{4\kappa}^n g(x),$$

where g is a polynomial function of degree at most $2q$ by a similar argument as in the previous proof. Integrating this upper bound with respect to the initial law ν and using **(H.1)** gives the second result. □

4.3. Laws of large numbers for the *P*-BMC

We now want to prove laws of large numbers for a family of functionals of the P-BMC (X_n^*). We are interested in polynomial functions on $\overline{\mathbb{R}}$ and $\overline{\mathbb{R}}^3$ multiplied by indicators. Precisely, for all $q \geq 1$, let F_q and G_q be the vector spaces generated by the following class of functions from $\overline{\mathbb{R}}^3$ to $\overline{\mathbb{R}}$ and from $\overline{\mathbb{R}}$ to $\mathbb{R}$ respectively,

$$\{x^\alpha y^\beta \mathbb{1}_{\mathbb{R}}(y),\ x^\alpha z^\tau \mathbb{1}_{\mathbb{R}}(z),\ x^\alpha y^\beta z^\tau \mathbb{1}_{\mathbb{R}^2}(y,z), \quad 0 \leq \alpha+\beta+\tau \leq q\},$$

$$\{x^\alpha \mathbb{1}_{\mathbb{R}}(x), \quad 0 \leq \alpha \leq q\},$$

where α, β, τ are integers. We first establish some technical results needed in the main proof.

Lemma 4.9. *Let $f \in F_q$ and $h \in G_q$. Under assumption* **(H.2)**,

(i) *if $q \leq 4\gamma$, $f \in L^1(P)$ and $Pf \in G_q$,*
(ii) *if $q \leq 4\gamma$, $h \in L^1(P_0, P_1, Q)$ and P_0h, P_1h and $Qh \in G_q$,*
(iii) *if $q \leq 2\gamma$, $h \otimes h \in L^1(P)$ and $P(h \otimes h) \in G_{2q}$.*

Proof. Take $q \leq 4\gamma$ and $p \leq 2\gamma$ and remark that $Pf(\partial) = 0$ for any $f \in F_q$, so that $Pf(x) = Pf(x)\mathbb{1}_{\mathbb{R}}(x)$ for all $x \in \overline{\mathbb{R}}$. Next, take $f_1(x,y,z) = x^\alpha y^\beta z^\tau \mathbb{1}_{\mathbb{R}^2}(y,z)$ and $f_2 = x^\delta y^\epsilon \mathbb{1}_{\mathbb{R}}(y)$ in F_q, $h(x) = x^\rho \mathbb{1}_{\mathbb{R}}(x)$ in G_q, and $l(y,z) = y^i \otimes z^j \mathbb{1}_{\mathbb{R}^2}(y,z)$ with $i, j \leq p$. Equation (4.2) yields, for $i \in \{0,1\}$,

$$\begin{aligned}
P|f_1|(x) &= p_{01}|x|^\alpha \mathbb{E}\big[\big|(b+\eta_2)x + a + \varepsilon_2\big|^\beta \big|(d+\eta_3)x + c + \varepsilon_3\big|^\tau\big],\\
P|f_2|(x) &= (p_{01}+p_0)|x|^\delta \mathbb{E}\big[\big|(b+\eta_2)x + a + \varepsilon_2\big|^\epsilon\big],\\
P_i|h|(x) &= (p_{01}+p_i)\mathbb{E}\big[\big|(b_{2+i}+\eta_{2+i})x + a_{2+i} + \varepsilon_{2+i}\big|^\rho\big],\\
P|l|(x) &= p_{01}\mathbb{E}\big[\big|(b+\eta_2)x + a + \varepsilon_2\big|^i \big|(d+\eta_3)x + c + \varepsilon_3\big|^j\big].
\end{aligned}$$

Assumption **(H.2)** entails that the 4γ-th moments of $\big((b+\eta_2)x + a + \varepsilon_2\big)$ and $\big((d+\eta_3)x + c + \varepsilon_3\big)$ are finite, which gives the integrability results, since $\beta + \tau \leq q$, $\epsilon \leq q$, $\rho \leq q$ and $i + j \leq 2p$. The integrability results are thus proved. It is then obvious that $Pl(x)$ is computed the same way as $Pf_1(x)$, and $P_ih(x)$ the same way as $Pf_2(x)$. But

$$\begin{aligned}
Pf_1(x) &= p_{01} \sum_{r=0}^{\beta} \sum_{s=0}^{\tau} C_\beta^r C_\tau^s \mathbb{E}\big[(b+\eta_2)^r (a+\varepsilon_2)^{\beta-r} (d+\eta_3)^s (c+\varepsilon_3)^{\tau-s}\big] x^{r+s+\alpha},\\
Pf_2(x) &= (p_{01}+p_0) \sum_{r=0}^{\epsilon} C_\epsilon^r \mathbb{E}\big[(b+\eta_2)^r (a+\varepsilon_2)^{\epsilon-r}\big] x^{r+\delta},
\end{aligned}$$

so that Pf_1, Pf_2 and P_ih are in G_q, since $\alpha+\beta+\tau \leq q$, $\delta+\epsilon \leq q$ and $\rho \leq q$. As for Pl, it belongs to G_{2p}, since $i + j \leq 2p$. □

We are now ready to prove the main result of this section.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Theorem 4.10. *Under assumptions* **(H.1−5)**, *for all function* $f \in F_\kappa$, *the following law of large numbers holds*

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{k\in\mathbb{G}_n^*} f(X_k^*, X_{2k}^*, X_{2k+1}^*) = \langle \mu, Pf\rangle W \quad a.s.$$

Proof. This result is similar to Theorems 11 of [13] and 3.1 of [11]. The proof follows essentially the same lines and is thus shortened here, the main difference being that the class of functions F_κ does not satisfy assumptions (i)−(vi) from [11,13] mainly because F_κ is not stable by multiplication and $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$ do not have moments of all order.

For all f in F_κ, Pf is well-defined from $\overline{\mathbb{R}}$ onto $\mathbb{R}$ thanks to Lemma 4.9 as $\kappa \le \gamma$. As $Pf(\partial) = 0$, by a slight abuse of notation we will also denote Pf its restriction to $\mathbb{R}$. Thus, Pf is μ-integrable by Lemma 4.5. One has

$$m^{-n} \sum_{k\in\mathbb{G}_n^*} f(X_k^*, X_{2k}^*, X_{2k+1}^*) - \langle \mu, Pf\rangle W = \frac{1}{m^n} \sum_{k\in\mathbb{G}_n^*} \left(f(X_k^*, X_{2k}^*, X_{2k+1}^*) - \langle \mu, Pf\rangle\right) + \langle \mu, Pf\rangle \left(\frac{|\mathbb{G}_n^*|}{m^n} - W\right).$$

By equation (2.3) the second term converges to 0 a.s. as n tends to infinity. In order to prove the a.s. convergence of the first term, as in [11,13], it is sufficient to prove that

$$\sum_{n\ge 0} m^{-2n} \mathbb{E}\left[\left(\sum_{k\in\mathbb{G}_n^*} g(X_k^*, X_{2k}^*, X_{2k+1}^*)\right)^2\right] < \infty, \tag{4.6}$$

with $g = f - \langle \mu, Pf\rangle \in F_\kappa$. Thanks to Lemma 4.9, $Pg \in G_\kappa$, and as $g^2 \in F_{2\kappa}$, one also has $Pg^2 \in G_{2\kappa}$. The expectation inside the sum decomposes as

$$\mathbb{E}\left[\left(\sum_{k\in\mathbb{G}_n^*} g(X_k^*, X_{2k}^*, X_{2k+1}^*)\right)^2\right] = \mathbb{E}\left[\left(\sum_{k\in\mathbb{G}_n^*} Pg(X_k^*)\right)^2\right] + \mathbb{E}\left[\sum_{k\in\mathbb{G}_n^*} \left(Pg^2 - (Pg)^2\right)(X_k^*)\right] = C_n + D_n.$$

We study the two terms C_n and D_n separately. Let us first prove that $\sum_{n\ge 0} m^{-2n} D_n < \infty$. We can rewrite $D_n = \mathbb{E}\big[\sum_{k\in\mathbb{G}_n^*} h(X_k^*)\big]$ with $h = Pg^2 - (Pg)^2$. As seen above, $h \in G_{2\kappa}$ and therefore h is μ-integrable thanks to Lemma 4.5. To investigate the limit of $\sum m^{-2n} D_n$, we prove that $m^{-n} D_n$ has a finite limit. More precisely, the following inequality holds

$$\begin{aligned} ||m^{-n} \sum_{k\in\mathbb{G}_n^*} h(X_k^*) - \langle \mu, h\rangle W||_2 &= ||m^{-n} \sum_{k\in\mathbb{G}_n^*} \left(h(X_k^*) - \langle \mu, h\rangle\right) + \langle \mu, h\rangle (m^{-n}|\mathbb{G}_n^*| - W)||_2 \\ &\le ||m^{-n} \sum_{k\in\mathbb{G}_n^*} \left(h(X_k^*) - \langle \mu, h\rangle\right)||_2 + |\langle \mu, h\rangle| \;\; ||m^{-n}|\mathbb{G}_n^*| - W||_2. \end{aligned}$$

The second term converges to zero. For the first term, again let $l := h - \langle \mu, h\rangle \in G_{2\kappa}$ and $\langle \mu, l\rangle = 0$, and by ([11], Eq. (15), p. 2504), the following equality holds

$$||m^{-n} \sum_{k\in\mathbb{G}_n^*} \left(h(X_k^*) - \langle \mu, h\rangle\right)||_2^2 = m^{-n} \mathbb{E}_\nu[l^2(Y_n)] + 2m^{-2} \sum_{\ell=0}^{n-1} m^{-\ell} \langle \nu, Q^\ell P(Q^{n-\ell-1} l \otimes Q^{n-\ell-1} l)\rangle. \tag{4.7}$$

Concerning the first term in equation (4.7), as $l^2 \in G_{4\kappa}$, by Lemma 4.8 one obtains $\lim_{n\to\infty} \mathbb{E}_\nu[l^2(Y_n)] = \langle \mu, l^2\rangle$ and $m^{-n}\mathbb{E}[l^2(Y_n)]$ converges to 0 a.s. Concerning the second term in equation (4.7), Lemma 4.8 yields $\lim_{n\to\infty} Q^{n-\ell-1} l(x) = \lim_{n\to\infty} \mathbb{E}_x[l(Y_{n-\ell-1}] = \langle \mu, l\rangle = 0$ and by Lemma 4.7, $Q^{n-r-1} l$ is dominated by some $\phi \in G_{2\kappa}$. Moreover, using Lemma 4.9, $\phi \otimes \phi$ belongs to $F_{4\kappa}$, it is P-integrable and $P(\phi \otimes \phi)$ belongs to $G_{4\kappa}$.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

By Lemma 4.7, $Q^\ell P(\phi\otimes\phi)$ is dominated by some $\psi \in G_{4\kappa}$, which is ν-integrable by assumption **(H.2)**. Lebesgue dominated convergence theorem yields

$$\lim_{n\to\infty} \langle \nu, Q^\ell P(Q^{n-\ell-1}l \otimes Q^{n-\ell-1}l)\rangle = 0,$$

and $|\langle \nu, Q^\ell P(Q^{n-\ell-1}l \otimes Q^{n-\ell-1}l)\rangle| \leq \langle \nu, \psi\rangle$. This upper bound allows us to deal with the limit of the second term of equation (4.7). Under assumption **(H.4)**, $\sum_{\ell=0}^{n-1} m^{-\ell}$ converges and for $\epsilon > 0$, there exists ℓ_ϵ such that $\sum_{\ell=\ell_\epsilon}^{n-1} m^{-\ell}\langle \nu, \psi\rangle \leq \epsilon$. Finally, for $n > \ell_\epsilon$, we obtain

$$\left| \sum_{\ell=0}^{n-1} m^{-\ell} \langle \nu, Q^\ell P(Q^{n-\ell-1}l \otimes Q^{n-\ell-1}l)\rangle \right| \leq \sum_{\ell=0}^{\ell_\epsilon-1} m^{-\ell} |\langle \nu, Q^\ell P(Q^{n-\ell-1}l \otimes Q^{n-\ell-1}l)\rangle| + \epsilon,$$

All the terms of the left sum converge to 0 with n, which finally proves the L_2-convergence of $m^{-n}\sum_{k\in\mathbb{G}_n^*} h(X_k^*)$ to $\langle \mu, h\rangle W$. This implies the convergence of the expectation $m^{-n}D_n$ to $\langle \mu, h\rangle \mathbb{E}[W]$ (recall that W is square-integrable). Therefore, one obtains $\sum_{n\geq 0} m^{-2n} D_n < \infty$ because $m > 1$.

Let us now prove that $\sum_{n\geq 0} m^{-2n} C_n < \infty$. Recall that $g \in F_\kappa$, $\langle \mu, Pg\rangle = 0$ and following equation (15) p. 2504 of [11], we obtain a new expression for C_n:

$$\begin{aligned} \frac{C_n}{m^{2n}} &= \left\| \frac{1}{m^n} \sum_{k\in\mathbb{G}_n^*} Pg(X_k^*) \right\|_2^2 \\ &= \frac{1}{m^n}\mathbb{E}_\nu[(Pg)^2(Y_n)] + \frac{2}{m^2}\sum_{\ell=0}^{n-1} \frac{\langle \nu, Q^\ell P(Q^{n-\ell-1}(Pg) \otimes Q^{n-\ell-1}(Pg))\rangle}{m^\ell}. \end{aligned} \tag{4.8}$$

The proof of the convergence of the first term of equation (4.8) is the same as the one of $\mathbb{E}_\nu[l^2(Y_q)]$, and $\sum_{n\geq 0} m^{-n}\mathbb{E}[(Pg)^2(Y_n)]$ converges. For the second term, setting $p = n - \ell - 1$, we can rewrite

$$\sum_{n\geq 0}\sum_{\ell=0}^{n-1} m^{-\ell} \langle \nu, Q^\ell P(Q^{n-\ell-1}(Pg) \otimes Q^{n-\ell-1}(Pg))\rangle = \sum_{\ell\geq 0} m^{-\ell} \left\langle \nu, Q^\ell P \left(\sum_{p\geq 0} (Q^p(Pg) \otimes Q^p(Pg)) \right) \right\rangle.$$

By Lemma 4.8, there exists $\varphi \in G_{\kappa+1}$, such that $|\mathbb{E}_x[(Pg)(Y_p)]| = |Q^p(Pg)(x)| \leq \varphi(x)\|B_1\|_{4\kappa}^p$ and therefore the following inequality holds

$$\left| \sum_{p\geq 0} (Q^p(Pg) \otimes Q^p(Pg)) \right| \leq (\varphi\otimes\varphi) \sum_{p\geq 0} \|B_1\|_{4\kappa}^p.$$

By assumption **(H.5)**, the series converges and there only remains to study the asymptotic behavior of $\sum_{\ell\geq 0} m^{-\ell}\langle \nu, Q^\ell P(\varphi\otimes\varphi)\rangle$. For this, let us remark that $\langle \nu, Q^\ell P(\varphi\otimes\varphi)\rangle = \mathbb{E}_\nu[P(\varphi\otimes\varphi)(Y_\ell)]$ with $P(\varphi\otimes\varphi) \in G_{2\kappa+2}$. By Lemma 4.8, $\lim_{\ell\to\infty} \mathbb{E}_\nu[P(\varphi\otimes\varphi)(Y_\ell)]$ is finite and the series converges because $m > 1$. We have thus proved that equation (4.6) holds, and hence the almost sure convergence of the series $m^{-n}\sum_{\ell\in\mathbb{G}_n^*} f(X_\ell^*, X_{2\ell}^*, X_{2\ell+1}^*)$ to $\langle \mu, Pf\rangle W$. □

4.4. Laws of large numbers for the R-BAR process

Let us now turn back to our R-BAR process and see how the law of large numbers given by Theorem 4.10 applies to our process.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proposition 4.11. *Under assumptions* **(H.1−5)**, *for all integers* $0 \leq q \leq \kappa$, *and all* $i \in \{0,1\}$, *the following laws of large numbers hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{|\mathbb{T}_n^*|} \sum_{k\in\mathbb{T}_n^*} \delta_{2k+i} X_k^q = \ell_i(q) \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{|\mathbb{T}_n^*|} \sum_{k\in\mathbb{T}_n^*} \delta_{2k}\delta_{2k+1} X_k^q = \ell_{01}(q) \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

with $\ell_i(q) = (p_{01}+p_i)\mathbb{E}[Y_\infty^q]$ *and* $\ell_{01}(q) = p_{01}\mathbb{E}[Y_\infty^q]$.

