Statistical inference of 2-type critical Galton–Watson processes with immigration

Kristóf Körmendi · Gyula Pap

Abstract In this paper the asymptotic behavior of the conditional least squares estimators of the offspring mean matrix for a 2-type critical positively regular Galton–Watson branching process with immigration is described. We also study this question for a natural estimator of the spectral radius of the offspring mean matrix, which we call criticality parameter. We discuss the subcritical case as well.

1 Introduction

Branching processes have a number of applications in biology, finance, economics, queueing theory etc., see e.g. Haccou, Jagers and Vatutin [6]. Many aspects of applications in epidemiology, genetics and cell kinetics were presented at the 2009 Badajoz Workshop on Branching Processes, see [18].

The estimation theory for single-type Galton–Watson branching processes with immigration has a long history, see the survey paper of Winnicki [21]. The critical case is the most interesting and complicated one. There are two multi-type critical Galton–Watson processes with immigration for which statistical inference is available: the unstable integer-valued autoregressive models of order 2 (which can be considered as a special 2-type Galton–Watson branching process with immigration), see Barczy et al. [3] and the 2-type doubly symmetric process, see Ispány et al. [9]. In the present paper the asymptotic behavior of the conditional least squares (CLS) estimator of the offspring means and criticality parameter for the general 2-type critical positively regular Galton–Watson process with immigration is described, see Theorem 3.1. It turns out that in a degenerate case

Kristóf Körmendi
MTA-SZTE Analysis and Stochastics Research Group,
Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.
Tel.: +36–30–6436554
E-mail: kormendi@math.u-szeged.hu

Gyula Pap
Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary
E-mail: papgy@math.u-szeged.hu

this estimator is not even weakly consistent. We also study the asymptotic behavior of a natural estimator of the spectral radius of the offspring mean matrix, which we call criticality parameter. We discuss the subcritical case as well, but the supercritical case still remains open.

Let us recall the results for a single-type Galton–Watson branching process $(X_k)_{k\in\mathbb{Z}_+}$ with immigration. Assuming that the immigration mean m_ε is known, the CLS estimator of the offspring mean m_ξ based on the observations $X_1,\ldots,X_n$ has the form

$$\widehat{m}_\xi^{(n)} = \frac{\sum_{k=1}^n X_{k-1}(X_k - m_\varepsilon)}{\sum_{k=1}^n X_{k-1}^2}$$

on the set $\sum_{k=1}^n X_{k-1}^2 > 0$, see Klimko ans Nelson [14]. Suppose that $m_\varepsilon > 0$, and the second moment of the branching and immigration distributions are finite.

If the process is subcritical, i.e., $m_\xi < 1$, then the probability of the existence of the estimator $\widehat{m}_\xi^{(n)}$ tends to 1 as $n \to \infty$, and the estimator $\widehat{m}_\xi^{(n)}$ is strongly consistent, i.e., $\widehat{m}_\xi^{(n)} \xrightarrow{\text{a.s.}} m_\xi$ as $n \to \infty$. If, in addition, the third moments of the branching and immigration distributions are finite, then

$$n^{1/2}(\widehat{m}_\xi^{(n)} - m_\xi) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{V_\xi\,\mathbb{E}(\widetilde{X}^3) + V_\varepsilon\,\mathbb{E}(\widetilde{X}^2)}{\left[\mathbb{E}(\widetilde{X}^2)\right]^2}\right) \qquad \text{as } n \to \infty, \tag{1.1}$$

where V_ξ and V_ε denote the offspring and immigration variance, respectively, and the distribution of the random variable $\widetilde{X}$ is the unique stationary distribution of the Markov chain $(X_k)_{k\in\mathbb{Z}_+}$. Klimko and Nelson [14] contains a similar results for the CLS estimator of (m_ξ, m_ε), and (1.1) can be derived by the method of that paper, see also Theorem 3.4. Note that $\mathbb{E}(\widetilde{X}^2)$ and $\mathbb{E}(\widetilde{X}^3)$ can be expressed by the first three moments of the branching and immigration distributions.

If the process is critical, i.e., $m_\xi = 1$, then the probability of the existence of the estimator $\widehat{m}_\xi^{(n)}$ tends to 1 as $n \to \infty$, and

$$n(\widehat{m}_\xi^{(n)} - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{X}_t\,\mathrm{d}(\mathcal{X}_t - m_\varepsilon t)}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t} \qquad \text{as } n \to \infty, \tag{1.2}$$

where the process $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{X}_t = m_\varepsilon\,\mathrm{d}t + \sqrt{V_\xi \mathcal{X}_t^+}\,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+,$$

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process, and x^+ denotes the positive part of $x \in \mathbb{R}$. Note that this so-called square-root process is also known as Feller diffusion, or Cox–Ingersoll–Ross [4] model in financial mathematics. Wei and Winnicki [20] proved a similar results for the CLS estimator of the offspring mean when the immigration mean is unknown, and (1.2) can be derived by the method of that paper. Note that $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ as $n \to \infty$ with $\mathcal{X}_t^{(n)} := n^{-1}X_{\lfloor nt \rfloor}$ for $t \in \mathbb{R}_+$, $n \in \mathbb{N}$, where $\lfloor x \rfloor$ denotes the (lower) integer part of $x \in \mathbb{R}$, see Wei and Winnicki [19]. We call the reader's attention that we use the notation $\xrightarrow{\mathcal{D}}$ for the weak convergenge in the Skorokhod space and also for the weak convergence in $\mathbb{R}$. Based on the context it should be clear which convergence do we think of. If, in addition, $V_\xi = 0$, then

$$n^{3/2}(\widehat{m}_\xi^{(n)} - 1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{3V_\varepsilon}{m_\varepsilon^2}\right) \qquad \text{as } n \to \infty, \tag{1.3}$$

see Ispány et al. [12].

If the process is supercritical, i.e., $m_\xi > 1$, then the probability of the existence of the estimator $\widehat{m}_\xi^{(n)}$ tends to 1 as $n \to \infty$, the estimator $\widehat{m}_\xi^{(n)}$ is strongly consistent, i.e., $\widehat{m}_\xi^{(n)} \xrightarrow{\text{a.s.}} m_\xi$ as $n \to \infty$, and

$$\left(\sum_{k=1}^n X_{k-1}\right)^{1/2}(\widehat{m}_\xi^{(n)} - m_\xi) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{(m_\xi + 1)^2}{m_\xi^2 + m_\xi + 1} V_\xi\right) \qquad \text{as } n \to \infty. \tag{1.4}$$

Wei and Winnicki [20] showed the same asymptotic behavior for the CLS estimator of the offspring mean when the immigration mean is unknown, and (1.4) can be derived by the method of that paper.

In Section 2 we recall some preliminaries on 2-type Galton–Watson models with immigration. Section 3 contains our main results. Section 4 contains a useful decomposition of the process. Sections 5 contains the proofs. In Appendix A we present estimates for the moments of the processes involved. Appendix B is devoted to the CLS estimators. An extended version of this paper with more detailed proofs is available on arXiv, see [15].