Proof. Set $q \leq \kappa$. We apply Theorem 4.10 to the function $f_0(x,y,z) = x^q \mathbb{1}_{\mathbb{R}}(y)$ if $i=0$ and $f_1(x,y,z) = x^q \mathbb{1}_{\mathbb{R}}(z)$ if $i=1$ for the first limit, and $f_{01}(x,y,z) = x^q \mathbb{1}_{\mathbb{R}^2}(y,z)$ for the second limit. The functions f_0, f_1 and f_{01} clearly belong to F_κ, and moreover $Pf_i(x) = (p_{01}+p_i)x^q$, $Pf_{01}(x) = p_{01}x^q$. Finally, notice that $\langle \mu, x^q\rangle = \mathbb{E}[Y_\infty^q]$. Theorem 4.10 thus yields

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{k\in\mathbb{G}_n^*} \delta_{2k+i} X_k^q = \ell_i(q) W, \quad \lim_{n\to\infty} \frac{1}{m^n} \sum_{k\in\mathbb{G}_n^*} \delta_{2k}\delta_{2k+1} X_k^q = \ell_{01}(q) W \quad \text{a.s.}$$

Now, for instance, the following decomposition holds

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{k\in\mathbb{T}_n^*} \delta_{2k+i} X_k^q = \sum_{\ell=0}^{n} \frac{1}{m^{n-\ell}} \left(\frac{1}{m^\ell} \sum_{k\in\mathbb{G}_\ell^*} \delta_{2k+i} X_k^q \right).$$

The sum above converges to $\ell_i(q) W m/(m-1)$ thanks to Lemma A.3 of [3] and we conclude using equation (2.4). □

Proposition 4.12. *Under assumptions* **(H.1−5)**, *for all integers* $0 \leq q \leq \kappa-1$, *and all* $i \in \{0,1\}$, *the following almost sure convergences hold*

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}^*} \delta_{2k+i} X_k^q X_{2k+i} = (p_{01}+p_i)(a_{2+i}\mathbb{E}[Y_\infty^q] + b_{2+i}\mathbb{E}[Y_\infty^{q+1}])\mathbb{1}_{\overline{\mathcal{E}}},$$

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}^*} \delta_{2k}\delta_{2k+1} X_k^q X_{2k+i} = p_{01}(a_{2+i}\mathbb{E}[Y_\infty^q] + b_{2+i}\mathbb{E}[Y_\infty^{q+1}])\mathbb{1}_{\overline{\mathcal{E}}},$$

and if $\kappa \geq 2$, *for all integers* $0 \leq q \leq \kappa-2$, *the following almost sure convergences holds*

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}^*} \delta_{2k+i} X_k^q X_{2k+i}^2 \\
&= (p_{01}+p_i)\big((a_{2+i}^2+\sigma_\varepsilon^2)\mathbb{E}[Y_\infty^q] + 2(a_{2+i}b_{2+i}+\rho_{ii})\mathbb{E}[Y_\infty^{q+1}] + (b_{2+i}^2+\sigma_\eta^2)\mathbb{E}[Y_\infty^{q+2}]\big)\mathbb{1}_{\overline{\mathcal{E}}}, \\
&\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}^*} \delta_{2k}\delta_{2k+1} X_k^q X_{2k} X_{2k+1} \\
&= p_{01}\big((ac+\rho_\varepsilon)\mathbb{E}[Y_\infty^q] + (ad+bc+2\rho)\mathbb{E}[Y_\infty^{q+1}] + (bd+\rho_\eta)\mathbb{E}[Y_\infty^{q+2}]\big)\mathbb{1}_{\overline{\mathcal{E}}}.
\end{aligned}$$

Proof. The proof follows the same lines as that of Proposition 4.11. □

We end this section by stating how to compute the moments of the invariant law μ.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Lemma 4.13. *Under assumptions* **(H.2)** *and* **(H.5)**, *the first moments of* Y_∞ *are*

$$\mathbb{E}[Y_\infty] = \frac{\mathbb{E}[A_1]}{1-\mathbb{E}[B_1]}, \qquad \mathbb{E}[Y_\infty^2] = \frac{\mathbb{E}[A_1^2]+2\mathbb{E}[A_1B_1]\mathbb{E}[Y_\infty]}{1-\mathbb{E}[B_1^2]},$$

and more generally, the moments of Y_∞ *can be calculated recursively for all* $1 \le q \le 4\kappa$ *thanks to the relation* $\mathbb{E}[Y_\infty^q] = \sum_{s=0}^{q} C_q^s \mathbb{E}[A_1^{q-s}B_1^s]\mathbb{E}[Y_\infty^s]$.

Proof. As Y_∞ is the stationary solution of equation $Y_n = A_n + B_nY_{n-1}$, Y_∞ has the same law as $A_0 + B_0Y_\infty$ where (A_0, B_0) is a copy of (A_1, B_1) independent from the sequence $(A_n, B_n)_{n\ge1}$. Hence, we can write $\mathbb{E}[Y_\infty] = \mathbb{E}[A_0 + B_0Y_\infty] = \mathbb{E}[A_1] + \mathbb{E}[B_1]\mathbb{E}[Y_\infty]$. Similarly, one has

$$\mathbb{E}[Y_\infty^2] = \mathbb{E}[(A_0 + B_0Y_\infty)^2] \;=\; \mathbb{E}[A_1^2] + 2\mathbb{E}[A_1B_1]\mathbb{E}[Y_\infty] + \mathbb{E}[B_1^2]\mathbb{E}[Y_\infty^2].$$

The general formula is obtained in the same way by developing the relation $\mathbb{E}[Y_\infty^q] = \mathbb{E}[(A_0 + B_0Y_\infty)^q]$. □

Note that one can easily compute the moments of A_1 and B_1 from their definition. In particular, the two following equalities hold

$$\mathbb{E}[Y_\infty] = \frac{am_0 + cm_1}{1 - bm_0 - dm_1},$$

and

$$\mathbb{E}[Y_\infty^2] = \frac{a^2m_0 + c^2m_1 + \sigma_\varepsilon^2 + 2\big((ab+\rho_{00})m_0 + (cd+\rho_{11})m_1\big)\mathbb{E}[Y_\infty]}{1 - (b^2m_0 + d^2m_1 + \sigma_\eta^2)},$$

with $m_0 = (p_{01} + p_0)/m$ and $m_1 = (p_{01} + p_1)/m$.

4.5. Consistency of the estimators

We are now able to prove the consistency of our estimators. We start with the computation of the limits of the normalizing matrices $\boldsymbol{S}_n$, $\boldsymbol{U}_n$ and $\boldsymbol{V}_n$, which is a direct consequence of Proposition 4.11

Proposition 4.14. *Under assumptions* **(H.1−5)**, *and if* $\kappa \ge 2$, *for* $i \in \{0,1\}$, *the following laws of large numbers hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\boldsymbol{S}_n^i}{|\mathbb{T}_n^*|} = \boldsymbol{S}^i \mathbb{1}_{\overline{\mathcal{E}}} = (p_{01}+p_i) \begin{pmatrix} 1 & \mathbb{E}[Y_\infty] \\ \mathbb{E}[Y_\infty] & \mathbb{E}[Y_\infty^2] \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

and

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\boldsymbol{S}_n}{|\mathbb{T}_n^*|} = \boldsymbol{S} \mathbb{1}_{\overline{\mathcal{E}}} = \begin{pmatrix} \boldsymbol{S}^0 & 0 \\ 0 & \boldsymbol{S}^1 \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

If in addition $\kappa \ge 4$, *the following convergences hold*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\boldsymbol{U}_n}{|\mathbb{T}_n^*|} \;=\; \boldsymbol{U} \mathbb{1}_{\overline{\mathcal{E}}} = m \begin{pmatrix} 1 & 2m_0\mathbb{E}[Y_\infty] & 2m_1\mathbb{E}[Y_\infty] & \mathbb{E}[Y_\infty^2] \\ 2m_0\mathbb{E}[Y_\infty] & 4m_0\mathbb{E}[Y_\infty^2] & 0 & 2m_0\mathbb{E}[Y_\infty^3] \\ 2m_1\mathbb{E}[Y_\infty] & 0 & 4m_1\mathbb{E}[Y_\infty^2] & 2m_1\mathbb{E}[Y_\infty^3] \\ \mathbb{E}[Y_\infty^2] & 2m_0\mathbb{E}[Y_\infty^3] & 2m_1\mathbb{E}[Y_\infty^3] & \mathbb{E}[Y_\infty^4] \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

and

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\boldsymbol{V}_n}{|\mathbb{T}_n^*|} = \boldsymbol{V} \mathbb{1}_{\overline{\mathcal{E}}} = p_{01} \begin{pmatrix} 1 & 2\mathbb{E}[Y_\infty] & \mathbb{E}[Y_\infty^2] \\ 2\mathbb{E}[Y_\infty] & 4\mathbb{E}[Y_\infty^2] & 2\mathbb{E}[Y_\infty^3] \\ \mathbb{E}[Y_\infty^2] & 2\mathbb{E}[Y_\infty^3] & \mathbb{E}[Y_\infty^4] \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

Besides, the matrices $\boldsymbol{S}^i$, $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *are invertible.*

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

We now turn to the consistency of our main estimators.

Proof of Theorem 3.1. As regards our main estimator $\widehat{\boldsymbol{\theta}}_n$, a direct application of Proposition 4.12 yields

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{|\mathbb{T}_{n-1}^*|} \boldsymbol{S}_{n-1}\widehat{\boldsymbol{\theta}}_n = m \begin{pmatrix} m_0(a+b\mathbb{E}[Y_\infty]) \\ m_0(a\mathbb{E}[Y_\infty]+b\mathbb{E}[Y_\infty^2]) \\ m_1(c+d\mathbb{E}[Y_\infty]) \\ m_1(c\mathbb{E}[Y_\infty]+d\mathbb{E}[Y_\infty^2]) \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} = \boldsymbol{S\theta}\mathbb{1}_{\overline{\mathcal{E}}} \text{ a.s.}$$

and the result follows from Proposition 4.14 and the definition of $\widehat{\boldsymbol{\theta}}_n$. The consistency of $\widehat{\boldsymbol{\sigma}}_n$ and $\widehat{\boldsymbol{\rho}}_n$ is a bit more complicated because their definition involves the $\widehat{\epsilon}_k$. We give a detailed proof of the convergence of $|\mathbb{T}_{n-1}^*|^{-1}\sum \widehat{\epsilon}_{2k}^2$, the other terms in $\boldsymbol{U}_{n-1}\widehat{\boldsymbol{\sigma}}_n$ and $\boldsymbol{V}_{n-1}\widehat{\boldsymbol{\rho}}_n$ being treated similarly. For $k \in \mathbb{G}_n$, one can develop

$$\begin{aligned} \widehat{\epsilon}_{2k}^2 &= \delta_{2k}(X_{2k} - \widehat{a}_n - \widehat{b}_n X_k)^2 \\ &= \delta_{2k}(\widehat{a}_n^2 + 2\widehat{a}_n\widehat{b}_n X_k + \widehat{b}_n^2 X_k^2 - 2\widehat{a}_n X_{2k} - 2\widehat{b}_n X_k X_{2k} + X_{2k}^2). \end{aligned}$$

Hence, the following equality holds

$$\begin{aligned} \sum_{k\in\mathbb{T}_{n-1}^*} \widehat{\epsilon}_{2k}^2 = &\sum_{\ell=1}^{n-1} \widehat{a}_\ell^2 \sum_{k\in\mathbb{G}_\ell} \delta_{2k} + 2\sum_{\ell=1}^{n-1} \widehat{a}_\ell\widehat{b}_\ell \sum_{k\in\mathbb{G}_\ell} \delta_{2k}X_k + \sum_{\ell=1}^{n-1} \widehat{b}_\ell^2 \sum_{k\in\mathbb{G}_\ell} \delta_{2k}X_k^2 \\ &-2\sum_{\ell=1}^{n-1} \widehat{a}_\ell \sum_{k\in\mathbb{G}_\ell} \delta_{2k}X_{2k} - 2\sum_{\ell=1}^{n-1} \widehat{b}_\ell \sum_{k\in\mathbb{G}_\ell} \delta_{2k}X_k X_{2k} + \sum_{k\in\mathbb{T}_{n-1}^*} \delta_{2k}X_{2k}^2. \end{aligned} \tag{4.9}$$

The limit of the last term is given by Proposition 4.12. The first term decomposes as

$$\frac{1}{m^{n-1}} \sum_{\ell=1}^{n-1} \widehat{a}_\ell^2 \sum_{k\in\mathbb{G}_\ell} \delta_{2k} = \sum_{\ell=1}^{n-1} \widehat{a}_\ell^2 \frac{m^\ell}{m^{n-1}} \frac{1}{m^\ell} \sum_{k\in\mathbb{G}_\ell} \delta_{2k}.$$

We apply Lemma A.3 of [3] to the sequence above. On the one hand,

$$\lim_{\ell\to\infty} \widehat{a}_\ell^2 \frac{1}{m^\ell} \sum_{k\in\mathbb{G}_\ell} \delta_{2k} = a^2(p_{01}+p_0)W \quad \text{a.s.}$$

thanks to the previous result on the consistency of $\widehat{\boldsymbol{\theta}}_n$ and Theorem 4.10. On the other hand, the series $\sum m^{-n}$ converges to $m/(m-1)$ under assumption **(H.4)**. Therefore, Lemma A.3 of [3] yields

$$\lim_{n\to\infty} \frac{1}{m^{n-1}} \sum_{\ell=1}^{n-1} \widehat{a}_\ell^2 \sum_{k\in\mathbb{G}_\ell} \delta_{2k} = \frac{m}{m-1} a^2(p_{01}+p_0)W \quad \text{a.s.}$$

and equation (2.4) finally yields

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{|\mathbb{T}_{n-1}^*|} \sum_{\ell=1}^{n-1} \widehat{a}_\ell^2 \sum_{k\in\mathbb{G}_\ell} \delta_{2k} = a^2(p_{01}+p_0)\mathbb{1}_{\overline{\mathcal{E}}} \quad \text{a.s.}$$

Note that the limit above is just the limit of $\widehat{a}_\ell^2$ multiplied by the limit of $|\mathbb{T}_{n-1}^*|^{-1}\sum \delta_{2k}$. The other terms in equation (4.9) are dealt with similarly using the results of Proposition 4.12. Finally, one obtains the almost sure

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

convergences

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{|\mathbb{T}_{n-1}^*|} \boldsymbol{U}_{n-1}\widehat{\boldsymbol{\sigma}}_n = \lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}} \begin{pmatrix} \widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2 \\ 2X_k\widehat{\epsilon}_{2k}^2 \\ 2X_k\widehat{\epsilon}_{2k+1}^2 \\ X_k^2(\widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2) \end{pmatrix} = \boldsymbol{U}\boldsymbol{\sigma}\mathbb{1}_{\overline{\mathcal{E}}},$$

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}1}{|\mathbb{T}_{n-1}^*|} \boldsymbol{V}_{n-1}\widehat{\boldsymbol{\rho}}_n = \lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}1}{|\mathbb{T}_{n-1}^*|} \sum_{k\in\mathbb{T}_{n-1}} \begin{pmatrix} \widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1} \\ 2X_k\widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1} \\ X_k^2\widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1} \end{pmatrix} = \boldsymbol{V}\boldsymbol{\rho}\mathbb{1}_{\overline{\mathcal{E}}},$$

hence the result using Proposition 4.14. □

5. Martingales and convergence rate

The aim of this section is to obtain sharper convergence results for our estimators, namely rates of convergence. The P-BMC approach does not allow this, therefore we now use martingale theory instead, as in [3,9]. However, we cannot directly apply the results therein mainly because our noise sequence ($\epsilon_k = \varepsilon_k + \eta_k X_{[k/2]}$) now contains the BAR process (X_k) and thus does not satisfy the assumptions of [3,9].

5.1. Martingales on binary trees

We start with a general result of convergence for martingales on a Galton−Watson binary tree, that we will make repeatedly use of in the following sections. Special cases of this result have already been proved and used in [3,9]. Note that in this binary tree context, we cannot use the standard asymptotic theory for vector martingales (see *e.g.* [12]) because the number of data is roughly multiplied by m at each generation.

Theorem 5.1. *Let $(\boldsymbol{M}_n)$ be a p-dimensional $\mathbb{F}^{\mathcal{O}}$-martingale on the GW-binary tree $\mathbb{T}^*$: $\boldsymbol{M}_n = \sum_{\ell=1}^n \sum_{k\in\mathbb{G}_\ell^*} \boldsymbol{W}_k$, with $\boldsymbol{W}_k = (w_k^1, w_k^2, \ldots, w_k^p)^t$. We make the following assumptions:*

(A.1) *$(\boldsymbol{M}_n)$ is square-integrable.*

Let $< \boldsymbol{M} >_n = \sum_{\ell=0}^{n-1} \boldsymbol{\Gamma}_\ell$ be the predictable quadratic variation of $(\boldsymbol{M}_n)$, with

$$\boldsymbol{\Gamma}_n = \mathbb{E}[\Delta\boldsymbol{M}_{n+1}\Delta\boldsymbol{M}_{n+1}^t \mid \mathcal{F}_n^{\mathcal{O}}].$$

(A.2) *$|\mathbb{T}_{n-1}^*|^{-1} < \boldsymbol{M} >_n$ converges almost surely to a positive semi-definite matrix $\boldsymbol{\Gamma}$ on $\overline{\mathcal{E}}$.*

(A.3) *The $p\times p$ $\mathbb{F}^{\mathcal{O}}$-matrix martingale $(\boldsymbol{K}_n)$ defined by*

$$\boldsymbol{K}_n = \sum_{\ell=1}^n |\mathbb{T}_\ell^*|^{-1}(\Delta\boldsymbol{M}_{\ell+1}\Delta\boldsymbol{M}_{\ell+1}^t - \mathbb{E}[\Delta\boldsymbol{M}_{\ell+1}\Delta\boldsymbol{M}_{\ell+1}^t \mid \mathcal{F}_\ell^{\mathcal{O}}])$$

is square-integrable and its component-wise predictable quadratic variationes are $\mathcal{O}(n)$ a.s. on $\overline{\mathcal{E}}$.

Let $(\boldsymbol{\Xi}_n)$ be a sequence of $p\times p$ invertible symmetric matrices such that

(A.4) *$|\mathbb{T}_n^*|^{-1}\boldsymbol{\Xi}_n$ converges a.s. to an invertible matrix $\boldsymbol{\Xi}$ on $\overline{\mathcal{E}}$;*

(A.5) *there exists a positive constant α such that on the non-extinction set $\overline{\mathcal{E}}$ and for all n and the following assumptions holds $\alpha(\boldsymbol{\Xi}_{n-1}^{-1} - \boldsymbol{\Xi}_n^{-1}) \geq \boldsymbol{\Xi}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Xi}_n^{-1}$, in the sense that $\alpha(\boldsymbol{\Xi}_{n-1}^{-1} - \boldsymbol{\Xi}_n^{-1}) - \boldsymbol{\Xi}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Xi}_n^{-1}$ is a positive semi-definite matrix.*

Then $\boldsymbol{M}_n^t\boldsymbol{\Xi}_{n-1}^{-1}\boldsymbol{M}_n = \mathcal{O}(n)$ and if $\boldsymbol{\Xi}$ is positive definite $\|\boldsymbol{M}_n\|^2 = \mathcal{O}(nm^n)$ a.s.
If in addition, the entries of $(\boldsymbol{M}_n)$ satisfy

(A.6) *$\sup_n \mathbb{E}[(m^{-n/2}\sum_{k\in\mathbb{G}_n^*} w_k^i)^4 \mid \mathcal{F}_{n-1}^{\mathcal{O}}] < \infty$ almost surely,*

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

then for all $\delta > 1/2$, $\|\boldsymbol{M}_n\|^2 = o(n^\delta m^n)$ *a.s. and*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell = \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

Proof of the first part of Theorem 5.1. The result is obvious on the extinction set $\mathcal{E}$. In the sequel, let us suppose that we are on the non-extinction set $\overline{\mathcal{E}}$. For all $n \geq 1$, denote $\mathcal{V}_n = \boldsymbol{M}_n^t \boldsymbol{\Xi}_{n-1}^{-1} \boldsymbol{M}_n$. The following equalities hold

$$\begin{aligned}\mathcal{V}_{n+1} &= \boldsymbol{M}_{n+1}^t \boldsymbol{\Xi}_n^{-1} \boldsymbol{M}_{n+1} = (\boldsymbol{M}_n + \Delta\boldsymbol{M}_{n+1})^t \boldsymbol{\Xi}_n^{-1} (\boldsymbol{M}_n + \Delta\boldsymbol{M}_{n+1}),\\ &= \mathcal{V}_n - \boldsymbol{M}_n^t (\boldsymbol{\Xi}_{n-1}^{-1} - \boldsymbol{\Xi}_n^{-1}) \boldsymbol{M}_n + 2\boldsymbol{M}_n^t \boldsymbol{\Xi}_n^{-1} \Delta\boldsymbol{M}_{n+1} + \Delta\boldsymbol{M}_{n+1}^t \boldsymbol{\Xi}_n^{-1} \Delta\boldsymbol{M}_{n+1},\end{aligned}$$

since $\boldsymbol{M}_n^t \boldsymbol{\Xi}_n^{-1} \Delta\boldsymbol{M}_{n+1}$ is scalar, and hence equal to its own transpose. By summing over the identity above, we obtain

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1}, \tag{5.1}$$

where

$$\mathcal{A}_n = \sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t (\boldsymbol{\Xi}_{\ell-1}^{-1} - \boldsymbol{\Xi}_\ell^{-1}) \boldsymbol{M}_\ell, \quad \mathcal{B}_{n+1} = 2\sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t \boldsymbol{\Xi}_\ell^{-1} \Delta\boldsymbol{M}_{\ell+1}, \quad \mathcal{W}_{n+1} = \sum_{\ell=1}^{n} \Delta\boldsymbol{M}_{\ell+1}^t \boldsymbol{\Xi}_\ell^{-1} \Delta\boldsymbol{M}_{\ell+1}.$$

The asymptotic behavior of the sequences $(\mathcal{W}_n)$ and $(\mathcal{B}_n)$ is given in the following lemmas.

Lemma 5.2. *Under assumptions* (**A.1**) *to* (**A.4**)*, the following almost sure convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \mathcal{W}_n = \frac{m-1}{m} \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

Lemma 5.3. *Under assumptions* (**A.1**) *to* (**A.5**)*, the following asymptotic property holds*

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

One then obtains

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\mathcal{V}_{n+1} + \mathcal{A}_n}{n} = \frac{m-1}{m} \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \qquad \text{a.s.} \tag{5.2}$$

As $(\mathcal{A}_n)$ is a sequence of positive real numbers, it follows that $\mathcal{V}_{n+1} = \mathcal{O}(n)$ a.s., which means that $\boldsymbol{M}_n^t \boldsymbol{\Xi}_{n-1}^{-1} \boldsymbol{M}_n = \mathcal{O}(n)$ a.s. As $\boldsymbol{\Xi}$ is positive definite, one obtains for large enough n, on the non-extinction set $\overline{\mathcal{E}}$

$$\|\boldsymbol{M}_n\|^2 = \boldsymbol{M}_n^t \boldsymbol{M}_n \leq \frac{\boldsymbol{M}_n^t \boldsymbol{\Xi}_{n-1}^{-1} \boldsymbol{M}_n}{\lambda_{\min}(\boldsymbol{\Xi}_{n-1}^{-1})},$$

where $\lambda_{\min}(\boldsymbol{\Xi}_{n-1}^{-1})$ denotes the smallest eigenvalue of $\boldsymbol{\Xi}_{n-1}^{-1}$. Finally, assumption **(A.4)** yields $\|\boldsymbol{M}_n\|^2 = \mathcal{O}(nm^n)$ a.s., which completes the proof of the first part of Theorem 5.1. □

It remains to prove Lemmas 5.2 and 5.3.

Proof of Lemma 5.2. First of all, we decompose $\mathcal{W}_{n+1} = \mathcal{T}_{n+1} + \mathcal{R}_{n+1}$ with

$$\mathcal{T}_{n+1} = \sum_{\ell=1}^{n} \frac{\Delta\boldsymbol{M}_{\ell+1}^t \boldsymbol{\Xi}^{-1} \Delta\boldsymbol{M}_{\ell+1}}{|\mathbb{T}_\ell^*|}, \quad \mathcal{R}_{n+1} = \sum_{\ell=1}^{n} \frac{\Delta\boldsymbol{M}_{\ell+1}^t (|\mathbb{T}_\ell^*| \boldsymbol{\Xi}_\ell^{-1} - \boldsymbol{\Xi}^{-1}) \Delta\boldsymbol{M}_{\ell+1}}{|\mathbb{T}_\ell^*|}.$$

We first prove that $\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \mathcal{T}_n = \frac{m-1}{m} \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}}$ a.s. As $\mathcal{T}_n$ is a scalar and the trace is commutative, we can rewrite $\mathcal{T}_{n+1} = \mathrm{tr}(\boldsymbol{H}_{n+1} \boldsymbol{\Xi}^{-1})$ where

$$\boldsymbol{H}_{n+1} = \sum_{\ell=1}^{n} \frac{\Delta\boldsymbol{M}_{\ell+1} \Delta\boldsymbol{M}_{\ell+1}^t}{|\mathbb{T}_\ell^*|} = \sum_{\ell=1}^{n} \frac{\boldsymbol{\Gamma}_\ell}{|\mathbb{T}_\ell^*|} + \boldsymbol{K}_n.$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

On the one hand, by assumption **(A.3)**, one has $\boldsymbol{K}_n = o(n)$ a.s. on $\overline{\mathcal{E}}$. On the other hand, assumption **(A.2)** yields

$$\mathbb{1}_{\{|\mathbb{G}_\ell^*|>0\}} \frac{\boldsymbol{\Gamma}_\ell}{|\mathbb{T}_\ell^*|} = \mathbb{1}_{\{|\mathbb{G}_\ell^*|>0\}} \left(\frac{< \boldsymbol{M} >_{\ell+1}}{|\mathbb{T}_\ell^*|} - \frac{|\mathbb{T}_{\ell-1}^*|}{|\mathbb{T}_\ell^*|} \frac{< \boldsymbol{M} >_\ell}{|\mathbb{T}_{\ell-1}^*|} \right)$$

so that H_n converges to $(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}/m)\,\mathbb{1}_{\overline{\mathcal{E}}} = \boldsymbol{\Gamma}(m-1)/m$ a.s. as ℓ tends to infinity. Hence, Cesaro convergence yields

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \boldsymbol{H}_n = \frac{m-1}{m} \boldsymbol{\Gamma} \mathbb{1}_{\overline{\mathcal{E}}} \qquad \text{a.s.} \tag{5.3}$$

As a consequence, $\lim \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \mathcal{T}_n/n = \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1})(m-1)/m \mathbb{1}_{\overline{\mathcal{E}}}$ a.s. We now turn to the asymptotic behavior of $\mathcal{R}_{n+1}$. We know from assumption **(A.4)** that $|\mathbb{T}_\ell^*|\boldsymbol{\Xi}_\ell^{-1} - \boldsymbol{\Xi}^{-1}$ goes to 0 as ℓ goes to infinity on $\overline{\mathcal{E}}$. Thus, for all positive ϵ, there exists ℓ_ϵ such that if $\ell \geq \ell_\epsilon$,

$$\mathbb{1}_{\{|\mathbb{G}_\ell^*|>0\}} |\Delta \boldsymbol{M}_{\ell+1}^t (|\mathbb{T}_\ell| \boldsymbol{\Xi}_\ell^{-1} - \boldsymbol{\Xi}^{-1}) \Delta \boldsymbol{M}_{\ell+1}| \leq 4\epsilon \Delta \boldsymbol{M}_{\ell+1}^t \Delta \boldsymbol{M}_{\ell+1} \mathbb{1}_{\{|\mathbb{G}_\ell^*|>0\}}.$$

Hence, there exists some positive real number c_ϵ such that, for $n \geq \ell_\epsilon$,

$$\begin{aligned} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} |\mathcal{R}_n| &\leq \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \left(4\epsilon \sum_{\ell=\ell_\epsilon}^{n-1} \mathbb{1}_{\{|\mathbb{G}_\ell^*|>0\}} \frac{\Delta \boldsymbol{M}_{\ell+1}^t \Delta \boldsymbol{M}_{\ell+1}}{|\mathbb{T}_\ell^*|} + c_\epsilon \right) \\ &\leq \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \left(4\epsilon \ \mathrm{tr}(\boldsymbol{H}_n) + c_\epsilon \right). \end{aligned}$$

This last inequality holding for any positive ϵ and large enough n, the limit given by equation (5.3) entails that $\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \mathcal{R}_n = 0$ a.s., which completes the proof. □

Proof of Lemma 5.3. Again, the result is obvious on the extinction set $\mathcal{E}$. Suppose now we are on the non-extinction set $\overline{\mathcal{E}}$. Recall that

$$\mathcal{B}_{n+1} = 2 \sum_{k=1}^{n} \boldsymbol{M}_k^t \boldsymbol{\Xi}_k^{-1} \Delta \boldsymbol{M}_{k+1}.$$

The process $(\mathcal{B}_n)$ is a real-valued $\mathbb{F}^{\mathcal{O}}$-martingale. In addition, the following equality clearly holds

$$\mathbb{E}[\Delta \mathcal{B}_{n+1}^2 | \mathcal{F}_n^{\mathcal{O}}] = 4 \boldsymbol{M}_n^t \boldsymbol{\Xi}_n^{-1} \boldsymbol{\Gamma}_n \boldsymbol{\Xi}_n^{-1} \boldsymbol{M}_n \qquad \text{a.s.}$$

Assumption **(A.5)** then yields

$$\langle \mathcal{B} \rangle_{n+1} \leq 4\alpha \sum_{k=1}^{n} \boldsymbol{M}_k^t (\boldsymbol{\Xi}_{k-1}^{-1} - \boldsymbol{\Xi}_k^{-1}) \boldsymbol{M}_k = 4\alpha \mathcal{A}_n. \qquad \text{a.s.}$$

Hence, the law of large number for real martingales yields $\mathcal{B}_n = o(\mathcal{A}_n)$ a.s. Finally, we deduce from decomposition (5.1) and Lemma 5.2 that

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

leading to $\mathcal{A}_n = \mathcal{O}(n)$ and $\mathcal{V}_{n+1} = \mathcal{O}(n)$ a.s. as both sequences are non-negative. This implies in turn that $\mathcal{B}_n = o(n)$ a.s. completing the proof. □

Proof of the second part of Theorem 5.1. Let us rewrite the entries M_n^q of the martingale $\boldsymbol{M}_n$ as

$$M_n^q = \sum_{l=1}^{n} \underbrace{m^{\ell/2}}_{u_\ell} \underbrace{\frac{1}{m^{\ell/2}} \sum_{k \in \mathbb{G}_l} w_k^q}_{\tilde{\varepsilon}_\ell^q}.$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Then one just has to apply Wei's lemma given in ([26], p. 1672) to the martingale difference sequence $\tilde{\varepsilon}_\ell$ and $u_\ell = m^{\ell/2}$, and for the function $f(x) = (\log x)^\delta$ for $\delta > 1/2$. Under assumption **(A.6)**, one obtains $M_n^q = o(m^{n/2} n^{\delta/2})$. As P_n^q is the qth entry of $\boldsymbol{M}_n$, one obtains $\|\boldsymbol{M}_n\|^2 = o(n^\delta m^n)$ a.s. Now recall that $\mathcal{V}_n = \boldsymbol{M}_n^t \boldsymbol{\Xi}_{n-1}^{-1} \boldsymbol{M}_n$, therefore, the following equality holds

$$\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \mathcal{V}_n = \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{M}_n^t \boldsymbol{\Xi}_{n-1}^{-1} \boldsymbol{M}_n = o(n^{2\delta}) \quad \text{a.s.},$$

for all $\delta > 1/4$. In particular, for $\delta = 1/2$, and we have the following order $\mathcal{V}_n = o(n)$. Lemmas 5.2 and 5.3 then yield

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\mathcal{A}_n}{n} = \frac{m-1}{m} \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \quad \text{a.s.} \tag{5.4}$$