2 Preliminaries on 2-type Galton–Watson models with immigration

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers, real numbers and non-negative real numbers, respectively. Every random variable will be defined on a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

For each $k, j \in \mathbb{Z}_+$ and $i, \ell \in \{1, 2\}$, the number of individuals of type i in the kth generation will be denoted by $X_{k,i}$, the number of type ℓ offsprings produced by the jth individual who is of type i belonging to the $(k-1)^{\text{th}}$ generation will be denoted by $\xi_{k,j,i,\ell}$, and the number of type i immigrants in the kth generation will be denoted by $\varepsilon_{k,i}$. Then we have

$$\begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} = \sum_{j=1}^{X_{k-1,1}} \begin{bmatrix} \xi_{k,j,1,1} \\ \xi_{k,j,1,2} \end{bmatrix} + \sum_{j=1}^{X_{k-1,2}} \begin{bmatrix} \xi_{k,j,2,1} \\ \xi_{k,j,2,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}, \qquad k \in \mathbb{N}. \tag{2.1}$$

Here $\big\{ \boldsymbol{X}_0, \boldsymbol{\xi}_{k,j,i}, \boldsymbol{\varepsilon}_k : k, j \in \mathbb{N}, \, i \in \{1,2\} \big\}$ are supposed to be independent, where

$$\boldsymbol{X}_k := \begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix}, \qquad \boldsymbol{\xi}_{k,j,i} := \begin{bmatrix} \xi_{k,j,i,1} \\ \xi_{k,j,i,2} \end{bmatrix}, \qquad \boldsymbol{\varepsilon}_k := \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}.$$

Moreover, $\{\boldsymbol{\xi}_{k,j,1} : k, j \in \mathbb{N}\}$, $\{\boldsymbol{\xi}_{k,j,2} : k, j \in \mathbb{N}\}$ and $\{\boldsymbol{\varepsilon}_k : k \in \mathbb{N}\}$ are supposed to consist of identically distributed random vectors.

We suppose $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2) < \infty$. Introduce the notations

$$\boldsymbol{m}_{\boldsymbol{\xi}_i} := \mathbb{E}(\boldsymbol{\xi}_{1,1,i}) \in \mathbb{R}_+^2, \qquad \boldsymbol{m}_{\boldsymbol{\xi}} := \begin{bmatrix} \boldsymbol{m}_{\boldsymbol{\xi}_1} & \boldsymbol{m}_{\boldsymbol{\xi}_2} \end{bmatrix} \in \mathbb{R}_+^{2\times 2}, \qquad \boldsymbol{m}_{\boldsymbol{\varepsilon}} := \mathbb{E}(\boldsymbol{\varepsilon}_1) \in \mathbb{R}_+^2,$$

$$\boldsymbol{V}_{\boldsymbol{\xi}_i} := \mathrm{Var}(\boldsymbol{\xi}_{1,1,i}) \in \mathbb{R}^{2\times 2}, \qquad \boldsymbol{V}_{\boldsymbol{\varepsilon}} := \mathrm{Var}(\boldsymbol{\varepsilon}_1) \in \mathbb{R}^{2\times 2}, \qquad i \in \{1,2\}.$$

We call the parameters $\boldsymbol{m}_{\boldsymbol{\xi}}$ and $\boldsymbol{m}_{\boldsymbol{\varepsilon}}$ the offspring mean matrix and the immigration mean vector respectively. Note that many authors define the offspring mean matrix as $\boldsymbol{m}_{\boldsymbol{\xi}}^\top$. For $k \in \mathbb{Z}_+$, let $\mathcal{F}_k := \sigma(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k)$. By (2.1),

$$\mathbb{E}(\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1}) = X_{k-1,1}\,\boldsymbol{m}_{\boldsymbol{\xi}_1} + X_{k-1,2}\,\boldsymbol{m}_{\boldsymbol{\xi}_2} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} = \boldsymbol{m}_{\boldsymbol{\xi}}\,\boldsymbol{X}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}. \tag{2.2}$$

Consequently,

$$\mathbb{E}(\boldsymbol{X}_k) = \boldsymbol{m}_{\boldsymbol{\xi}}\,\mathbb{E}(\boldsymbol{X}_{k-1}) + \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N},$$

which implies

$$\mathbb{E}(\boldsymbol{X}_k) = \boldsymbol{m}_{\boldsymbol{\xi}}^k\,\mathbb{E}(\boldsymbol{X}_0) + \sum_{j=0}^{k-1} \boldsymbol{m}_{\boldsymbol{\xi}}^j\,\boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}. \tag{2.3}$$

Hence, the asymptotic behavior of the sequence $(\mathbb{E}(\boldsymbol{X}_k))_{k \in \mathbb{Z}_+}$ depends on the asymptotic behavior of the powers $(\boldsymbol{m}_{\boldsymbol{\xi}}^k)_{k \in \mathbb{N}}$ of the offspring mean matrix, which is related to the spectral radius $r(\boldsymbol{m}_{\boldsymbol{\xi}}) := \varrho \in \mathbb{R}_+$ of $\boldsymbol{m}_{\boldsymbol{\xi}}$ (see the Perron–Frobenius theorem, e.g., Horn and Johnson [7, Theorems 8.2.11 and 8.5.1]). A 2-type Galton–Watson process $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ with immigration is referred to respectively as *subcritical, critical* or

supercritical if $\varrho < 1$, $\varrho = 1$ or $\varrho > 1$ (see, e.g., Athreya and Ney [1, V.3] or Quine [17]). We will write the offspring mean matrix of a 2-type Galton–Watson process with immigration in the form

$$\boldsymbol{m_\xi} := \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}. \tag{2.4}$$

Then its spectral radius is

$$\varrho = \frac{\alpha + \delta + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2}. \tag{2.5}$$

We study only *critical* 2-type Galton–Watson processes with immigration, i.e., when $\varrho = 1$, which is equivalent to $\alpha, \delta \in [0, 1]$ and $\beta, \gamma \in [0, \infty)$ with $\beta\gamma = (1 - \alpha)(1 - \delta)$. We will focus only on the case when the offpring mean matrix is *positively regular*, i.e., when there is a positive integer $k \in \mathbb{N}$ such that the entries of $\boldsymbol{m_\xi}^k$ are positive (see Kesten and Stigum [13]), which is equivalent to $\beta, \gamma \in (0, \infty)$, $\alpha, \delta \in \mathbb{R}_+$ with $\alpha + \delta > 0$. Then the matrix $\boldsymbol{m_\xi}$ has eigenvalues 1 and

$$\lambda := \alpha + \delta - 1 \in (-1, 1).$$

By the Perron–Frobenius theorem (see, e.g., Horn and Johnson [7, Theorems 8.2.11 and 8.5.1]),

$$\boldsymbol{m_\xi}^k \to \boldsymbol{u}_{\mathrm{right}} \boldsymbol{u}_{\mathrm{left}}^\top \qquad \text{as} \;\; k \to \infty,$$

where $\boldsymbol{u}_{\mathrm{right}}$ is the unique right eigenvector of $\boldsymbol{m_\xi}$ (called the right Perron vector of $\boldsymbol{m_\xi}$) corresponding to the eigenvalue 1 such that the sum of its coordinates is 1, and $\boldsymbol{u}_{\mathrm{left}}$ is the unique left eigenvector of $\boldsymbol{m_\xi}$ (called the left Perron vector of $\boldsymbol{m_\xi}$) corresponding to the eigenvalue 1 such that $\langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{u}_{\mathrm{left}} \rangle = 1$, hence we have

$$\boldsymbol{u}_{\mathrm{right}} = \frac{1}{\beta + 1 - \alpha} \begin{bmatrix} \beta \\ 1 - \alpha \end{bmatrix}, \qquad \boldsymbol{u}_{\mathrm{left}} = \frac{1}{1 - \lambda} \begin{bmatrix} \gamma + 1 - \delta \\ \beta + 1 - \alpha \end{bmatrix}.$$