First of all, $\mathcal{A}_n$ may be rewritten as

$$\mathcal{A}_n = \sum_{\ell=1}^n \boldsymbol{M}_\ell^t (\boldsymbol{\Xi}_{\ell-1}^{-1} - \boldsymbol{\Xi}_\ell^{-1}) \boldsymbol{M}_\ell = \sum_{\ell=1}^n \boldsymbol{M}_\ell^t \boldsymbol{\Delta}_\ell \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell,$$

where $\boldsymbol{\Delta}_n = \mathbf{I}_4 - \boldsymbol{\Xi}_n^{-1} \boldsymbol{\Xi}_{n-1}$. Thanks to assumption **(A.4)** we know that

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{\Delta}_n = \frac{m-1}{m} \mathbf{I}_4 \mathbb{1}_{\overline{\mathcal{E}}} \qquad \text{a.s.}$$

Besides, equation (5.4) yields that $\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \mathcal{A}_n \sim n(m-1) m^{-1} \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}}$ a.s. Plugging these two results into the equality

$$\mathcal{A}_n = \frac{m-1}{m} \sum_{\ell=1}^n \boldsymbol{M}_\ell^t \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell + \sum_{\ell=1}^n \boldsymbol{M}_\ell^t \left(\boldsymbol{\Delta}_\ell - \frac{m-1}{m} \mathbf{I}_4 \right) \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell$$

gives that $\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \sum_{\ell=1}^n \boldsymbol{M}_\ell^t \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell \sim \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \mathcal{A}_n m (m-1)^{-1}$ a.s. Thus one obtains

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{\ell=1}^n \boldsymbol{M}_\ell^t \boldsymbol{\Xi}_{\ell-1}^{-1} \boldsymbol{M}_\ell = \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Xi}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \quad \text{a.s.}$$

which is the expected result. □

5.2. Rate of convergence for $\widehat{\boldsymbol{\theta}}_n$

We apply Theorem 5.1 to a suitably chosen martingale. Recall that

$$\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta} = \boldsymbol{S}_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} (\epsilon_{2k}, X_k \epsilon_{2k}, \epsilon_{2k+1}, X_k \epsilon_{2k+1})^t = \boldsymbol{S}_{n-1}^{-1} \boldsymbol{M}_n, \tag{5.5}$$

where

$$\boldsymbol{M}_n = \sum_{k \in \mathbb{T}_{n-1}} (\epsilon_{2k}, X_k \epsilon_{2k}, \epsilon_{2k+1}, X_k \epsilon_{2k+1})^t .$$

Under assumptions **(H.1−3)**, for all $n \geq 0$, $k \in \mathbb{G}_n$, $\mathbb{E}[\epsilon_{2k+i} | \mathcal{F}_n^{\mathcal{O}}] = \mathbb{E}[X_k \epsilon_{2k+i} | \mathcal{F}_n^{\mathcal{O}}] = 0$ and $(\boldsymbol{M}_n)$ is a square-integrable $(\mathcal{F}_n^{\mathcal{O}})$-martingale, so that assumption **(A.1)** of Theorem 5.1 holds. Let us compute the predictable quadratic variation of $(\boldsymbol{M}_n)$

$$\mathbb{E}[\Delta \boldsymbol{M}_{n+1} \Delta \boldsymbol{M}_{n+1}^t | \mathcal{F}_n^{\mathcal{O}}] = \boldsymbol{\Gamma}_n = \sum_{k \in \mathbb{G}_n} \boldsymbol{\gamma}_k \otimes \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}, \tag{5.6}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

where

$$\boldsymbol{\gamma}_k = \begin{pmatrix} \delta_{2k}(\sigma_\varepsilon^2 + 2X_k\rho_{00} + X_k^2\sigma_\eta^2) & \delta_{2k}\delta_{2k+1}(\rho_\varepsilon + 2X_k\rho + X_k^2\rho_\eta) \\ \delta_{2k}\delta_{2k+1}(\rho_\varepsilon + 2X_k\rho + X_k^2\rho_\eta) & \delta_{2k+1}(\sigma_\varepsilon^2 + 2X_k\rho_{11} + X_k^2\sigma_\eta^2) \end{pmatrix}. \tag{5.7}$$

Thus the predictable quadratic variation of $(\boldsymbol{M}_n)$ is given by

$$\langle \boldsymbol{M} \rangle_n = \sum_{\ell=0}^{n-1} \boldsymbol{\Gamma}_\ell = \sum_{k \in \mathbb{T}_{n-1}} \boldsymbol{\gamma}_k \otimes \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}.$$

Lemma 5.4. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 4$*, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\langle \boldsymbol{M} \rangle_n}{|\mathbb{T}_{n-1}^*|} = \boldsymbol{\Gamma} \mathbb{1}_{\overline{\mathcal{E}}} = \begin{pmatrix} \boldsymbol{\Gamma}^0 & \boldsymbol{\Gamma}^{01} \\ \boldsymbol{\Gamma}^{01} & \boldsymbol{\Gamma}^1 \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.,$$

where $\boldsymbol{\Gamma}^0$, $\boldsymbol{\Gamma}^{01}$ *and* $\boldsymbol{\Gamma}^1$ *are the* 2×2 *matrices defined by*

$$\boldsymbol{\Gamma}^i = \begin{pmatrix} \sigma_\varepsilon^2\ell_i(0) + 2\rho_{ii}\ell_i(1) + \sigma_\eta^2\ell_i(2) & \sigma_\varepsilon^2\ell_i(1) + 2\rho_{ii}\ell_i(2) + \sigma_\eta^2\ell_i(3) \\ \sigma_\varepsilon^2\ell_i(1) + 2\rho_{ii}\ell_i(2) + \sigma_\eta^2\ell_i(3) & \sigma_\varepsilon^2\ell_i(2) + 2\rho_{ii}\ell_i(3) + \sigma_\eta^2\ell_i(4) \end{pmatrix},$$

and

$$\boldsymbol{\Gamma}^{01} = \begin{pmatrix} \rho_\varepsilon\ell_{01}(0) + 2\rho\ell_{01}(1) + \rho_\eta\ell_{01}(2) & \rho_\varepsilon\ell_{01}(1) + 2\rho\ell_{01}(2) + \rho_\eta\ell_{01}(3) \\ \rho_\varepsilon\ell_{01}(1) + 2\rho\ell_{01}(2) + \rho_\eta\ell_{01}(3) & \rho_\varepsilon\ell_{01}(2) + 2\rho\ell_{01}(3) + \rho_\eta\ell_{01}(4) \end{pmatrix}.$$

In addition, $\boldsymbol{\Gamma}$ *is positive definite.*

Proof. This is a direct consequence of Proposition 4.11. □

Hence, assumption **(A.2)** holds if $\kappa \geq 4$. The process $(\boldsymbol{K}_n)$ is clearly a square-integrable martingale if $\gamma \geq 2$. It is not difficult to check that its component-wise predictable quadratic variation is at worst of the order of

$$\sum_{\ell=1}^{n} \frac{1}{|\mathbb{T}_\ell^*|^2} \sum_{k \in \mathbb{G}_\ell} \delta_{2k+i} X_k^8.$$

Proposition 4.11 ensures that $|\mathbb{T}_n^*|^{-1} \sum_{k\in\mathbb{G}_n} \delta_{2k+i} X_k^4$ converges almost surely on $\overline{\mathcal{E}}$ provided $\kappa \geq 4$, it is therefore bounded by some constant C. As a result, its square is also bounded by C^2 and $|\mathbb{T}_\ell^*|^{-2} \sum_{k\in\mathbb{G}_\ell} \delta_{2k+i} X_k^8 \leq C^2$ a.s. on $\overline{\mathcal{E}}$. Finally, one obtains that

$$\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \sum_{\ell=1}^{n} \frac{1}{|\mathbb{T}_\ell^*|^2} \sum_{k \in \mathbb{G}_\ell} \delta_{2k+i} X_k^8 \leq C^2 n \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}},$$

so that assumption **(A.3)** holds if $\kappa \geq 4$.

We now introduce a new sequence of matrices $\boldsymbol{\Sigma}_n$. They are defined as a standardized version of the predictable quadratic variation of $(\boldsymbol{M}_n)$, with the variance coefficients σ_ε^2 and σ_η^2 set to 1 and all the covariance coefficients ρ_ε, ρ_η, ρ_{ij} set to 0, namely

$$\boldsymbol{\Sigma}_n = \sum_{\ell=1}^{n} \boldsymbol{\Phi}_\ell \boldsymbol{\Phi}_\ell^t = \sum_{k \in \mathbb{T}_n} (1 + X_k^2) \begin{pmatrix} \delta_{2k} & 0 \\ 0 & \delta_{2k+1} \end{pmatrix} \otimes \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix},$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

where $\boldsymbol{\Phi}_n$ is the 4×2^n matrix of the collection of the 4×1 vectors $(1+X_k^2)^{1/2}(\delta_{2k}, \delta_{2k}X_k, \delta_{2k+1}, \delta_{2k+1}X_k)^t$ for $k \in \mathbb{G}_n = \{2^n, 2^n+1, \dots, 2^{n+1}-1\}$

$$\boldsymbol{\Phi}_n = \begin{pmatrix} \delta_{2(2^n)}\sqrt{1+X_{2^n}^2} & \dots & \delta_{2(2^{n+1}-1)}\sqrt{1+X_{2^{n+1}-1}^2} \\ \delta_{2(2^n)}X_{2^n}\sqrt{1+X_{2^n}^2} & \dots & \delta_{2(2^{n+1}-1)}X_{2^{n+1}-1}\sqrt{1+X_{2^{n+1}-1}^2} \\ \delta_{2(2^n)+1}\sqrt{1+X_{2^n}^2} & \dots & \delta_{2(2^{n+1}-1)+1}\sqrt{1+X_{2^{n+1}-1}^2} \\ \delta_{2(2^n)+1}X_{2^n}\sqrt{1+X_{2^n}^2} & \dots & \delta_{2(2^{n+1}-1)+1}X_{2^{n+1}-1}\sqrt{1+X_{2^{n+1}-1}^2} \end{pmatrix}.$$

Note that $\boldsymbol{\Phi}_\ell \boldsymbol{\Phi}_\ell^t$ and hence $\boldsymbol{\Sigma}_n$ is positive definite as soon as the X_k are not constant. The next result is again a direct consequence of Proposition 4.11.

Lemma 5.5. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 4$*, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{\boldsymbol{\Sigma}_n}{|\mathbb{T}_n^*|} = \boldsymbol{\Sigma}\mathbb{1}_{\overline{\mathcal{E}}} = \begin{pmatrix} \boldsymbol{\Sigma}^0 & 0 \\ 0 & \boldsymbol{\Sigma}^1 \end{pmatrix} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.,$$

where $\boldsymbol{\Sigma}^i$ *is the* 2×2 *matrix* $\boldsymbol{\Sigma}^i = \begin{pmatrix} \ell_i(0)+\ell_i(2) & \ell_i(1)+\ell_i(3) \\ \ell_i(1)+\ell_i(3) & \ell_i(2)+\ell_i(4) \end{pmatrix}$*. In addition,* $\boldsymbol{\Sigma}$ *is positive definite.*

As a result, Assumption **(A.4)** also holds if $\kappa \geq 4$, and we now turn to Assumption **(A.5)**.

Lemma 5.6. *Under assumptions* **(H.1–5)**, *for all* $\alpha > \max\{2\sigma_\varepsilon^2, 2\sigma_\eta^2, \mu^0, \mu^1, \nu\}$ *and for all* n*, the following inequality holds* $\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Sigma}_n^{-1} \leq \alpha(\boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_n^{-1})$*, where*

$$\mu^i = \frac{1}{2}\left(\sigma_\varepsilon^2 + \sigma_\eta^2 + \left((\sigma_\varepsilon^2 - \sigma_\eta^2)^2 + 4\rho_{ii}^2\right)^{\frac{1}{2}}\right), \ \nu = \sigma_\varepsilon^2 + \sigma_\eta^2 + \left((\sigma_\varepsilon^2 - \sigma_\eta^2)^2 + (\rho_{00}+\rho_{11})^2\right)^{\frac{1}{2}}.$$

Proof. We first prove that for all such α, $\boldsymbol{\Gamma}_n \leq \alpha\boldsymbol{\Phi}_n\boldsymbol{\Phi}_n^t$ holds. For all $k \in \mathbb{G}_n$, let $D_k^0 = \alpha(1+X_k^2) - (\sigma_\varepsilon^2 + 2X_k\rho_{00} + X_k^2\sigma_\eta^2)$, $D_k^1 = \alpha(1+X_k^2) - (\sigma_\varepsilon^2 + 2X_k\rho_{11} + X_k^2\sigma_\eta^2)$ and $D_k^{01} = \rho_\varepsilon^2 + 2X_k\rho + X_k^2\rho_\eta^2$ be the coefficients of $\alpha\boldsymbol{\Phi}_n\boldsymbol{\Phi}_n^t - \boldsymbol{\Gamma}_n$ up to the sum over $\mathbb{G}_n$. We first need to prove that $D_k^i > 0$ for all k. One can rewrite $D_k^i = \alpha - \sigma_\varepsilon^2 - 2\rho_{ii}X_k + (\alpha - \sigma_\eta^2)X_k^2$, so that it is sufficient to prove that this second order polynomial in X_k has no real root, as both its terms of degree 0 and 2 are positive by assumption. Its discriminant in a function of α given by

$$\Delta(\alpha) = -(\alpha^2 - \alpha(\sigma_\varepsilon^2 + \sigma_\eta^2) + \sigma_\varepsilon^2\sigma_\eta^2 - \rho_{ii}^2).$$

This discriminant $\Delta(\alpha)$ is again a second order polynomial in α. Therefore it is negative as soon as α is larger than its largest root μ^i. Second, we want to prove that $(D_k^{01})^2 \leq D_k^0 D_k^1$. The Cauchy–Schwarz inequality yields

$$(D_k^{01})^2 \leq (\sigma_\varepsilon^2 + 2X_k\rho_{00} + X_k^2\sigma_\eta^2)(\sigma_\varepsilon^2 + 2X_k\rho_{11} + X_k^2\sigma_\eta^2).$$

Hence, one just has to check that

$$(\sigma_\varepsilon^2 + 2X_k\rho_{00} + X_k^2\sigma_\eta^2)(\sigma_\varepsilon^2 + 2X_k\rho_{11} + X_k^2\sigma_\eta^2) \leq D_k^0 D_k^1,$$

which boils down to proving that the second order polynomial

$$\alpha - 2\sigma_\varepsilon^2 - 2(\rho_{00}+\rho_{11})X_k + (\alpha - 2\sigma_\eta^2)X_k^2$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

is non negative. Similar arguments as above yield that the preceding polynomial is nonnegative as soon as $\alpha > \max\{2\sigma_\varepsilon^2, 2\sigma_\eta^2, \nu\}$. Thus, for $u = (u_1, u_2, u_3, u_4)^t \in \mathbb{R}^4$ the following lower bound holds

$$\begin{aligned} u^t(\alpha \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^t - \boldsymbol{\Gamma}_n)u = \sum_{k \in \mathbb{G}_n} & \left(u_1 \delta_{2k} (D_k^0)^{1/2} + u_2 \delta_{2k} X_k (D_k^0)^{1/2} - u_3 \delta_{2k+1} \frac{D_k}{(D_k^0)^{1/2}} - u_4 \delta_{2k+1} X_k \frac{D_k}{(D_k^0)^{1/2}} \right)^2 \\ & + \left(u_3 \delta_{2k+1} \left(D_k^1 - \frac{(D_k^{01})^2}{D_k^0} \right)^{1/2} + u_4 \delta_{2k+1} X_k \left(D_k^1 - \frac{(D_k^{01})^2}{D_k^0} \right)^{1/2} \right)^2 \\ & \geq 0, \end{aligned}$$

hence $\boldsymbol{\Gamma}_n \leq \alpha \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^t$. To obtain the expected result, we use Riccati equation (see *e.g.* Lem. B.1 of [3]) and the definition of $\boldsymbol{\Sigma}_n$ to obtain

$$\boldsymbol{\Sigma}_n^{-1} = \boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1} \boldsymbol{\Phi}_n^t \boldsymbol{\Sigma}_{n-1}^{-1}, \tag{5.8}$$

where $\boldsymbol{l}_n = \boldsymbol{\Phi}_n^t \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n$. By multiplying both sides by $\boldsymbol{\Phi}_n$, we obtain

$$\begin{aligned} \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\Phi}_n &= \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n - \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1} \boldsymbol{l}_n, \\ &= \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n - \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1} (\mathbf{I}_{2^n} + \boldsymbol{l}_n - \mathbf{I}_{2^n}), \\ &= \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1}. \end{aligned}$$