More precisely, using the so-called Putzer's spectral formula (see, e.g., Putzer [16]) the powers of $\boldsymbol{m_\xi}$ can be written in the form

$$\boldsymbol{m_\xi}^k = \frac{1}{1 - \lambda} \begin{bmatrix} 1 - \delta & \beta \\ \gamma & 1 - \alpha \end{bmatrix} + \frac{\lambda^k}{1 - \lambda} \begin{bmatrix} 1 - \alpha & -\beta \\ -\gamma & 1 - \delta \end{bmatrix}$$

$$= \boldsymbol{u}_{\mathrm{right}} \boldsymbol{u}_{\mathrm{left}}^\top + \lambda^k \boldsymbol{v}_{\mathrm{right}} \boldsymbol{v}_{\mathrm{left}}^\top, \qquad k \in \mathbb{N}, \tag{2.6}$$

where $\boldsymbol{v}_{\mathrm{right}}$ and $\boldsymbol{v}_{\mathrm{left}}$ are appropriate right and left eigenvectors of $\boldsymbol{m_\xi}$, respectively, belonging to the eigenvalue λ, for instance,

$$\boldsymbol{v}_{\mathrm{right}} = \frac{1}{1 - \lambda} \begin{bmatrix} -\beta - 1 + \alpha \\ \gamma + 1 - \delta \end{bmatrix}, \qquad \boldsymbol{v}_{\mathrm{left}} = \frac{1}{\beta + 1 - \alpha} \begin{bmatrix} -1 + \alpha \\ \beta \end{bmatrix}.$$

Next we will recall a convergence result for positively regular and critical 2-type CBI processes. For each $n \in \mathbb{N}$, consider the random step process

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad t \in \mathbb{R}_+.$$

The following theorem is a special case of the main result in Ispány and Pap [11, Theorem 3.1].

2.1 Theorem. *Let $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ be a 2-type Galton–Watson process with immigration such that $\alpha, \delta \in [0, 1)$ and $\beta, \gamma \in (0, \infty)$ with $\alpha + \delta > 0$ and $\beta\gamma = (1 - \alpha)(1 - \delta)$ (hence it is critical and positively regular), $\boldsymbol{X}_0 = \boldsymbol{0}$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2) < \infty$. Then*

$$(\boldsymbol{\mathcal{X}}_t^{(n)})_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{D}} (\boldsymbol{\mathcal{X}}_t)_{t \in \mathbb{R}_+} := (\mathcal{Z}_t \boldsymbol{u}_{\mathrm{right}})_{t \in \mathbb{R}_+} \qquad \text{as} \;\; n \to \infty \tag{2.7}$$

in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, where $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$\mathrm{d}\mathcal{Z}_t = \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m_\varepsilon} \rangle \, \mathrm{d}t + \sqrt{\langle \overline{\boldsymbol{V}_\xi} \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{u}_{\mathrm{left}} \rangle \mathcal{Z}_t^+} \, \mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+, \qquad \mathcal{Z}_0 = 0, \tag{2.8}$$

where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ *is a standard Brownian motion and*

$$\overline{V_\xi} := \sum_{i=1}^{2} \langle e_i, u_{\text{right}} \rangle V_{\xi_i} = \frac{\beta V_{\xi_1} + (1-\alpha) V_{\xi_2}}{\beta + 1 - \alpha} \tag{2.9}$$

is a mixed offspring variance matrix.

In fact, in Ispány and Pap [11, Theorem 3.1], the above result has been prooved under the higher moment assumptions $\mathbb{E}(\|\xi_{1,1,1}\|^4) < \infty$, $\mathbb{E}(\|\xi_{1,1,2}\|^4) < \infty$ and $\mathbb{E}(\|\varepsilon_1\|^4) < \infty$, which have been relaxed in Danka and Pap [5, Theorem 3.1].

2.2 Remark. By Ikeda and Watanabe [8, Example 8.2, page 221], the SDE (2.8) has a unique strong solution $(\mathcal{Y}_t^{(y)})_{t \in \mathbb{R}_+}$ for all initial values $\mathcal{Y}_0^{(y)} = y \in \mathbb{R}$, and if $y \geqslant 0$, then $\mathcal{Y}_t^{(y)}$ is nonnegative for all $t \in \mathbb{R}_+$ with probability one, hence $\mathcal{Y}_t^+$ may be replaced by $\mathcal{Y}_t$ under the square root in (2.8), see, e.g., Barczy et al. [2, Remark 3.3]. $\qquad\square$

Clearly, $\overline{V_\xi}$ depends only on the branching distributions, i.e., on the distributions of $\xi_{1,1,1}$ and $\xi_{1,1,2}$. Note that $\overline{V_\xi} = \text{Var}(Y_1 \,|\, Y_0 = u_{\text{right}})$, where $(Y_k)_{k \in \mathbb{Z}_+}$ is a 2-type Galton–Watson process without immigration such that its branching distributions are the same as that of $(X_t)_{k \in \mathbb{Z}_+}$, since for each $i \in \{1, 2\}$, $V_{\xi_i} = \text{Var}(Y_1 \,|\, Y_0 = e_i)$.

For the sake of simplicity, we consider a zero start Galton–Watson process with immigration, that is, we suppose $X_0 = 0$. The general case of nonzero initial value may be handled in a similar way, but we renounce to consider it. In the sequel we always assume $m_\varepsilon \neq 0$, otherwise $X_k = 0$ for all $k \in \mathbb{N}$.

3 Main results

For each $n \in \mathbb{N}$, any CLS estimator

$$\widehat{m}_\xi^{(n)} = \begin{bmatrix} \widehat{\alpha}_n & \widehat{\beta}_n \\ \widehat{\gamma}_n & \widehat{\delta}_n \end{bmatrix}$$

of the offspring mean matrix m_ξ based on a sample $X_1, \ldots, X_n$ has the form

$$\widehat{m}_\xi^{(n)} = B_n A_n^{-1} \tag{3.1}$$

on the set

$$\Omega_n := \{\omega \in \Omega : \det(A_n(\omega)) > 0\}, \tag{3.2}$$

where

$$A_n := \sum_{k=1}^{n} X_{k-1} X_{k-1}^\top, \qquad B_n := \sum_{k=1}^{n} (X_k - m_\varepsilon) X_{k-1}^\top, \tag{3.3}$$

see Lemma B.1. The spectral radius ϱ given in (2.5) can be called *criticality parameter*, and its natural estimator is

$$\widehat{\varrho}_n = r\left(\widehat{m}_\xi^{(n)}\right) := \frac{\widehat{\alpha}_n + \widehat{\delta}_n + \sqrt{(\widehat{\alpha}_n - \widehat{\delta}_n)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n}}{2}, \tag{3.4}$$

on the set on the set $\Omega_n \cap \widetilde{\Omega}_n$ with

$$\widetilde{\Omega}_n := \left\{\omega \in \Omega_n : (\widehat{\alpha}_n(\omega) - \widehat{\delta}_n(\omega))^2 + 4\widehat{\beta}_n(\omega)\widehat{\gamma}_n(\omega) \geqslant 0\right\}. \tag{3.5}$$

By Lemma B.4, if $\langle \overline{V_\xi} v_{\text{left}}, v_{\text{left}} \rangle > 0$ and the assumptions of Theorem 3.1 hold, then $\mathbb{P}(\Omega_n \cap \widetilde{\Omega}_n) \to 1$ as $n \to \infty$.