In particular, as $\boldsymbol{l}_n$ is positive definite, one obtains

$$\boldsymbol{\Sigma}_n^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{1/2} = \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1/2}.$$

Taking the square of each side of the above equation then yields

$$\boldsymbol{\Sigma}_n^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n) \boldsymbol{\Phi}_n^t \boldsymbol{\Sigma}_n^{-1} = \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Phi}_n (\mathbf{I}_{2^n} + \boldsymbol{l}_n)^{-1} \boldsymbol{\Phi}_n^t \boldsymbol{\Sigma}_{n-1}^{-1} = \boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_n^{-1},$$

by equation (5.8). As $\mathbf{I}_{2^n} + \boldsymbol{l}_n \geq \mathbf{I}_{2^n}$ in the sense of positive semi-definite matrices, one obtains $\boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_n^{-1} \geq \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^t \boldsymbol{\Sigma}_n^{-1}$. This inequality together with $\boldsymbol{\Gamma}_n \leq \alpha \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^t$ yield the result. □

Lemma 5.7. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 4$, *for* $i \in \{0, 1\}$ *and* $q \in \{0, 1\}$, *one obtains*

$$\sup_n \left\{ m^{-2n} \mathbb{E} \left[\left(\sum_{k \in \mathbb{G}_n^*} X_k^q \epsilon_{2k+i} \right)^4 \mid \mathcal{F}_n^{\mathcal{O}} \right] \right\} < \infty \qquad a.s.$$

Proof. The following inequality is easily proved

$$\begin{aligned} m^{-2n} \mathbb{E} \left[\left(\sum_{k \in \mathbb{G}_n^*} X_k^q \epsilon_{2k+i} \right)^4 \mid \mathcal{F}_n^{\mathcal{O}} \right] \leq & \, C \left(\frac{1}{m^n} \sum_{k \in \mathbb{G}_n^*} \delta_{2k+i} X_k^{2q} (1 + X_k + X_k^2) \right)^2 \\ & + C \frac{1}{m^{2n}} \sum_{k \in \mathbb{G}_n^*} \delta_{2k+i} X_k^{4q} (1 + X_k + X_k^2 + X^3 + X^4), \end{aligned}$$

where C is a constant depending only on the moments of $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$ up to order 4. The result follows from Proposition 4.11. □

We have now proved that assumptions **(A.1−6)** of Theorem 5.1 hold for the martingale $(\boldsymbol{M}_n)$ and the sequence of positive definite matrices $(\boldsymbol{\Xi}_n) = (\boldsymbol{\Sigma}_n)$, thus we obtain the following result.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proposition 5.8. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 4$*, one obtains*

$$\boldsymbol{M}_n^t \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{M}_n = \mathcal{O}(n), \quad \text{and} \quad \|\boldsymbol{M}_n\|^2 = \mathcal{O}(nm^n) \quad a.s.$$

In addition, for all $\delta > 1/2$, $\|\boldsymbol{M}_n\|^2 = o(n^\delta m^n)$ *a.s. and*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t \boldsymbol{\Sigma}_{\ell-1}^{-1} \boldsymbol{M}_\ell = \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Sigma}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}}, \quad a.s.$$

Now recall that $\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta} = \boldsymbol{S}_{n-1}^{-1} \boldsymbol{M}_n$. One then readily obtains Theorem 3.2.

Proof of Theorem 3.2. As $\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta} = \boldsymbol{S}_{n-1}^{-1} \boldsymbol{M}_n$, one obtains

$$\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\|^2 = \boldsymbol{M}_n^t \boldsymbol{S}_{n-1}^{-2} \boldsymbol{M}_n \leq \|\boldsymbol{M}_n\|^2 \lambda_{\max}(\boldsymbol{S}_{n-1}^{-2}),$$

where $\lambda_{\max}(\boldsymbol{S}_{n-1}^{-2})$ denotes the highest eigenvalue of matrix $\boldsymbol{S}_{n-1}^{-2}$. We use Proposition 4.14 to conclude that $\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\|^2 = o(n^\delta m^{-n})$ a.s. For the quadratic strong law, Proposition 5.8 yields

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{\ell=1}^{n} (\widehat{\boldsymbol{\theta}}_\ell - \boldsymbol{\theta})^t \boldsymbol{S}_{\ell-1} \boldsymbol{\Sigma}_{\ell-1}^{-1} \boldsymbol{S}_{\ell-1} (\widehat{\boldsymbol{\theta}}_\ell - \boldsymbol{\theta}) = \mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{\Sigma}^{-1}) \mathbb{1}_{\overline{\mathcal{E}}} \qquad \text{a.s.}$$

and the result is obtained by using Proposition 4.14 and Lemma 5.5. A similar argument as in the Proof of Lemma 5.2 is used to replace $\boldsymbol{S}_{\ell-1} \boldsymbol{\Sigma}_{\ell-1}^{-1} \boldsymbol{S}_{\ell-1}$ by its equivalent $|\mathbb{T}_{\ell-1}^*|^{-1} \boldsymbol{S}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}$. □

5.3. Rate of convergence for $\widehat{\boldsymbol{\sigma}}_n$

We proceed in two steps. Recall that

$$\widehat{\boldsymbol{\sigma}}_n = \boldsymbol{U}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \left(\widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2, 2X_k\widehat{\epsilon}_{2k}^2, 2X_k\widehat{\epsilon}_{2k+1}^2, X_k^2(\widehat{\epsilon}_{2k}^2 + \widehat{\epsilon}_{2k+1}^2)\right)^t,$$

is our estimator of $\boldsymbol{\sigma} = (\sigma_\varepsilon^2, \rho_{00}, \rho_{11}, \sigma_\eta^2)^t$, and

$$\boldsymbol{\sigma}_n = \boldsymbol{U}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \left(\epsilon_{2k}^2 + \epsilon_{2k+1}^2, 2X_k\epsilon_{2k}^2, 2X_k\epsilon_{2k+1}^2, X_k^2(\epsilon_{2k}^2 + \epsilon_{2k+1}^2)\right)^t.$$

Our first step is to prove the convergence of $\boldsymbol{\sigma}_n$ to $\boldsymbol{\sigma}$. The second step is the convergence of $\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n$ with a convergence rate.

5.3.1. Convergence of $\boldsymbol{\sigma}_n$

The convergence of $\boldsymbol{\sigma}_n$ to $\boldsymbol{\sigma}$ is directly obtained using the usual law of large numbers for square-integrable vector-valued martingales. Note that one could obtain a convergence rate under stronger moment assumptions using Theorem 5.1.

Lemma 5.9. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 8$*, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{\sigma}_n = \boldsymbol{\sigma} \mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proof. Set

$$\boldsymbol{M}_n^{\boldsymbol{\sigma}} = \boldsymbol{U}_{n-1}(\boldsymbol{\sigma}_n - \boldsymbol{\sigma}) = \sum_{\ell=1}^{n-1} \sum_{k \in \mathbb{G}_\ell} \begin{pmatrix} \epsilon_{2k}^2 + \epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k}^2 + \epsilon_{2k+1}^2 \mid \mathcal{F}_\ell^{\mathcal{O}}] \\ 2X_k(\epsilon_{2k}^2 - \mathbb{E}[\epsilon_{2k}^2 \mid \mathcal{F}_\ell^{\mathcal{O}}]) \\ 2X_k(\epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k+1}^2 \mid \mathcal{F}_\ell^{\mathcal{O}}]) \\ X_k^2(\epsilon_{2k}^2 + \epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k}^2 + \epsilon_{2k+1}^2 \mid \mathcal{F}_\ell^{\mathcal{O}}]) \end{pmatrix}.$$

Hence, $(\boldsymbol{M}_n^{\boldsymbol{\sigma}})$ is a square-integrable $(\mathcal{F}_n^{\mathcal{O}})$-martingale. One can compute all the entries of its predictable quadratic variation and prove that they all equal a constant (depending on the moments of $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$ up to order 4) multiplied by $\sum \delta_{2k+i} X_k^q$ or $\sum \delta_{2k}\delta_{2k+1} X_k^q$ with $q \leq 8$. Hence, the laws of large numbers given in Proposition 4.11 ensure that $m^{-n} < \boldsymbol{M}^{\boldsymbol{\sigma}} >_n$ converges almost surely to a constant matrix on the non-extinction set $\overline{\mathcal{E}}$. The standard law of large numbers for square-integrable martingales then implies that $(\boldsymbol{M}_n^{\boldsymbol{\sigma}}) = o(m^n)$ a.s. Besides, $m^{-n}\boldsymbol{U}_n$ also converges to a fixed matrix on the non-extinction set $\overline{\mathcal{E}}$ by Proposition 4.14. Therefore $\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}(\boldsymbol{\sigma}_n - \boldsymbol{\sigma}) = \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{U}_{n-1}^{-1} \boldsymbol{M}_n^{\boldsymbol{\sigma}}$ tends to 0 a.s. when n tends to infinity. □

5.3.2. Convergence of $\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n$

We now turn to the convergence of $\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n$. One can rewrite $\boldsymbol{U}_{n-1}(\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n)$ as

$$\boldsymbol{U}_{n-1}(\widehat{\boldsymbol{\sigma}}_n - \boldsymbol{\sigma}_n) = \boldsymbol{P}_n^{\boldsymbol{\sigma}} + 2\boldsymbol{R}_n^{\boldsymbol{\sigma}}, \tag{5.9}$$

with

$$\boldsymbol{P}_n^{\boldsymbol{\sigma}} = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} (\widehat{\epsilon}_{2k} - \epsilon_{2k})^2 + (\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})^2 \\ 2X_k(\widehat{\epsilon}_{2k} - \epsilon_{2k})^2 \\ 2X_k(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})^2 \\ X_k^2\big((\widehat{\epsilon}_{2k} - \epsilon_{2k})^2 + (\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})^2\big) \end{pmatrix},$$

and

$$\boldsymbol{R}_n^{\boldsymbol{\sigma}} = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} \epsilon_{2k}(\widehat{\epsilon}_{2k} - \epsilon_{2k}) + \epsilon_{2k+1}(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1}) \\ 2X_k\epsilon_{2k}(\widehat{\epsilon}_{2k} - \epsilon_{2k}) \\ 2X_k\epsilon_{2k+1}(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1}) \\ X_k^2\big(\epsilon_{2k}(\widehat{\epsilon}_{2k} - \epsilon_{2k}) + \epsilon_{2k+1}(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})\big) \end{pmatrix}.$$

We are going to study separately the asymptotic properties of $\boldsymbol{P}_n^{\boldsymbol{\sigma}}$ and $\boldsymbol{R}_n^{\boldsymbol{\sigma}}$.

Lemma 5.10. *Under assumptions* **(H.1–5)** *and if $\kappa \geq 4$, one obtains*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{k \in \mathbb{T}_n} (\widehat{\epsilon}_{2k} - \epsilon_{2k})^2 = q_0(0)\mathbb{1}_{\overline{\mathcal{E}}} = (m-1)\mathrm{tr}(\boldsymbol{\Gamma}^0(\boldsymbol{S}^0)^{-1})\mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

where $\boldsymbol{\Gamma}^0$ is defined in Lemma 5.4 and $\boldsymbol{S}^0$ in Proposition 4.14.

Proof. We are going to apply Theorem 5.1 to the first two entries of the martingale $(\boldsymbol{M}_n)$. Indeed, let $\boldsymbol{M}_n^0$ be the 2-component vector corresponding to the first two entries of $\boldsymbol{M}_n$

$$\boldsymbol{M}_n^0 = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} \epsilon_{2k} \\ X_k\epsilon_{2k} \end{pmatrix}.$$

Let $\boldsymbol{\theta}^0 = (a, b)^t$, $\widehat{\boldsymbol{\theta}}_n^0 = (\widehat{a}_n, \widehat{b}_n)^t$. Clearly, one has $(\widehat{\boldsymbol{\theta}}_n^0 - \boldsymbol{\theta}^0) = (\boldsymbol{S}_{n-1}^0)^{-1}\boldsymbol{M}_n^0$, therefore one obtains

$$\begin{aligned} \sum_{k \in \mathbb{G}_n} (\widehat{\epsilon}_{2k} - \epsilon_{2k})^2 &= \sum_{k \in \mathbb{G}_n} \delta_{2k} (\widehat{\boldsymbol{\theta}}_n^0 - \boldsymbol{\theta}^0)^t \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix} (\widehat{\boldsymbol{\theta}}_n^0 - \boldsymbol{\theta}^0) \\ &= (\boldsymbol{M}_n^0)^t (\boldsymbol{S}_{n-1}^0)^{-1} (\boldsymbol{S}_n^0 - \boldsymbol{S}_{n-1}^0)(\boldsymbol{S}_{n-1}^0)^{-1} \boldsymbol{M}_n^0. \end{aligned}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Set also $\boldsymbol{\Sigma}_n^0$ to be the 2×2 matrix defined by

$$\boldsymbol{\Sigma}_n^0 = \sum_{k\in\mathbb{T}_n}(1+X_k^2)\delta_{2k}\begin{pmatrix}1 & X_k\\ X_k & X_k^2\end{pmatrix}.$$

Thus, Proposition 4.14 and Lemma 5.5 yield

$$\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}(\boldsymbol{\Sigma}_n^0)^{1/2}(\boldsymbol{S}_n^0)^{-1}(\boldsymbol{S}_{n+1}^0-\boldsymbol{S}_n^0)(\boldsymbol{S}_n^0)^{-1}(\boldsymbol{\Sigma}_n^0)^{1/2}=\boldsymbol{\Delta}^{\mathbf{0}}\mathbb{1}_{\overline{\mathcal{E}}}\quad a.s.$$

where $\boldsymbol{\Delta}^{\mathbf{0}} = (m-1)(\boldsymbol{\Sigma}^0)^{1/2}(\boldsymbol{S}^0)^{-1}(\boldsymbol{\Sigma}^0)^{1/2}$. Note that $\boldsymbol{\Delta}^{\mathbf{0}}$ is positive definite and the matrices $\boldsymbol{\Sigma}_n^0$, $\boldsymbol{S}_n^0$, $\boldsymbol{\Sigma}^0$, $\boldsymbol{S}^0$ and $\boldsymbol{\Delta}^{\mathbf{0}}$ commute. We now use Theorem 5.1 for the martingale $(\boldsymbol{M}_n^0)$, with the sequence $\big(\Xi_n = (\boldsymbol{\Delta}^0)^{-1/2}\boldsymbol{\Sigma}_n^0(\boldsymbol{\Delta}^0)^{-1/2}\big)$. As $\boldsymbol{\Delta}^0$ is a fixed positive definite matrix, it is clear that all the assumptions of Theorem 5.1 hold, as in Section 5.2. Thus one obtains the a.s. limit

$$\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{\ell=1}^{n}(\boldsymbol{M}_\ell^0)^t((\boldsymbol{\Delta}^0)^{-1/2}\boldsymbol{\Sigma}_{\ell-1}^0(\boldsymbol{\Delta}^0)^{-1/2})^{-1}\boldsymbol{M}_\ell^0=\operatorname{tr}(\boldsymbol{\Gamma}^0(\boldsymbol{\Sigma}^0)^{-1}\boldsymbol{\Delta}^0)\mathbb{1}_{\overline{\mathcal{E}}} = (m-1)\operatorname{tr}(\boldsymbol{\Gamma}^0(\boldsymbol{S}^0)^{-1}).$$

Finally, a similar argument as in the Proof of Lemma 5.2 is used to replace $\boldsymbol{\Delta}^0$ by its asymptotic equivalent $(\boldsymbol{\Sigma}_n^0)^{1/2}(\boldsymbol{S}_n^0)^{-1}(\boldsymbol{S}_{n+1}^0-\boldsymbol{S}_n^0)(\boldsymbol{S}_n^0)^{-1}(\boldsymbol{\Sigma}_n^0)^{1/2}$ to obtain

$$\begin{aligned}\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{k\in\mathbb{T}_n}(\widehat{\epsilon}_{2k}-\epsilon_{2k})^2 &= \lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{\ell=1}^{n}(\boldsymbol{M}_\ell^0)^t(\boldsymbol{\Sigma}_{\ell-1}^0)^{-1/2}\boldsymbol{\Delta}^0(\boldsymbol{\Sigma}_{\ell-1}^0)^{-1/2}\boldsymbol{M}_\ell^0\\ &= \lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{\ell=1}^{n}(\boldsymbol{M}_\ell^0)^t\big((\boldsymbol{\Delta}^0)^{-1/2}\boldsymbol{\Sigma}_{\ell-1}^0(\boldsymbol{\Delta}^0)^{-1/2}\big)^{-1}\boldsymbol{M}_\ell^0,\end{aligned}$$

hence the result. □

A similar proof yields the following results for odd indices.