3.1 Theorem. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration such that* $\alpha,\delta\in[0,1)$ *and* $\beta,\gamma\in(0,\infty)$ *with* $\alpha+\delta>0$ *and* $\beta\gamma=(1-\alpha)(1-\delta)$ *(hence it is critical and and positively regular),* $\boldsymbol{X}_0=\boldsymbol{0}$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8)<\infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8)<\infty$, *and* $\boldsymbol{m}_{\boldsymbol{\varepsilon}}\neq\boldsymbol{0}$. *Suppose* $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle>0$. *Then the probability of the existence of the estimators* $\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}$, *and* $\widehat{\varrho}_n$ *tends to 1 as* $n\to\infty$. *Furthermore*

$$n^{1/2}(\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}-\boldsymbol{m}_{\boldsymbol{\xi}})\xrightarrow{\mathcal{D}}\frac{(1-\lambda^2)^{1/2}}{\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}}\frac{\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\int_0^1\mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t}{\int_0^1\mathcal{Y}_t\,\mathrm{d}t}\boldsymbol{v}_{\mathrm{left}}^{\top}, \tag{3.6}$$

$$n(\widehat{\varrho}_n-1)\xrightarrow{\mathcal{D}}\frac{\int_0^1\mathcal{Y}_t\,\mathrm{d}(\mathcal{Y}_t-t\langle\boldsymbol{u}_{\mathrm{left}},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle)}{\int_0^1\mathcal{Y}_t^2\,\mathrm{d}t}, \tag{3.7}$$

as $n\to\infty$, *where* $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ *is the pathwise unique strong solution of the SDE* (2.8), *and* $(\widetilde{\boldsymbol{\mathcal{W}}}_t)_{t\in\mathbb{R}_+}$ *is a 2-dimenional standard Wiener processes independent of* $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$.

3.2 Remark. We note that in the critical positively regular case the limit distribution for the CLS estimator of the offspring mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ is concentrated on the 2-dimensional subspace $\mathbb{R}^2\boldsymbol{v}_{\mathrm{left}}^{\top}\subset\mathbb{R}^{2\times2}$. Surprisingly, the scaling factor of the CLS estimators of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is $\sqrt{n}$, which is the same as in the subcritical case. The reason of this strange phenomenon can be understood from the joint asymptotic behavior of $\det(\boldsymbol{A}_n)$ and $\boldsymbol{D}_n\widetilde{\boldsymbol{A}}_n$ given in Theorem 4.1. One of the decisive tools in deriving the needed asymptotic behavior is a good bound for the moments of the involved processes, see Corollary A.3. $\square$

3.3 Remark. One of the assumptions of Theorem 3.1 is that $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle>0$. Since $\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}$ is a positive semidefinite matrix the only time this condition can fail is when $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle=0$. In this case the scaling factor of the CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is 1 instead of $\sqrt{n}$, therefore this estimator is not consistent. This is due to the fact that if $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle=0$ then the limit in the second and fourth convergence of Theorem 4.1 is 0. The extended paper on arXiv contains the explicit limit distribution along with the full proofs in this degenerate case. $\square$

It would be useful to know the asymptotics of these estimations not just in the critical case, but in the sub- and supercritical cases as well. We include here the results for the subcritical case. The proof is based on the martingale central limit theorem, and it can be found in the extended version of this paper on the arXiv. The same problem in the supercritical case is still open.

3.4 Theorem. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration such that* $\alpha,\delta\in[0,1)$ *and* $\beta,\gamma\in(0,\infty)$ *with* $\alpha+\delta>0$ *and* $\beta\gamma<(1-\alpha)(1-\delta)$ *(hence it is subcritical and positively regular),* $\boldsymbol{X}_0=\boldsymbol{0}$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2)<\infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2)<\infty$, $\boldsymbol{m}_{\boldsymbol{\varepsilon}}\neq\boldsymbol{0}$, *and at least one of the matrices* $\boldsymbol{V}_{\boldsymbol{\xi}_1}$, $\boldsymbol{V}_{\boldsymbol{\xi}_2}$, $\boldsymbol{V}_{\boldsymbol{\varepsilon}}$ *is invertible. Then the probability of the existence of the estimators* $\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}$ *and* $\widehat{\varrho}_n$ *tends to 1 as* $n\to\infty$, *and the estimators* $\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}$ *and* $\widehat{\varrho}_n$ *are strongly consistent, i.e.,* $\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}\xrightarrow{\mathrm{a.s.}}\boldsymbol{m}_{\boldsymbol{\xi}}$ *and* $\widehat{\varrho}_n\xrightarrow{\mathrm{a.s.}}\varrho$ *as* $n\to\infty$.

If, in addition, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^6)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^6)<\infty$ *and* $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^6)<\infty$, *then*

$$n^{1/2}(\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}-\boldsymbol{m}_{\boldsymbol{\xi}})\xrightarrow{\mathcal{D}}\boldsymbol{Z}, \tag{3.8}$$

$$n^{1/2}(\widehat{\varrho}_n-\varrho)\xrightarrow{\mathcal{D}}\mathrm{Tr}(\boldsymbol{R}\boldsymbol{Z})\overset{\mathcal{D}}{=}\mathcal{N}\big(0,\mathrm{Tr}\big[\boldsymbol{R}^{\otimes2}\,\mathbb{E}(\boldsymbol{Z}^{\otimes2})\big]\big), \tag{3.9}$$

as $n\to\infty$, *where* $\boldsymbol{Z}$ *is a* 2×2 *random matrix having a normal distribution with zero mean and with*

$$\mathbb{E}(\boldsymbol{Z}^{\otimes2})=\Bigg\{\sum_{i=1}^2\mathbb{E}\big[(\boldsymbol{\xi}_{1,1,i}-\mathbb{E}(\boldsymbol{\xi}_{1,1,i}))^{\otimes2}\big]\,\mathbb{E}\Big[\widetilde{X}_i\big(\widetilde{\boldsymbol{X}}^{\top}\big)^{\otimes2}\Big]$$

$$+\mathbb{E}\big[(\boldsymbol{\varepsilon}_1-\mathbb{E}(\boldsymbol{\varepsilon}_1))^{\otimes2}\big]\,\mathbb{E}\Big[\big(\widetilde{\boldsymbol{X}}^{\top}\big)^{\otimes2}\Big]\Bigg\}\Big(\big[\mathbb{E}\big(\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^{\top}\big)\big]^{\otimes2}\Big)^{-1},$$

where the distribution of the 2-dimensional random vector $\widetilde{X}$ is the unique stationary distribution of the Markov chain $(X_k)_{k\in\mathbb{Z}_+}$, and

$$R := (\nabla r(m_{\boldsymbol{\xi}}))^{\top} = \frac{1}{2}I_2 + \frac{1}{2\sqrt{(\alpha-\delta)^2 + 4\beta\gamma}}\begin{bmatrix} \alpha-\delta & 2\beta \\ 2\gamma & \delta-\alpha \end{bmatrix}.$$

4 Decomposition of the process

Applying (2.2), let us introduce the sequence

$$M_k := X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = X_k - m_{\boldsymbol{\xi}}X_{k-1} - m_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}, \tag{4.1}$$

of martingale differences with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$. By (4.1), the process $(X_k)_{k\in\mathbb{Z}_+}$ satisfies the recursion

$$X_k = m_{\boldsymbol{\xi}}X_{k-1} + m_{\boldsymbol{\varepsilon}} + M_k, \qquad k \in \mathbb{N}. \tag{4.2}$$