Lemma 5.11. *Under assumptions* **(H.1–5)** *and if* $\kappa\geq 4$, *the following convergence holds*

$$\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{k\in\mathbb{T}_n}(\widehat{\epsilon}_{2k+1}-\epsilon_{2k+1})^2=q_1(0)\mathbb{1}_{\overline{\mathcal{E}}}=(m-1)\operatorname{tr}(\boldsymbol{\Gamma}^1(\boldsymbol{S}^1)^{-1})\mathbb{1}_{\overline{\mathcal{E}}}\quad a.s.$$

where $\boldsymbol{\Gamma}^1$ *is defined in Lemma 5.4 and* $\boldsymbol{S}^1$ *in Proposition 4.14.*

The proof of Lemma 5.10 can also be adapted to obtain the two following results.

Lemma 5.12. *Under assumptions* **(H.1–5)** *and if* $\kappa\geq 4$, *for all* $i\in\{0,1\}$, *the almost sure convergence holds*

$$\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\sum_{k\in\mathbb{T}_n}X_k(\widehat{\epsilon}_{2k+i}-\epsilon_{2k+i})^2=q_i(1)\mathbb{1}_{\overline{\mathcal{E}}}=(m-1)\operatorname{tr}(\boldsymbol{\Gamma}^i(\boldsymbol{S}^i)^{-2}\boldsymbol{T}^i)\mathbb{1}_{\overline{\mathcal{E}}},$$

where $\boldsymbol{T}^i$ *is the* 2×2 *matrix defined by* $\boldsymbol{T}^i=\begin{pmatrix}\ell_i(1) & \ell_i(2)\\ \ell_i(2) & \ell_i(3)\end{pmatrix}$.

Proof. We prove the result for $i=0$, the other case being similar. With the notation of the proof of Lemma 5.10, one obtains

$$\begin{aligned}\sum_{k\in\mathbb{G}_n}X_k(\widehat{\epsilon}_{2k}-\epsilon_{2k})^2 &= \sum_{k\in\mathbb{G}_n}\delta_{2k}(\widehat{\boldsymbol{\theta}}_n^0-\boldsymbol{\theta}^0)^t\begin{pmatrix}X_k & X_k^2\\ X_k^2 & X_k^3\end{pmatrix}(\widehat{\boldsymbol{\theta}}_n^0-\boldsymbol{\theta}^0)\\ &= (\boldsymbol{M}_n^0)^t(\boldsymbol{S}_{n-1}^0)^{-1}(\boldsymbol{T}_n^0-\boldsymbol{T}_{n-1}^0)(\boldsymbol{S}_{n-1}^0)^{-1}\boldsymbol{M}_n^0,\end{aligned}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

with $\boldsymbol{T}_n^0 = \sum_{k\in\mathbb{T}_n} \delta_{2k} \begin{pmatrix} X_k & X_k^2 \\ X_k^2 & X_k^3 \end{pmatrix}$. Proposition 4.11 yields the following convergence

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} (\boldsymbol{\Sigma}_n^0)^{1/2} (\boldsymbol{S}_n^0)^{-1} (\boldsymbol{T}_{n+1}^0 - \boldsymbol{T}_n^0)(\boldsymbol{S}_n^0)^{-1} (\boldsymbol{\Sigma}_n^0)^{1/2} = \boldsymbol{\Delta}^{\mathbf{0}} \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

with a new matrix $\boldsymbol{\Delta}^{\mathbf{0}}$ defined by $\boldsymbol{\Delta}^{\mathbf{0}} = (m-1)\boldsymbol{\Sigma}^0 (\boldsymbol{S}^0)^{-2} \boldsymbol{T}^0$. Note that this new $\boldsymbol{\Delta}^{\mathbf{0}}$ is again positive definite and the matrices $\boldsymbol{\Sigma}_n^0$, $\boldsymbol{S}_n^0$, $\boldsymbol{T}_n^0$, $\boldsymbol{\Sigma}^0$, $\boldsymbol{S}^0$, $\boldsymbol{T}^0$, and $\boldsymbol{\Delta}^{\mathbf{0}}$ still commute. The end of the proof is similar to that of Lemma 5.10 with the new matrix $\boldsymbol{\Delta}^{\mathbf{0}}$. □

Lemma 5.13. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 4$*, for all* $i \in \{0,1\}$*, the almost sure convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \sum_{k\in\mathbb{T}_n} X_k^2 (\widehat{\epsilon}_{2k+i} - \epsilon_{2k+i})^2 = q_i(2) \mathbb{1}_{\overline{\mathcal{E}}} = (m-1)\mathrm{tr}(\boldsymbol{\Gamma}^i (\boldsymbol{S}^i)^{-2} \boldsymbol{W}^i) \mathbb{1}_{\overline{\mathcal{E}}},$$

where $\boldsymbol{W}^i$ *is the* 2×2 *matrix defined by* $\boldsymbol{W}^i = \begin{pmatrix} \ell_i(2) & \ell_i(3) \\ \ell_i(3) & \ell_i(4) \end{pmatrix}$.

Lemmas 5.10, 5.11, 5.12 and 5.13 give the almost sure convergence of the sequence $(\boldsymbol{P}_n^{\boldsymbol{\sigma}})$.

Lemma 5.14. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 4$*, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \boldsymbol{P}_n^{\boldsymbol{\sigma}} = \big(q_0(0) + q_1(0), 2q_0(1), 2q_1(1), q_0(2) + q_1(2)\big)^t \mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

It remains to give the limit of the sequence $(\boldsymbol{R}_n^{\boldsymbol{\sigma}})$.

Lemma 5.15. *Under assumptions* **(H.1–5)** *and if* $\kappa \geq 8$*, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \boldsymbol{R}_n^{\boldsymbol{\sigma}} = 0 \qquad a.s.$$

Proof. It is sufficient to prove that $(\boldsymbol{R}_n^{\boldsymbol{\sigma}})$ is a martingale and that its predictable quadratic variation is almost surely $\mathcal{O}(n)$. For all $k \in \mathbb{G}_n$, one obtains

$$\mathbb{E}[\epsilon_{2k}(\widehat{\epsilon}_{2k} - \epsilon_{2k}) \mid \mathcal{F}_n^{\mathcal{O}}] = \delta_{2k}\big((a - \widehat{a}_n) + (b - \widehat{b}_n)X_k\big)\mathbb{E}[\epsilon_{2k} \mid \mathcal{F}_n^{\mathcal{O}}] = 0,$$

and we have the same result for the other entries of $\boldsymbol{R}_n^{\boldsymbol{\sigma}}$. Hence, $(\boldsymbol{R}_n^{\boldsymbol{\sigma}})$ is a $(\mathcal{F}_n^{\mathcal{O}})$-martingale. It is also square-integrable. We are going to study $(\boldsymbol{R}_n^{\boldsymbol{\sigma}})$ component-wise. We give the details for the last entry, the others being treated similarly. For $i \in \{0,1\}$, set

$$Q_n^i = \sum_{\ell=1}^{n-1} (\boldsymbol{\theta}^i - \widehat{\boldsymbol{\theta}}_n^i)^t \sum_{k\in\mathbb{G}_\ell} \delta_{2k+i} \begin{pmatrix} X_k^2 \\ X_k^3 \end{pmatrix} \epsilon_{2k+i}.$$

The last entry of $\boldsymbol{R}_n^{\boldsymbol{\sigma}}$ can then be rewritten as $Q_n^0 + Q_n^1$. The processes Q_n^i are clearly $(\mathcal{F}_n^{\mathcal{O}})$-martingales with predictable quadratic variation equal to

$$\langle Q^i \rangle_n = \sum_{\ell=1}^{n-1} (\boldsymbol{M}_\ell^i)^t (\boldsymbol{S}_{\ell-1}^i)^{-1} \boldsymbol{\Delta}_\ell^i (\boldsymbol{S}_{\ell-1}^i)^{-1} \boldsymbol{M}_\ell^i,$$

with $\boldsymbol{\Delta}_n^i = \sum_{k\in\mathbb{G}_n} \delta_{2k} (\sigma_\varepsilon^2 + 2\rho_{ii} X_k + \sigma_\eta^2 X_k^2) \begin{pmatrix} X_k^4 & X_k^5 \\ X_k^5 & X_k^6 \end{pmatrix}$. Thanks to Proposition 4.11, the sequence of matrices $(\boldsymbol{\Sigma}_n^i)^{1/2} (\boldsymbol{S}_{n-1}^i)^{-1} \boldsymbol{\Delta}_n^i (\boldsymbol{S}_{n-1}^i)^{-1} (\boldsymbol{\Sigma}_n^i)^{1/2}$ converges almost surely on the non-extinction set $\overline{\mathcal{E}}$ to a fixed positive definite matrix $\boldsymbol{\Delta}^i$. We now use Theorem 5.1 along the same lines as in the proof of Lemma 5.10 to obtain that $\langle Q^i \rangle_n = \mathcal{O}(n)$, and thus $Q_n^i = o(n)$. The other entries of $(\boldsymbol{R}_n^{\boldsymbol{\sigma}})$ are dealt with similarly, yielding the result. □

Proof of Theorem 3.3. It is a direct consequence of equation (5.9), Proposition 4.14 and Lemmas 5.9, 5.14 and 5.15. □

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

5.4. Rate of convergence for $\widehat{\boldsymbol{\rho}}_n$

We proceed again in two steps. Recall that

$$\widehat{\boldsymbol{\rho}}_n = \boldsymbol{V}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \left(\widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1}, 2X_k\widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1}, X_k^2\widehat{\epsilon}_{2k}\widehat{\epsilon}_{2k+1}\right)^t,$$

is our estimator of $\boldsymbol{\rho} = (\rho_\varepsilon, \rho, \rho_\eta)^t$, and

$$\boldsymbol{\rho}_n \;=\; \boldsymbol{V}_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \left(\epsilon_{2k}\epsilon_{2k+1}, 2X_k\epsilon_{2k}\epsilon_{2k+1}, X_k^2\epsilon_{2k}\epsilon_{2k+1}\right)^t.$$

Our first step is to prove the convergence of $\boldsymbol{\rho}_n$ to $\boldsymbol{\rho}$. The second step is the convergence of $\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n$ with a convergence rate.

5.4.1. Convergence of $\boldsymbol{\rho}_n$

The convergence of $\boldsymbol{\rho}_n$ to $\boldsymbol{\rho}$ is again directly obtained using the standard law of large numbers for square-integrable vector-valued martingales. Note that one could also obtain a convergence rate under stronger moment assumptions using Theorem 5.1.

Lemma 5.16. *Under assumptions* **(H.1–5)** *and if $\kappa \geq 8$, the following convergence holds*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\boldsymbol{\rho}_n = \boldsymbol{\rho}\mathbb{1}_{\overline{\mathcal{E}}} \qquad a.s.$$

Proof. Set

$$\boldsymbol{M}_n^{\boldsymbol{\rho}} = \boldsymbol{V}_{n-1}(\boldsymbol{\rho}_n - \boldsymbol{\rho}) = \sum_{\ell=1}^{n-1}\sum_{k\in\mathbb{G}_\ell} \begin{pmatrix} \epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_\ell^{\mathcal{O}}] \\ 2X_k\big(\epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_\ell^{\mathcal{O}}]\big) \\ X_k^2\big(\epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_\ell^{\mathcal{O}}]\big) \end{pmatrix}.$$

Hence, $(\boldsymbol{M}_n^{\boldsymbol{\rho}})$ is a square-integrable $(\mathcal{F}_n^{\mathcal{O}})$-martingale. One can compute all the entries of its predictable quadratic variation and prove that they all equal a constant (depending on the moments of $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$ up to order 4) multiplied by $\sum \delta_{2k}\delta_{2k+1}X_k^q$ with $q \leq 8$. Therefore, the laws of large numbers given in Proposition 4.11 ensure that $m^{-n}\langle \boldsymbol{M}^{\boldsymbol{\rho}}\rangle_n$ converges almost surely to a constant matrix on the non-extinction set $\overline{\mathcal{E}}$. The standard law of large numbers for square-integrable martingales then implies that $(\boldsymbol{M}_n^{\boldsymbol{\rho}}) = o(m^n)$ a.s. Besides, $m^{-n}\boldsymbol{V}_n$ also converges to a fixed matrix on the non-extinction set $\overline{\mathcal{E}}$ by Proposition 4.14. Thus $\boldsymbol{\rho}_n - \boldsymbol{\rho} = \boldsymbol{V}_{n-1}^{-1}\boldsymbol{M}_n^{\boldsymbol{\rho}}$ tends to 0 a.s. on $\overline{\mathcal{E}}$ when n tends to infinity. □

5.4.2. Convergence of $\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n$

We now turn to the convergence of $\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n$. We follow the same steps as in Section 5.3.2. One can rewrite $\boldsymbol{V}_{n-1}(\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n)$ as

$$\begin{aligned} \boldsymbol{V}_{n-1}(\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_n) &= \boldsymbol{P}_n^{\boldsymbol{\rho}} + \boldsymbol{R}_n^{\boldsymbol{\rho}}, \quad \text{with} \\ \boldsymbol{P}_n^{\boldsymbol{\rho}} &= \sum_{k\in\mathbb{T}_{n-1}} (\widehat{\epsilon}_{2k} - \epsilon_{2k})(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})\left(1, 2X_k, X_k^2\right)^t, \quad \text{and} \\ \boldsymbol{R}_n^{\boldsymbol{\rho}} &= \sum_{k\in\mathbb{T}_{n-1}} \left(\epsilon_{2k+1}(\widehat{\epsilon}_{2k} - \epsilon_{2k}) + \epsilon_{2k}(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1})\right)\left(1, 2X_k, X_k^2\right)^t. \end{aligned} \tag{5.10}$$

We are going to study separately the asymptotic properties of $\boldsymbol{P}_n^{\boldsymbol{\rho}}$ and $\boldsymbol{R}_n^{\boldsymbol{\rho}}$. The limit of $\boldsymbol{R}_n^{\boldsymbol{\rho}}$ is obtained as in Lemma 5.15.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Lemma 5.17. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 8$, *one obtains*

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \frac{1}{n} \boldsymbol{R}_n^{\boldsymbol{\rho}} = 0 \qquad a.s.$$

Lemma 5.18. *Under assumptions* **(H.1−5)** *and if* $\kappa \geq 4$, *the following almost sure convergence holds*

$$\lim_{n\to\infty} \frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{n} \sum_{k\in\mathbb{T}_n} (\widehat{\epsilon}_{2k} - \epsilon_{2k})(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1}) = q_{01}(0)\mathbb{1}_{\overline{\mathcal{E}}} = \frac{m-1}{2}\mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{S}^{-2}\boldsymbol{J}^{01})\mathbb{1}_{\overline{\mathcal{E}}}$$

where $\boldsymbol{J}^{01} = \begin{pmatrix} 0 & \boldsymbol{S}^{01} \\ \boldsymbol{S}^{01} & 0 \end{pmatrix}$, *and* $\boldsymbol{S}^{01} = \begin{pmatrix} \ell_{01}(0) & \ell_{01}(1) \\ \ell_{01}(1) & \ell_{01}(2) \end{pmatrix}$.

Proof. First, notice that for all $k \in \mathbb{G}_n$, the following decomposition holds

$$2(\widehat{\epsilon}_{2k} - \epsilon_{2k})(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1}) = \delta_{2k}\delta_{2k+1}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})^t \begin{pmatrix} 0 & 0 & 1 & X_k \\ 0 & 0 & X_k & X_k^2 \\ 1 & X_k & 0 & 0 \\ X_k & X_k^2 & 0 & 0 \end{pmatrix} (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}).$$

Hence, one obtains

$$2\sum_{k\in\mathbb{T}_n} (\widehat{\epsilon}_{2k} - \epsilon_{2k})(\widehat{\epsilon}_{2k+1} - \epsilon_{2k+1}) = \sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t \boldsymbol{S}_{\ell-1}^{-1}(\boldsymbol{J}_\ell^{01} - \boldsymbol{J}_{\ell-1}^{01})\boldsymbol{S}_{\ell-1}^{-1}\boldsymbol{M}_\ell,$$

with $\boldsymbol{J}_n^{01} = \begin{pmatrix} 0 & \boldsymbol{S}_n^{01} \\ \boldsymbol{S}_n^{01} & 0 \end{pmatrix}$ and $\boldsymbol{S}_n^{01} = \sum_{k\in\mathbb{T}_n} 2\delta_{2k}\delta_{2k+1} \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix}$. Set $\boldsymbol{\Delta}_n = \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{S}_n^{-1}(\boldsymbol{J}_{n+1}^{01} - \boldsymbol{J}_n^{01})\boldsymbol{S}_n^{-1}\boldsymbol{\Sigma}_n^{1/2}$. Proposition 4.11 yields

$$\lim_{n\to\infty} \mathbb{1}_{\{|\mathbb{G}_n^*|>0\}} \boldsymbol{\Delta}_n = \boldsymbol{\Delta}\mathbb{1}_{\overline{\mathcal{E}}} = (m-1)\boldsymbol{\Sigma}^{1/2}\boldsymbol{S}^{-1}\boldsymbol{J}^{01}\boldsymbol{S}^{-1}\boldsymbol{\Sigma}^{1/2}\mathbb{1}_{\overline{\mathcal{E}}} \quad a.s.$$

Hence, as in the proof of Lemma 5.10, it is sufficient to study the convergence of $\sum_{\ell=1}^{n} \boldsymbol{M}_\ell^t \boldsymbol{\Sigma}_{\ell-1}^{-1/2}\boldsymbol{\Delta}\boldsymbol{\Sigma}_{\ell-1}^{-1/2}\boldsymbol{M}_\ell$. To this end, we apply Theorem 5.1 to the martingale $(\boldsymbol{M}_n)$ and with the sequence of matrices $\boldsymbol{\Xi}_n = \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_n^{1/2}$. As we have seen before, assumptions **(A.1−3)** and **(A.6)** hold. Assumption **(A.4)** is a direct consequence of Lemma 5.5. We will not investigate assumption **(A.5)** but directly prove Lemma 5.3, a part of Theorem 5.1, in our specific context that is

$$\mathcal{B}'_{n+1} = 2\sum_{k=1}^{n} \boldsymbol{M}_k^t \boldsymbol{\Xi}_k^{-1} \Delta\boldsymbol{M}_{k+1} = o(n).$$

Note that the matrix $\boldsymbol{J}^{01}$ has a special property, namely

$$(\boldsymbol{J}^{01})^2 = \begin{pmatrix} (\boldsymbol{S}^{01})^2 & 0 \\ 0 & (\boldsymbol{S}^{01})^2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}^{01} & 0 \\ 0 & \boldsymbol{S}^{01} \end{pmatrix}^2 = (\boldsymbol{I}^{01})^2,$$

so that although $\boldsymbol{J}^{01}$ is not positive definite, $\boldsymbol{I}^{01}$ is. As a result, as the matrices $\boldsymbol{J}_n^{01}$, $\boldsymbol{\Sigma}_n$ and $\boldsymbol{S}_n$ commute, one has $\boldsymbol{\Delta}^2 = (\boldsymbol{\Delta}')^2$ with the positive definite matrix $\boldsymbol{\Delta}' = (m-1)\boldsymbol{\Sigma}^{1/2}\boldsymbol{S}^{-1}\boldsymbol{I}^{01}\boldsymbol{S}^{-1}\boldsymbol{\Sigma}^{1/2}$. We have seen that the coefficient α given by Lemma 5.6 satisfies $\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Sigma}_n^{-1} \leq \alpha(\boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_n^{-1})$. In view of the property of $\boldsymbol{J}^{01}$ and the fact that the matrices $\boldsymbol{\Delta}$, $\boldsymbol{\Delta}'$, $\boldsymbol{\Sigma}_n$ and $\boldsymbol{\Gamma}_n$ commute, we obtain

$$\begin{aligned}\left(\boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_n^{1/2}\right)^{-1}\boldsymbol{\Gamma}_n\left(\boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}_n^{1/2}\right)^{-1} &= \boldsymbol{\Delta}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\Delta}' \\ &\leq \alpha\boldsymbol{\Delta}'(\boldsymbol{\Sigma}_{n-1}^{-1} - \boldsymbol{\Sigma}_n^{-1})\boldsymbol{\Delta}'.\end{aligned}$$

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Thanks to Lemma 5.5, we have the convergence

$$\lim_{n\to\infty}\frac{\lambda_{\max}(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})}{\lambda_{\min}(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})}=\frac{\lambda_{\max}(\boldsymbol{\Sigma}^{-1})}{\lambda_{\min}(\boldsymbol{\Sigma}^{-1})}>0.$$

So there exists $n_0>0$ and $\beta>0$ such that

$$\beta\lambda_{\min}(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})\geq\lambda_{\max}(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})\lambda_{\max}((\boldsymbol{\Delta}')^2)\text{ for }n\geq n_0,$$

which implies $\boldsymbol{\Delta}'(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})\boldsymbol{\Delta}'\leq\beta(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})$, and

$$\mathbb{E}[\Delta\mathcal{B}_{n+1}'^2|\mathcal{F}_n^{\mathcal{O}}]=4\boldsymbol{M}_n^t\boldsymbol{\Xi}_n^{-1}\boldsymbol{\Gamma}_n\boldsymbol{\Xi}_n^{-1}\boldsymbol{M}_n\leq\alpha\beta\boldsymbol{M}_n^t(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})\boldsymbol{M}_n\text{ for }n\geq n_0.$$

$\boldsymbol{M}_n^t(\boldsymbol{\Sigma}_{n-1}^{-1}-\boldsymbol{\Sigma}_n^{-1})\boldsymbol{M}_n$ is the increment of $\mathcal{A}_n$ defined in equation 5.1 when $\Xi_n=\Sigma_n$. By the arguments used in the proof of Proposition 5.8, $\sum_{\ell=1}^n\boldsymbol{M}_\ell^t(\boldsymbol{\Sigma}_{\ell-1}^{-1}-\boldsymbol{\Sigma}_\ell^{-1})\boldsymbol{M}_\ell=\mathcal{O}(n)$ which implies $\mathcal{B}_{n+1}'=o(n)$. We can then apply Theorem 5.1 which immediately yields the result. □

Similarly, one obtains the two following convergences.

Lemma 5.19. *Under assumptions* **(H.1−5)** *and if* $\kappa\geq4$*, the following almost sure convergence holds*

$$\lim_{n\to\infty}\frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}}{n}\sum_{k\in\mathbb{T}_n}X_k(\widehat{\epsilon}_{2k}-\epsilon_{2k})(\widehat{\epsilon}_{2k+1}-\epsilon_{2k+1})=q_{01}(1)\mathbb{1}_{\overline{\mathcal{E}}}=\frac{m-1}{2}\mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{S}^{-2}\boldsymbol{K}^{01})\mathbb{1}_{\overline{\mathcal{E}}}$$

where $\boldsymbol{K}^{01}=\begin{pmatrix}0&\boldsymbol{T}^{01}\\\boldsymbol{T}^{01}&0\end{pmatrix}$ *and* $\boldsymbol{T}^{01}=\begin{pmatrix}\ell_{01}(1)&\ell_{01}(2)\\\ell_{01}(2)&\ell_{01}(3)\end{pmatrix}$.

Lemma 5.20. *Under assumptions* **(H.1−5)** *and if* $\kappa\geq4$*, the following almost sure convergence holds*

$$\lim_{n\to\infty}\frac{\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}1}{n}\sum_{k\in\mathbb{T}_n}X_k^2(\widehat{\epsilon}_{2k}-\epsilon_{2k})(\widehat{\epsilon}_{2k+1}-\epsilon_{2k+1})=q_{01}(2)\mathbb{1}_{\overline{\mathcal{E}}}=\frac{m-1}{2}\mathrm{tr}(\boldsymbol{\Gamma}\boldsymbol{S}^{-2}\boldsymbol{L}^{01})\mathbb{1}_{\overline{\mathcal{E}}}$$

where $\boldsymbol{L}^{01}=\begin{pmatrix}0&\boldsymbol{W}^{01}\\\boldsymbol{W}^{01}&0\end{pmatrix}$, *and* $\boldsymbol{W}^{01}=\begin{pmatrix}\ell_{01}(2)&\ell_{01}(3)\\\ell_{01}(3)&\ell_{01}(4)\end{pmatrix}$.

Thus, one obtains the following limit.

Lemma 5.21. *Under assumptions* **(H.1−5)** *and if* $\kappa\geq4$*, the following almost sure convergence holds*

$$\lim_{n\to\infty}\mathbb{1}_{\{|\mathbb{G}_n^*|>0\}}\frac{1}{n}\boldsymbol{P}_n^{\boldsymbol{\sigma}}=\big(q_{01}(0),2q_{01}(1),q_{01}(2)\big)^t\mathbb{1}_{\overline{\mathcal{E}}}\qquad a.s.$$

Proof of Theorem 3.4. It is a direct consequence of equation (5.10), Proposition 4.14 and Lemmas 5.16, 5.17 and 5.21. □

6. Asymptotic normality

To derive the central limit theorems (CLT), we use a CLT for martingales given in ([12], Thm. 3.II.10). To this aim, we use a new filtration. Namely, instead of using the observed generation-wise filtration, we will use the sister pair-wise one. Let

$$\mathcal{G}_p^{\mathcal{O}}=\mathcal{O}\vee\sigma\{\delta_1X_1,\ (\delta_{2k}X_{2k},\delta_{2k+1}X_{2k+1}),\ 1\leq k\leq p\}$$

be the σ-algebra generated by the whole history $\mathcal{O}$ of the Galton−Watson process and all observed individuals up to the offspring of individual p. Hence $(\delta_{2k}\epsilon_{2k},\delta_{2k+1}\epsilon_{2k+1})$ is $\mathcal{G}_k^{\mathcal{O}}$-measurable. In all the sequel, we will work on the non-extinction probability space $(\overline{\mathcal{E}},\mathbb{P}_{\overline{\mathcal{E}}})$ and we denote by $\mathbb{E}_{\overline{\mathcal{E}}}$ the corresponding expectation.

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proof of Theorem 3.5, first step. For a fixed integer $n \geq 1$, let us define the $\mathcal{G}_p^{\mathcal{O}}$-martingale $(\boldsymbol{M}_p^{(n)})_{\{p\geq 1\}}$ by

$$\boldsymbol{M}_p^{(n)} = \frac{1}{|\mathbb{T}_n^*|^{1/2}} \sum_{k=1}^{p} \boldsymbol{D}_k \qquad \text{with} \qquad \boldsymbol{D}_k = (\epsilon_{2k}, X_k\epsilon_{2k}, \epsilon_{2k+1}, X_k\epsilon_{2k+1})^t .$$

Indeed, under (**H.1**−**5**), D_k is clearly a $\mathcal{G}_k^{\mathcal{O}}$-martingale difference sequence. Set $\nu_n = |\mathbb{T}_n| = 2^{n+1} - 1$. Therefore the following equality holds

$$\boldsymbol{M}_{\nu_n}^{(n)} = \frac{1}{|\mathbb{T}_n^*|^{1/2}} \sum_{k=1}^{|\mathbb{T}_n|} \boldsymbol{D}_k = \frac{1}{|\mathbb{T}_n^*|^{1/2}} \boldsymbol{M}_n.$$

We want to apply Theorem 3.II.10 of [12] to the process $(\boldsymbol{M}_{\nu_n}^{(n)})$. As the non-extinction set $\overline{\mathcal{E}}$ is in $\mathcal{G}_k^{\mathcal{O}}$ for every $k \geq 1$, it is easy to prove that

$$\mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k\boldsymbol{D}_k^t|\mathcal{G}_{k-1}^{\mathcal{O}}] = \mathbb{E}[\boldsymbol{D}_k\boldsymbol{D}_k^t|\mathcal{G}_{k-1}^{\mathcal{O}}] = \boldsymbol{\gamma}_k \otimes \begin{pmatrix} 1 & X_k \\ X_k & X_k^2 \end{pmatrix},$$

where γ_k is defined in equation (5.7). Lemma 5.4 gives the $\mathbb{P}_{\overline{\mathcal{E}}}$ almost sure limit of the following process

$$\frac{1}{|\mathbb{T}_n^*|} \sum_{k\in\mathbb{T}_n^*} \mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k\boldsymbol{D}_k^t|\mathcal{G}_{k-1}^{\mathcal{O}}] \xrightarrow[n\to\infty]{} \boldsymbol{\Gamma} \qquad \text{a.s.}$$

Therefore, the first assumption of Theorem 3.II.10 from [12] holds under $\mathbb{P}_{\overline{\mathcal{E}}}$. Thanks to $\gamma \geq \kappa \geq 4$, we can easily prove that for some $r > 2$, one obtains

$$\sup_{k\geq 0} \mathbb{E}[\|\boldsymbol{D}_k\|^r|\mathcal{G}_{k-1}^{\mathcal{O}}] < \infty \qquad \text{a.s.}$$

which in turn implies the Lindeberg condition. We can now conclude that under $\mathbb{P}_{\overline{\mathcal{E}}}$ the following convergence holds

$$\frac{1}{|\mathbb{T}_{n-1}^*|^{1/2}} \sum_{k\in\mathbb{T}_{n-1}} \boldsymbol{D}_k = \frac{1}{|\mathbb{T}_{n-1}^*|^{1/2}} \boldsymbol{M}_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Gamma}).$$

Finally, result (3.3) follows from equation (5.5) and Proposition 4.14 together with Slutsky's Lemma. □

Proof of Theorem 3.5, second step. We apply Theorem 3.II.10 of [12] again to the sequences $(\boldsymbol{M}_p^{(\boldsymbol{\sigma},n)})_{\{p\geq 1\}}$ of $\mathcal{G}_p^{\mathcal{O}}$-martingales defined by

$$|\mathbb{T}_{n-1}^*|^{1/2}\boldsymbol{M}_p^{(\boldsymbol{\sigma},n)} = \sum_{k=1}^{p} \boldsymbol{D}_k^{\boldsymbol{\sigma}} = \sum_{k=1}^{p} \begin{pmatrix} \epsilon_{2k}^2 + \epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k}^2 + \epsilon_{2k+1}^2 \mid \mathcal{F}_{r_k}^{\mathcal{O}}] \\ 2X_k(\epsilon_{2k}^2 - \mathbb{E}[\epsilon_{2k}^2 \mid \mathcal{F}_{r_k}^{\mathcal{O}}]) \\ 2X_k(\epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k+1}^2 \mid \mathcal{F}_{r_k}^{\mathcal{O}}]) \\ X_k^2(\epsilon_{2k}^2 + \epsilon_{2k+1}^2 - \mathbb{E}[\epsilon_{2k}^2 + \epsilon_{2k+1}^2 \mid \mathcal{F}_{r_k}^{\mathcal{O}}]) \end{pmatrix},$$

where r_k denotes the generation of k. Set $\nu_n = |\mathbb{T}_{n-1}| = 2^n - 1$. One obtains $|\mathbb{T}_{n-1}^*|^{1/2}\boldsymbol{M}_{\nu_n}^{(\boldsymbol{\sigma},n)} = \boldsymbol{U}_{n-1}(\boldsymbol{\sigma}_n - \boldsymbol{\sigma})$. We have to study the limit $\boldsymbol{\Gamma}^{\sigma}$ of the process $\frac{1}{|\mathbb{T}_{n-1}^*|}\sum_{k=1}^{\mathbb{T}_{n-1}} \mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k^{\boldsymbol{\sigma}}(\boldsymbol{D}_k^{\boldsymbol{\sigma}})^t \mid \mathcal{G}_{k-1}^{\mathcal{O}}]$. In order to compute the