By (3.1), for each $n \in \mathbb{N}$, we have

$$\widehat{m}_{\boldsymbol{\xi}}^{(n)} - m_{\boldsymbol{\xi}} = D_n A_n^{-1}$$

on the set Ω_n given in (3.2), where A_n is defined in (3.3), and

$$D_n := \sum_{k=1}^{n} M_k X_{k-1}^{\top}, \qquad n \in \mathbb{N}.$$

By (2.7) and the continous mapping theorem one can derive

$$n^{-3}A_n = \frac{1}{n^3}\sum_{k=1}^{n} X_k X_k^{\top} \xrightarrow{\mathcal{D}} \int_0^1 \mathcal{X}_t \mathcal{X}_t^{\top} \,\mathrm{d}t = \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t\, u_{\mathrm{right}}u_{\mathrm{right}}^{\top} =: \mathcal{A}$$

as $n \to \infty$. However, since $\det(\mathcal{A}) = 0$, the continuous mapping theorem can not be used for determining the weak limit of the sequence $(n^3 A_n^{-1})_{n\in\mathbb{N}}$. To avoid this, we can write

$$\widehat{m}_{\boldsymbol{\xi}}^{(n)} - m_{\boldsymbol{\xi}} = D_n A_n^{-1} = \frac{1}{\det(A_n)}D_n \widetilde{A}_n, \qquad n \in \mathbb{N}, \tag{4.3}$$

on the set Ω_n, where $\widetilde{A}_n$ denotes the adjugate of A_n (also called the matrix of cofactors) given by

$$\widetilde{A}_n := \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,2}^2 & -X_{k-1,1}X_{k-1,2} \\ -X_{k-1,1}X_{k-1,2} & X_{k-1,1}^2 \end{bmatrix}, \qquad n \in \mathbb{N}.$$

In order to prove Theorem 3.1 we will find the asymptotic behavior of the sequence $(\det(A_n), D_n\widetilde{A}_n)_{n\in\mathbb{N}}$. First we derive a useful decomposition for X_k, $k \in \mathbb{N}$. Let us introduce the sequence

$$U_k := \langle u_{\mathrm{left}}, X_k \rangle = \frac{(\gamma+1-\delta)X_{k,1} + (\beta+1-\alpha)X_{k,2}}{1-\lambda}, \qquad k \in \mathbb{Z}_+.$$

One can observe that $U_k \geqslant 0$ for all $k \in \mathbb{Z}_+$, and

$$U_k = U_{k-1} + \langle u_{\mathrm{left}}, m_{\boldsymbol{\varepsilon}} \rangle + \langle u_{\mathrm{left}}, M_k \rangle, \qquad k \in \mathbb{N}, \tag{4.4}$$

since $\langle u_{\mathrm{left}}, m_{\boldsymbol{\xi}}X_{k-1} \rangle = u_{\mathrm{left}}^{\top}m_{\boldsymbol{\xi}}X_{k-1} = u_{\mathrm{left}}^{\top}X_{k-1} = U_{k-1}$, because u_{left} is a left eigenvector of the mean matrix $m_{\boldsymbol{\xi}}$ belonging to the eigenvalue 1. Hence $(U_k)_{k\in\mathbb{Z}_+}$ is a nonnegative unstable AR(1) process with positive drift $\langle u_{\mathrm{left}}, m_{\boldsymbol{\varepsilon}} \rangle$ and with heteroscedastic innovation $(\langle u_{\mathrm{left}}, M_k \rangle)_{k\in\mathbb{N}}$. Note that the solution of the recursion (4.4) is

$$U_k = \sum_{j=1}^{k} \langle u_{\mathrm{left}}, M_j + m_{\boldsymbol{\varepsilon}} \rangle, \qquad k \in \mathbb{N}, \tag{4.5}$$

and, by the continous mapping theorem

$$(n^{-1}U_{\lfloor nt\rfloor})_{t\in\mathbb{R}_+} = (\langle\boldsymbol{u}_{\text{left}},\boldsymbol{\mathcal{X}}_t^{(n)}\rangle)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{D}} (\langle\boldsymbol{u}_{\text{left}},\boldsymbol{\mathcal{X}}_t\rangle)_{t\in\mathbb{R}_+} \overset{\mathcal{D}}{=} (\mathcal{Y}_t)_{t\in\mathbb{R}_+} \qquad \text{as } n\to\infty, \qquad (4.6)$$

where $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ is the pathwise unique strong solution of the SDE (2.8). Moreover, let

$$V_k := \langle\boldsymbol{v}_{\text{left}},\boldsymbol{X}_k\rangle = \frac{-(1-\alpha)X_{k,1}+\beta X_{k,2}}{\beta+1-\alpha}, \qquad k\in\mathbb{Z}_+.$$

Note that we have

$$V_k = \lambda V_{k-1} + \langle\boldsymbol{v}_{\text{left}},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle + \langle\boldsymbol{v}_{\text{left}},\boldsymbol{M}_k\rangle, \qquad k\in\mathbb{N}, \qquad (4.7)$$

since $\langle\boldsymbol{v}_{\text{left}},\boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{k-1}\rangle = \boldsymbol{v}_{\text{left}}^\top\boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{k-1} = \lambda\boldsymbol{v}_{\text{left}}^\top\boldsymbol{X}_{k-1} = \lambda V_{k-1},$ because $\boldsymbol{v}_{\text{left}}$ is a left eigenvector of the mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue λ. Thus $(V_k)_{k\in\mathbb{N}}$ is a stable AR(1) process with drift $\langle\boldsymbol{v}_{\text{left}},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle$ and with heteroscedastic innovation $(\langle\boldsymbol{v}_{\text{left}},\boldsymbol{M}_k\rangle)_{k\in\mathbb{N}}$. Note that the solution of the recursion (4.7) is

$$V_k = \sum_{j=1}^{k}\lambda^{k-j}\langle\boldsymbol{v}_{\text{left}},\boldsymbol{M}_j+\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle, \qquad k\in\mathbb{N}. \qquad (4.8)$$

By (2.1) and (4.1), we obtain the decomposition

$$\boldsymbol{M}_k = \sum_{j=1}^{X_{k-1,1}}\big(\boldsymbol{\xi}_{k,j,1}-\mathbb{E}(\boldsymbol{\xi}_{k,j,1})\big) + \sum_{j=1}^{X_{k-1,2}}\big(\boldsymbol{\xi}_{k,j,2}-\mathbb{E}(\boldsymbol{\xi}_{k,j,2})\big) + \big(\boldsymbol{\varepsilon}_k-\mathbb{E}(\boldsymbol{\varepsilon}_k)\big), \qquad k\in\mathbb{N}. \qquad (4.9)$$

The recursion (4.2) has the solution

$$\boldsymbol{X}_k = \sum_{j=1}^{k}\boldsymbol{m}_{\boldsymbol{\xi}}^{k-j}(\boldsymbol{m}_{\boldsymbol{\varepsilon}}+\boldsymbol{M}_j), \qquad k\in\mathbb{N}.$$

Consequently, using (2.6),

$$\boldsymbol{X}_k = \sum_{j=1}^{k}\big(\boldsymbol{u}_{\text{right}}\boldsymbol{u}_{\text{left}}^\top + \lambda^{k-j}\boldsymbol{v}_{\text{right}}\boldsymbol{v}_{\text{left}}^\top\big)(\boldsymbol{m}_{\boldsymbol{\varepsilon}}+\boldsymbol{M}_j)$$

$$= \boldsymbol{u}_{\text{right}}\boldsymbol{u}_{\text{left}}^\top\sum_{j=1}^{k}(\boldsymbol{X}_j-\boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{j-1}) + \boldsymbol{v}_{\text{right}}\boldsymbol{v}_{\text{left}}^\top\sum_{j=1}^{k}\lambda^{k-j}(\boldsymbol{X}_j-\boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{j-1})$$

$$= \boldsymbol{u}_{\text{right}}\boldsymbol{u}_{\text{left}}^\top\sum_{j=1}^{k}(\boldsymbol{X}_j-\boldsymbol{X}_{j-1}) + \boldsymbol{v}_{\text{right}}\boldsymbol{v}_{\text{left}}^\top\sum_{j=1}^{k}\big[\lambda^{k-j}\boldsymbol{X}_j-\lambda^{k-j+1}\boldsymbol{X}_{j-1}\big]$$

$$= \boldsymbol{u}_{\text{right}}\boldsymbol{u}_{\text{left}}^\top\boldsymbol{X}_k + \boldsymbol{v}_{\text{right}}\boldsymbol{v}_{\text{left}}^\top\boldsymbol{X}_k = U_k\boldsymbol{u}_{\text{right}} + V_k\boldsymbol{v}_{\text{right}},$$

hence

$$\boldsymbol{X}_k = \begin{bmatrix}X_{k,1}\\X_{k,2}\end{bmatrix} = \begin{bmatrix}\boldsymbol{u}_{\text{right}} & \boldsymbol{v}_{\text{right}}\end{bmatrix}\begin{bmatrix}U_k\\V_k\end{bmatrix} = \begin{bmatrix}\frac{\beta}{\beta+1-\alpha}U_k - \frac{\beta+1-\alpha}{1-\lambda}V_k\\ \frac{1-\alpha}{\beta+1-\alpha}U_k + \frac{\gamma+1-\delta}{1-\lambda}V_k\end{bmatrix}, \qquad k\in\mathbb{Z}_+. \qquad (4.10)$$

This decomposition yields

$$\det(\boldsymbol{A}_n) = \left(\sum_{k=1}^{n-1}U_k^2\right)\left(\sum_{k=1}^{n-1}V_k^2\right) - \left(\sum_{k=1}^{n-1}U_kV_k\right)^2, \qquad (4.11)$$

since

$$\det(\boldsymbol{A}_n) = \det\left(\sum_{k=1}^{n} \boldsymbol{X}_{k-1}\boldsymbol{X}_{k-1}^{\top}\right)$$

$$= \det\left(\begin{bmatrix} \boldsymbol{u}_{\text{right}} & \boldsymbol{v}_{\text{right}} \end{bmatrix}\sum_{k=1}^{n}\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^{\top}\begin{bmatrix} \boldsymbol{u}_{\text{right}} & \boldsymbol{v}_{\text{right}} \end{bmatrix}^{\top}\right)$$

$$= \det\left(\sum_{k=1}^{n}\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^{\top}\right)\left[\det\left(\begin{bmatrix} \boldsymbol{u}_{\text{right}} & \boldsymbol{v}_{\text{right}} \end{bmatrix}\right)\right]^2,$$

where

$$\det\left(\begin{bmatrix} \boldsymbol{u}_{\text{right}} & \boldsymbol{v}_{\text{right}} \end{bmatrix}\right) = 1. \tag{4.12}$$

Theorem 3.1 will follow from the following statement by the continuous mapping theorem and by Slutsky's lemma.

4.1 Theorem. *Suppose that the assumptions of Theorem 3.1 hold. If* $\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle > 0,$ *then*

$$\sum_{k=1}^{n} n^{-5/2}U_{k-1}V_{k-1} \xrightarrow{\mathbb{P}} 0 \qquad as \ \ n \to \infty,$$

$$\sum_{k=1}^{n}\begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-2}V_{k-1}^2 \\ n^{-2}\boldsymbol{M}_k U_{k-1} \\ n^{-3/2}\boldsymbol{M}_k V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle}{1-\lambda^2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}t \\ \int_0^1 \mathcal{Y}_t\,\mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle^{1/2}}{(1-\lambda^2)^{1/2}}\overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix} \qquad as \ \ n \to \infty.$$

Proof of Theorem 3.1. In order to derive the statements, we can use the continuous mapping theorem and Slutsky's lemma.

Theorem 4.1 implies (3.6). Indeed, we can use the representation (4.3), where the adjugate $\widetilde{\boldsymbol{A}}_n$ can be written in the form

$$\widetilde{\boldsymbol{A}}_n = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\sum_{\ell=1}^{n} \boldsymbol{X}_{\ell-1}\boldsymbol{X}_{\ell-1}^{\top}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Using (4.10), we have

$$\boldsymbol{D}_n\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^{n}\boldsymbol{M}_k\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^{\top}\begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}^{\top}\sum_{\ell=1}^{n}\begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix}\begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix}^{\top}\begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Here we have

$$\begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}^{\top} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} \boldsymbol{u}_{\text{right}}^{\top} \\ \boldsymbol{v}_{\text{right}}^{\top} \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\boldsymbol{v}_{\text{left}}^{\top} \\ \boldsymbol{u}_{\text{left}}^{\top} \end{bmatrix}.$$

Theorem 4.1 implies asymptotic expansions

$$\sum_{k=1}^{n}\boldsymbol{M}_k\begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^{\top} = n^2\boldsymbol{D}_{n,1} + n^{3/2}\boldsymbol{D}_{n,2},$$

$$\sum_{\ell=1}^{n}\begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix}\begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix}^{\top} = n^3\boldsymbol{A}_{n,1} + n^{5/2}\boldsymbol{A}_{n,2} + n^2\boldsymbol{A}_{n,3},$$

where

$$\boldsymbol{D}_{n,1} := n^{-2} \sum_{k=1}^{n} \boldsymbol{M}_k U_{k-1} \boldsymbol{e}_1^\top \xrightarrow{\mathcal{D}} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t \, \boldsymbol{e}_1^\top =: \boldsymbol{\mathcal{D}}_1,$$

$$\boldsymbol{D}_{n,2} := n^{-3/2} \sum_{k=1}^{n} \boldsymbol{M}_k V_{k-1} \boldsymbol{e}_2^\top \xrightarrow{\mathcal{D}} \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda^2)^{1/2}} \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{e}_2^\top =: \boldsymbol{\mathcal{D}}_2,$$

$$\boldsymbol{A}_{n,1} := n^{-3} \sum_{\ell=1}^{n} \begin{bmatrix} U_{\ell-1}^2 & 0 \\ 0 & 0 \end{bmatrix} \xrightarrow{\mathcal{D}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} =: \boldsymbol{\mathcal{A}}_1,$$

$$\boldsymbol{A}_{n,2} := n^{-5/2} \sum_{\ell=1}^{n} \begin{bmatrix} 0 & U_{\ell-1}V_{\ell-1} \\ U_{\ell-1}V_{\ell-1} & 0 \end{bmatrix} \xrightarrow{\mathcal{D}} \boldsymbol{0},$$

$$\boldsymbol{A}_{n,3} := n^{-2} \sum_{\ell=1}^{n} \begin{bmatrix} 0 & 0 \\ 0 & V_{\ell-1}^2 \end{bmatrix} \xrightarrow{\mathcal{D}} \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda^2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} =: \boldsymbol{\mathcal{A}}_3$$