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

conditional expectation, let us denote, for $k \geq 1$

$$A_i(k) = \delta_{2k+i}\Big(\sum_{r=0}^{4} C_4^l \theta((1-i)(4-r),(1-i)r,i(4-r),ir) X_k^r \\ -\big(\sigma_\varepsilon^4 + 4\rho_{ii}\sigma_\varepsilon^2 X_k + (4\rho_{ii}^2 + 2\sigma_\varepsilon^2\sigma_\eta^2)X_k^2 + 4\rho_{ii}\sigma_\eta^2 X_k^3 + \sigma_\eta^4 X_k^4\big)\Big),$$

$$A_{01}(k) = \delta_{2k}\delta_{2k+1}\Big(\sum_{r=0}^{2}\sum_{s=0}^{2} C_2^r C_2^s \theta(2-r,r,2-s,s) X_k^{r+s} \\ -\big(\sigma_\varepsilon^4 + 2\sigma_\varepsilon^2(\rho_{00}+\rho_{11})X_k + (2\sigma_\varepsilon^2\sigma_\eta^2 + 4\rho_{00}\rho_{11})X_k^2 + 2\sigma_\eta^2(\rho_{00}+\rho_{11})X_k^3 + \sigma_\eta^4 X_k^4\big)\Big),$$

and $B_i(k) = A_i(k) + A_{01}(k)$. Using these notations, we obtain

$$\mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k^{\boldsymbol{\sigma}}(\boldsymbol{D}_k^{\boldsymbol{\sigma}})^t \mid \mathcal{G}_{k-1}^{\mathcal{O}}] \\ = \begin{pmatrix} (B_0+B_1)(k) & 2X_kB_0(k) & 2X_kB_1(k) & X_k^2(B_0+B_1)(k) \\ 2X_kB_0(k) & 4X_k^2A_0(k) & 4X_k^2A_{01}(k) & 2X_k^3B_0(k) \\ 2X_kB_1(k) & 4X_k^2A_{01}(k) & 4X_k^2A_1(k) & 2X_k^3B_1(k) \\ X_k^2(B_0+B_1)(k) & 2X_k^3B_0(k) & 2X_k^3B_1(k) & X_k^4(B_0+B_1)(k) \end{pmatrix}.$$

The almost sure limit of the above quantity is given by Proposition 4.11. Indeed, the following convergences hold

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_{n-1}^*|} \sum_{k=1}^{\mathbb{T}_{n-1}} A_i(k) X_k^q = A_i^q \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{|\mathbb{T}_{n-1}^*|} \sum_{k=1}^{\mathbb{T}_{n-1}} A_{01}(k) X_k^q = A_{01}^q \quad \text{a.s.}$$

with

$$A_i^q = \sum_{r=0}^{4} C_4^l \theta((1-i)(4-r),(1-i)r,i(4-r),ir)\ell_i(r+q) \\ -\big(\sigma_\varepsilon^4\ell_i(q) + 4\rho_{ii}\sigma_\varepsilon^2\ell_i(1+q) + (4\rho_{ii}^2 + 2\sigma_\varepsilon^2\sigma_\eta^2)\ell_i(2+q) + 4\rho_{ii}\sigma_\eta^2\ell_i(3+q) + \sigma_\eta^4\ell_i(4+q)4\big),$$

$$A_{01}^q = \sum_{r=0}^{2}\sum_{s=0}^{2} C_2^r C_2^s \theta(2-r,r,2-s,s)\ell_{01}(r+s+q) \\ -\big(\sigma_\varepsilon^4\ell_{01}(q) + 2\sigma_\varepsilon^2(\rho_{00}+\rho_{11})\ell_{01}(1+q) + (2\sigma_\varepsilon^2\sigma_\eta^2 + 4\rho_{00}\rho_{11})\ell_{01}(2+q) \\ +2\sigma_\eta^2(\rho_{00}+\rho_{11})\ell_{01}(3+q) + \sigma_\eta^4\ell_{01}(4+q)\big).$$

We also set $B_i^q = A_i^q + A_{01}^q$. With these notations, we are able to explicit the limit matrix $\boldsymbol{\Gamma}^\sigma$ of the process $\frac{1}{|\mathbb{T}_{n-1}^*|}\sum_{k=1}^{\mathbb{T}_{n-1}} \mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k^{\boldsymbol{\sigma}}(\boldsymbol{D}_k^{\boldsymbol{\sigma}})^t \mid \mathcal{G}_{k-1}^{\mathcal{O}}]$

$$\boldsymbol{\Gamma}^\sigma = \begin{pmatrix} B_0^0 + B_1^0 & 2B_0^1 & 2B_1^1 & B_0^2 + B_1^2 \\ 2B_0^1 & 4A_0^2 & 4A_{01}^2 & 2B_0^3 \\ 2B_1^1 & 4A_{01}^2 & 4A_1^2 & 2B_1^3 \\ B_0^2 + B_1^2 & 2B_0^3 & 2B_1^3 & B_0^4 + B_1^4 \end{pmatrix}. \tag{6.1}$$

The first assumption of Theorem 3.II.10 from [12] holds under $\mathbb{P}_{\overline{\mathcal{E}}}$ and we easily prove the second one to conclude that under $\mathbb{P}_{\overline{\mathcal{E}}}$

$$\boldsymbol{M}_{\nu_n}^{\boldsymbol{\sigma},n} = |\mathbb{T}_{n-1}^*|^{-1/2} \sum_{k\in\mathbb{T}_{n-1}} \boldsymbol{D}_k^{\boldsymbol{\sigma}} = |\mathbb{T}_{n-1}^*|^{-1/2} \boldsymbol{U}_{n-1}(\boldsymbol{\sigma}_n - \boldsymbol{\sigma}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Gamma}^\sigma).$$

We conclude using Proposition 4.14 and Theorem 3.3. □

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

Proof of Theorem 3.5, third step. We apply again Theorem 3.II.10 of [12] with to the sequence of $\mathcal{G}_p^{\mathcal{O}}$-martingales $(\boldsymbol{M}_p^{(\boldsymbol{\rho},n)})_{\{p\geq 1\}}$ defined by

$$|\mathbb{T}_{n-1}^*|^{1/2}\boldsymbol{M}_p^{(\boldsymbol{\rho},n)} = \sum_{k=1}^{p}\boldsymbol{D}_k^{\boldsymbol{\rho}} = \sum_{k=1}^{p}\begin{pmatrix} \epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_{r_k}^{\mathcal{O}}] \\ 2X_k\big(\epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_{r_k}^{\mathcal{O}}]\big) \\ X_k^2\big(\epsilon_{2k}\epsilon_{2k+1} - \mathbb{E}[\epsilon_{2k}\epsilon_{2k+1} \mid \mathcal{F}_{r_k}^{\mathcal{O}}]\big) \end{pmatrix}.$$

Set $\nu_n = |\mathbb{T}_{n-1}| = 2^n - 1$. Thus one can rewrite $|\mathbb{T}_{n-1}^*|^{1/2}\boldsymbol{M}_{\nu_n}^{(\rho,n)} = \boldsymbol{V}_{n-1}(\boldsymbol{\rho}_n - \boldsymbol{\rho})$. Let us denote

$$\begin{aligned} C(k) = \delta_{2k}\delta_{2k+1}\Big(&(\theta(2,0,2,0) - \rho_\varepsilon^2) + 2\big(\theta(2,0,1,1) + \theta(1,1,2,0) - 2\rho\rho_\varepsilon\big)X_k \\ &+ \big(\theta(0,2,2,0) + \theta(2,0,0,2) + 4\theta(1,1,1,1) - 4\rho^2 - 2\rho_\varepsilon\ \rho_\eta\big)X_k^2 \\ &+ 2\big(\theta(0,2,1,1) + \theta(1,1,0,2) - 2\rho\rho_\eta\big)X_k^3 + \big(\theta(2,0,2,0) - \rho_\eta^2\big)X_k^4\Big). \end{aligned}$$

We are now able to write

$$\mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k^{\boldsymbol{\rho}}(\boldsymbol{D}_k^{\boldsymbol{\rho}})^t \mid \mathcal{G}_{k-1}^{\mathcal{O}}] = C(k)\begin{pmatrix} 1 & 2X_k & X_k^2 \\ 2X_k & 4X_k^2 & 2X_k^3 \\ X_k^2 & 2X_k^3 & X_k^4 \end{pmatrix}.$$

For the determination of the limit $\boldsymbol{\Gamma}^{\boldsymbol{\rho}}$ of $|\mathbb{T}_{n-1}^*|^{-1}\sum_{k=1}^{\mathbb{T}_{n-1}}\mathbb{E}_{\overline{\mathcal{E}}}[\boldsymbol{D}_k^{\boldsymbol{\rho}}(\boldsymbol{D}_k^{\boldsymbol{\rho}})^t \mid \mathcal{G}_{k-1}^{\mathcal{O}}]$, let us remark, using Proposition 4.11 that

$$\lim_{n\to\infty}\frac{1}{|\mathbb{T}_{n-1}^*|}\sum_{k=1}^{\mathbb{T}_{n-1}}C(k)X_k^q = C^q \quad \text{a.s.}$$

with

$$\begin{aligned} C^q = &(\theta(2,0,2,0) - \rho_\varepsilon^2)\ell_{01}(q) + 2\big(\theta(2,0,1,1) + \theta(1,1,2,0) - 2\rho\rho_\varepsilon\big)\ell_{01}(1+q) \\ &+ \big(\theta(0,2,2,0) + \theta(2,0,0,2) + 4\theta(1,1,1,1) - 4\rho^2 - 2\rho_\varepsilon\ \rho_\eta\big)\ell_{01}(2+q) \\ &+ 2\big(\theta(0,2,1,1) + \theta(1,1,0,2) - 2\rho\rho_\eta\big)\ell_{01}(3+q) + \big(\theta(2,0,2,0) - \rho_\eta^2\big)\ell_{01}(4+q). \end{aligned}$$

The matrix $\boldsymbol{\Gamma}^{\boldsymbol{\rho}}$ is thus given by

$$\boldsymbol{\Gamma}^{\rho} = \begin{pmatrix} C^0 & 2C^1 & C^2 \\ 2C^1 & 4C^2 & 2C^3 \\ C^2 & 2C^3 & C^4 \end{pmatrix}. \tag{6.2}$$

The first assumption of Theorem 3.II.10 from [12] holds under $\mathbb{P}_{\overline{\mathcal{E}}}$ and we easily prove the second one to conclude that under $\mathbb{P}_{\overline{\mathcal{E}}}$

$$\boldsymbol{M}_{\nu_n}^{\boldsymbol{\rho},n} = |\mathbb{T}_{n-1}^*|^{-1/2}\sum_{k\in\mathbb{T}_{n-1}}\boldsymbol{D}_k^{\boldsymbol{\rho}} = |\mathbb{T}_{n-1}^*|^{-1/2}\boldsymbol{V}_{n-1}(\boldsymbol{\rho}_n - \boldsymbol{\rho}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Gamma}^{\boldsymbol{\rho}}).$$

We conclude using Proposition 4.14 and Theorem 3.4. □

7. Application to real data

We have applied our estimation procedure to the Escherichia coli data of [25] (these data are available on request from the corresponding author of [25]). E. coli is a rod-shaped bacterium that reproduces by dividing in the middle. Each cell has thus a new end (or *pole*), and an older one. The cell that inherits the old pole of its mother is called the old pole cell, the cell that inherits the new pole of its mother is called the new pole cell. Therefore, each cell has a *type*: old pole (even) or new pole (odd), inducing asymmetry in the cell division. Stewart *et al.* [25] filmed colonies of dividing cells, determining the complete lineages and the growth rate of

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

TABLE 1. Estimation of $\boldsymbol{\theta}$ on the data set *penna-2002-10-04-4*.

a	b	c	d
0.0363	0.0266	0.0306	0.1706
$[0.0275, 0.0450]$	$[-0.2094, 0.2627]$	$[0.0216, 0.0396]$	$[-0.0709, 0.4120]$

TABLE 2. Estimation of noise variances on the data set *penna-2002-10-04-4*.

σ_ε^2	σ_η^2
0.0004	0.2431
$[-0, 0002, 0.0010]$	$[-0.0750, 0.5613]$

each cell. Several attempts have already been made to fit BAR processes to these data, see [9,11,13,14], but only with fixed coefficients models. In particular, [14] suggests that such models cannot explain all the randomness of the data.

We have run our estimators on the data set `penna-2002-10-04-4` from the experiments of [25]. It is the largest data set of the experiment. It contains 663 cells up to generation 9 (note that there would be 1023 cells in a full tree up to generation 9). For each of the 663 observed cells, the measure of the growth rate is available. For each cell, its length was recorded from birth to division and the corresponding growth curve was fitted by an exponential function $t \mapsto \exp(\lambda t)$ where λ is called the growth rate of the cell. Growth rates go from the minimum value 0.009 to the maximal 0.067. The 0.01-quantile equals 0.024 and the 0.99-quantile equals 0.049. Mean and median equal 0.037 (std: 0.004). Note that even though the number of observed generations $n = 9$ is low, the rate of convergence of our estimators is $|\mathbb{T}_n^*|^{-1/2}$ which is of order $\pi^{-n/2}$. Here for $n = 9$, $|\mathbb{T}_9^*| = 663$. Table 1 gives the estimation $\widehat{\boldsymbol{\theta}}_9$ of $\boldsymbol{\theta}$ with the 95% Confidence Interval (C.I.) of each coefficient. Note that our estimator $\widehat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}$ is exactly the same as in [9], and of course we obtain the same point estimation. The confidence intervals are wider, as the variance is different. More precisely, the variance is given by the CLT for $\boldsymbol{\theta}$ in equation (3.3). We have approximated it by $\boldsymbol{S}_8^{-1}\boldsymbol{\Gamma}_8\boldsymbol{S}_8^{-1}$ thanks to the convergences given in Proposition 4.14 and Lemma 5.4. Table 2 gives the estimation of the variance coefficients σ_ε^2 and σ_η^2 of $\boldsymbol{\sigma}$ (other covariance coefficients of $\widehat{\boldsymbol{\sigma}}_9$ and $\widehat{\boldsymbol{\rho}}_9$ can be computed but are less easy to interpret). The variance of these parameters is again given by the central limit Theorem 3.5. To obtain confidence intervals, one needs an estimation of the joint moments of $(\varepsilon_2, \eta_2, \varepsilon_3, \eta_3)$ up to the order 4. Such estimators can be easily derived following the same ideas as in Section 3.1. A Wald's test for the positivity of σ_ε^2 (resp. σ_η^2) can be derived from Theorem 3.5. It rejects the null hypothesis $H_0 : \sigma_\varepsilon^2 = 0$ (resp. $H_0 : \sigma_\eta^2 = 0$) with p-value $p = 0.0799$ (resp. $p = 0.0671$). We are not far from supporting the validity of the random coefficients model.

REFERENCES

[1] V. Bansaye, Proliferating parasites in dividing cells: Kimmel's branching model revisited. *Ann. Appl. Probab.* **18** (2008) 967–996.

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

[3] B. Bercu, B. de Saporta and A. Gégout-Petit, Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron. J. Probab.* **14** (2009) 2492–2526.

[4] V. Blandin, Asymptotic results for bifurcating random coefficient autoregressive processes (2012). Preprint `ArXiv: 1204.2926`.

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

[6] Q.M. Bui and R.M. Huggins, Inference for the random coefficients bifurcating autoregressive model for cell lineage studies. *J. Statist. Plann. Inference* **81** (1999) 253–262.

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

[8] B. de Saporta, Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_n Y_n + b_n$ with Markovian coefficients. *Stochastic Process. Appl.* **115** (2005) 1954–1978.

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

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press

[10] B. de Saporta, A. Gégout Petit and L. Marsalle, Asymmetry tests for bifurcating autoregressive processes with missing data. *Stat. Probab. Lett.* **82** (2012) 1439–1444.

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

[12] M. Duflo, Random iterative models, Applications of Mathematics, vol. 34. Springer-Verlag, Berlin (1997).

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

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

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

[16] J.D. Hamilton, Time series analysis. Princeton University Press, Princeton, NJ (1994).

[17] T.E. Harris, The theory of branching processes. *Die Grundlehren der Mathematischen Wissenschaften*, Bd. 119. Springer-Verlag, Berlin (1963).

[18] R.M. Huggins, Robust inference for variance components models for single trees of cell lineage data. *Ann. Statist.* **24** (1996) 1145–1160.

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

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

[21] R.M. Huggins and R.G. Staudte, Variance components models for dependent cell populations. *J. AMS* **89** (1994) 19–29.

[22] S.Y. Hwang and I.V. Basawa, Branching Markov processes and related asymptotics. *J. Multivariate Anal.* **100** (2009) 1155–1167.

[23] S.Y. Hwang and I.V. Basawa, Asymptotic optimal inference for multivariate branching-Markov processes via martingale estimating functions and mixed normality. *J. Multivariate Anal.* **102** (2011) 1018–1031.

[24] Nicholls, D. F., and Quinn, B. G. Random coefficient autoregressive models: an introduction. In vol. 11, *Lect. Notes Statist.* Springer-Verlag, New York (1982).

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

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

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

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

https://doi.org/10.1051/ps/2013042 Published online by Cambridge University Press