jointly as $n \to \infty$. Consequently, we obtain an asymptotic expansion

$$\boldsymbol{D}_n \widetilde{\boldsymbol{A}}_n = (n^2 \boldsymbol{D}_{n,1} + n^{3/2} \boldsymbol{D}_{n,2}) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} (n^3 \boldsymbol{A}_{n,1} + n^{5/2} \boldsymbol{A}_{n,2} + n^2 \boldsymbol{A}_{n,3}) \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}$$

$$= (n^5 \boldsymbol{C}_{n,1} + n^{9/2} \boldsymbol{C}_{n,2} + n^4 \boldsymbol{C}_{n,3} + n^{7/2} \boldsymbol{C}_{n,4}) \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix},$$

where

$$\boldsymbol{C}_{n,1} := \boldsymbol{D}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} = n^{-5} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{M}_k U_{k-1} U_{\ell-1}^2 \boldsymbol{e}_1^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0}$$

for all $n \in \mathbb{N}$, and

$$\boldsymbol{C}_{n,2} := \boldsymbol{D}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} + \boldsymbol{D}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{D}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1,$$

$$\boldsymbol{C}_{n,3} := \boldsymbol{D}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,3} + \boldsymbol{D}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{D}}_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3,$$

$$\boldsymbol{C}_{n,4} := \boldsymbol{D}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,3} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{D}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3$$

as $n \to \infty$. Using again Theorem 4.1 and (4.11), we conclude

$$\begin{bmatrix} n^{-5} \det(\boldsymbol{A}_n) \\ n^{-9/2} \boldsymbol{D}_n \widetilde{\boldsymbol{A}}_n \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda^2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \\ \boldsymbol{\mathcal{D}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix} \end{bmatrix} \qquad \text{as } n \to \infty.$$

Here

$$\boldsymbol{\mathcal{D}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\text{left}}^\top \\ \boldsymbol{u}_{\text{left}}^\top \end{bmatrix}$$

$$= \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle^{1/2}}{(1-\lambda^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{e}_2^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{v}_{\text{left}}^\top \\ \boldsymbol{u}_{\text{left}}^\top \end{bmatrix}$$

$$= \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle^{1/2}}{(1-\lambda^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{v}_{\text{left}}^\top.$$

Since $\boldsymbol{m_\varepsilon} \neq \boldsymbol{0}$, by the SDE (2.8), we have $\mathbb{P}(\mathcal{Y}_t = 0 \text{ for all } t \in [0,1]) = 0$, which implies that $\mathbb{P}(\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t > 0) = 1$, hence the continuous mapping theorem implies (3.6).

The proof of (3.7) can be carried out similarly. For the details see the extended paper on arXiv. □

5 Proof of Theorem 4.1

The first convergence in Theorem 4.1 follows from Lemma B.2.

For the second convergence in Theorem 4.1, consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \boldsymbol{\mathcal{M}}_t^{(n)} \\ \boldsymbol{\mathcal{N}}_t^{(n)} \\ \boldsymbol{\mathcal{P}}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Z}_k^{(n)} \qquad \text{with} \qquad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1} \boldsymbol{M}_k \\ n^{-2} \boldsymbol{M}_k U_{k-1} \\ n^{-3/2} \boldsymbol{M}_k V_{k-1} \end{bmatrix} = \begin{bmatrix} n^{-1} \\ n^{-2} U_{k-1} \\ n^{-3/2} V_{k-1} \end{bmatrix} \otimes \boldsymbol{M}_k$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$, where $\otimes$ denotes Kronecker product of matrices. The second convergence in Theorem 4.1 follows from Lemma B.3 and the following theorem (this will be explained after Theorem 5.1).

5.1 Theorem. *Suppose that the assumptions of Theorem 4.1 hold. Then we have*

$$\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}} \qquad \text{as } n \to \infty, \tag{5.1}$$

where the process $(\boldsymbol{\mathcal{Z}}_t)_{t \in \mathbb{R}_+}$ *with values in* $(\mathbb{R}^2)^3$ *is the unique strong solution of the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t) \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{5.2}$$

with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ *and* $(\widetilde{\boldsymbol{\mathcal{W}}}_t)_{t \in \mathbb{R}_+}$ *are independent 2-dimensional standard Wiener processes, and* $\gamma : \mathbb{R}_+ \times (\mathbb{R}^2)^3 \to (\mathbb{R}^{2\times2})^{3\times2}$ *is defined by*

$$\gamma(t, \boldsymbol{x}) := \begin{bmatrix} (\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{x}_1 + t\boldsymbol{m_\varepsilon} \rangle^+)^{1/2} & 0 \\ (\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{x}_1 + t\boldsymbol{m_\varepsilon} \rangle^+)^{3/2} & 0 \\ 0 & \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle^{1/2}}{(1-\lambda^2)^{1/2}} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{x}_1 + t\boldsymbol{m_\varepsilon} \rangle \end{bmatrix} \otimes \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2}$$

for $t \in \mathbb{R}_+$ *and* $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \in (\mathbb{R}^2)^3$.

Note that the statement of Theorem 5.1 holds even if $\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle = 0$, when the last 2-dimensional coordinate process of the unique strong solution $(\boldsymbol{\mathcal{Z}}_t)_{t \in \mathbb{R}_+}$ is $\boldsymbol{0}$.

The SDE (5.2) has the form

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \mathrm{d}\boldsymbol{\mathcal{N}}_t \\ \mathrm{d}\boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} (\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m_\varepsilon}\rangle^+)^{1/2} \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ (\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m_\varepsilon}\rangle^+)^{3/2} \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle^{1/2}}{(1-\lambda^2)^{1/2}} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m_\varepsilon}\rangle \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+. \tag{5.3}$$

One can prove that the first 2-dimensional equation of the SDE (5.3) has a pathwise unique strong solution $(\boldsymbol{\mathcal{M}}_t^{(\boldsymbol{y}_0)})_{t\in\mathbb{R}_+}$ with arbitrary initial value $\boldsymbol{\mathcal{M}}_0^{(\boldsymbol{y}_0)} = \boldsymbol{y}_0 \in \mathbb{R}^2$, see the proof of Ispány and Pap [11, Theorem 3.1]. Thus the SDE (5.2) has a pathwise unique strong solution with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$, and we have

$$\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \boldsymbol{\mathcal{M}}_t \\ \boldsymbol{\mathcal{N}}_t \\ \boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s + s\boldsymbol{m_\varepsilon}\rangle^{1/2} \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_s \\ \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s + s\boldsymbol{m_\varepsilon}\rangle \, \mathrm{d}\boldsymbol{\mathcal{M}}_s \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle^{1/2}}{(1-\lambda^2)^{1/2}} \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s + s\boldsymbol{m_\varepsilon}\rangle \, \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

By the method of the proof of $\boldsymbol{\mathcal{X}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{X}}$ in Theorem 3.1 in Barczy et al. [2] one can derive

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{X}}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{5.4}$$

where

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad \widetilde{\boldsymbol{\mathcal{X}}}_t := \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m_\varepsilon}\rangle \boldsymbol{u}_{\text{right}}, \qquad t \in \mathbb{R}_+, \qquad n \in \mathbb{N}.$$

Next, similarly to the proof of (B.3), by the continuous mapping theorem, convergence (5.4) with $U_{k-1} = \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{X}_{k-1}\rangle$ and Lemma B.3 implies

$$\sum_{k=1}^n \begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-2}V_{k-1}^2 \\ n^{-2}\boldsymbol{M}_k U_{k-1} \\ n^{-3/2}\boldsymbol{M}_k V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \boldsymbol{u}_{\text{left}}, \widetilde{\boldsymbol{\mathcal{X}}}_t\rangle^2 \, \mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle}{1-\lambda^2} \int_0^1 \langle \boldsymbol{u}_{\text{left}}, \widetilde{\boldsymbol{\mathcal{X}}}_t\rangle \, \mathrm{d}t \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle^{1/2}}{(1-\lambda^2)^{1/2}} \int_0^1 \mathcal{Y}_t \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix} \qquad \text{as } n \to \infty.$$

This limiting random vector can be written in the form as given in Theorem 4.1, since $\langle \boldsymbol{u}_{\text{left}}, \widetilde{\boldsymbol{\mathcal{X}}}_t\rangle = \mathcal{Y}_t$ for all $t \in \mathbb{R}_+$.

Proof of Theorem 5.1. In order to show convergence $\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}}$, we apply a theorem concerning the convergence of random step processes (see [10], Corollary 2.2) with the special choices $\boldsymbol{\mathcal{U}} := \boldsymbol{\mathcal{Z}}$, $\boldsymbol{U}_k^{(n)} := \boldsymbol{Z}_k^{(n)}$, $n, k \in \mathbb{N}$, $(\mathcal{F}_k^{(n)})_{k\in\mathbb{Z}_+} := (\mathcal{F}_k)_{k\in\mathbb{Z}_+}$ and the function γ which is defined in Theorem 5.1. Note that the discussion after Theorem 5.1 shows that the SDE (5.2) admits a unique strong solution $(\boldsymbol{\mathcal{Z}}_t^{\boldsymbol{z}})_{t\in\mathbb{R}_+}$ for all initial values $\boldsymbol{\mathcal{Z}}_0^{\boldsymbol{z}} = \boldsymbol{z} \in (\mathbb{R}^2)^3$. The conditional variance has the form

$$\mathrm{Var}\big(\boldsymbol{Z}_k^{(n)} \,|\, \mathcal{F}_{k-1}\big) = \begin{bmatrix} n^{-2} & n^{-3}U_{k-1} & n^{-5/2}V_{k-1} \\ n^{-3}U_{k-1} & n^{-4}U_{k-1}^2 & n^{-7/2}U_{k-1}V_{k-1} \\ n^{-5/2}V_{k-1} & n^{-7/2}U_{k-1}V_{k-1} & n^{-3}V_{k-1}^2 \end{bmatrix} \otimes \boldsymbol{V}_{\boldsymbol{M}_k}$$

for $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, with $\boldsymbol{V}_{\boldsymbol{M}_k} := \mathrm{Var}(\boldsymbol{M}_k \,|\, \mathcal{F}_{k-1})$, and $\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})^\top$ has the form

$$\begin{bmatrix} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon}\rangle & \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon}\rangle^2 & \boldsymbol{0} \\ \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon}\rangle^2 & \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon}\rangle^3 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle}{1-\lambda^2} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon}\rangle^2 \end{bmatrix} \otimes \overline{\boldsymbol{V}_{\boldsymbol{\xi}}}$$

for $s \in \mathbb{R}_+$, where we used that $\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle^+ = \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$. Indeed, by (4.1), we get

$$\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle = \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{X}_k - \boldsymbol{m_\xi} \boldsymbol{X}_{k-1} - \boldsymbol{m_\varepsilon} \rangle + \langle \boldsymbol{u}_{\text{left}}, s\boldsymbol{m_\varepsilon} \rangle$$

$$= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{X}_k - \boldsymbol{X}_{k-1} - \boldsymbol{m_\varepsilon} \rangle + s \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m_\varepsilon} \rangle \tag{5.5}$$

$$= \frac{1}{n} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{X}_{\lfloor ns \rfloor} \rangle + \frac{ns - \lfloor ns \rfloor}{n} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m_\varepsilon} \rangle = \frac{1}{n} U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m_\varepsilon} \rangle \in \mathbb{R}_+$$

for $s \in \mathbb{R}_+$, $n \in \mathbb{N}$, since $\boldsymbol{u}_{\text{left}}^\top \boldsymbol{m_\xi} = \boldsymbol{u}_{\text{left}}^\top$ implies $\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m_\xi} \boldsymbol{X}_{k-1} \rangle = \boldsymbol{u}_{\text{left}}^\top \boldsymbol{m_\xi} \boldsymbol{X}_{k-1} = \boldsymbol{u}_{\text{left}}^\top \boldsymbol{X}_{k-1} = \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{X}_{k-1} \rangle$.

We need to prove that for each $T > 0$,

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{V}_{M_k} - \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle \overline{\boldsymbol{V_\xi}} \, \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.6}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \boldsymbol{V}_{M_k} - \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle^2 \overline{\boldsymbol{V_\xi}} \, \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.7}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \boldsymbol{V}_{M_k} - \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle^3 \overline{\boldsymbol{V_\xi}} \, \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.8}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V}_{M_k} - \frac{\langle \overline{\boldsymbol{V_\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle}{1 - \lambda^2} \int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle^2 \overline{\boldsymbol{V_\xi}} \, \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.9}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \boldsymbol{V}_{M_k} \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.10}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \boldsymbol{V}_{M_k} \right\| \xrightarrow{\mathbb{P}} 0 \tag{5.11}$$

as $n \to \infty$.

First we show (5.6). By (5.5),

$$\int_0^t \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m_\varepsilon} \rangle \, \mathrm{d}s = \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k + \frac{nt - \lfloor nt \rfloor}{n^2} U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2}{2n^2} \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m_\varepsilon} \rangle.$$

Using Lemma A.1, we have $\boldsymbol{V}_{M_k} = U_{k-1} \overline{\boldsymbol{V_\xi}} + V_{k-1} \widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \boldsymbol{V_\varepsilon}$, thus, in order to show (5.6), it suffices to prove

$$n^{-2} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0, \qquad n^{-2} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0, \tag{5.12}$$

$$n^{-2} \sup_{t \in [0,T]} \left[\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2 \right] \to 0 \tag{5.13}$$

as $n \to \infty$. Using (A.3) with $(\ell, i, j) = (2, 0, 1)$ and (A.4) with $(\ell, i, j) = (2, 1, 0)$, we have (5.12). Clearly, (5.13) follows from $|nt - \lfloor nt \rfloor| \leqslant 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, thus we conclude (5.6).

One can verify conditions (5.7) and (5.8) the same way. Checking conditions (5.8), (5.9) and (5.10) requires some more work. We only prove (5.9) here, but the same idea of decomposition can be used for the latter conditions as well. In order to prove (5.9) first we show that

$$
n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V_{M_k}} - \frac{\langle \overline{\boldsymbol{V_\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle}{1 - \lambda^2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\boldsymbol{V_\xi}} \right\| \overset{\mathbb{P}}{\longrightarrow} 0 \tag{5.14}
$$

as $n \to \infty$ for all $T > 0$. Using Lemma A.1, we obtain

$$
\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V_{M_k}} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 \overline{\boldsymbol{V_\xi}} + \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^3 \widetilde{\boldsymbol{V_\xi}} + \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V_\varepsilon}. \tag{5.15}
$$

Using (A.3) with $(\ell, i, j) = (6, 0, 3)$ and $(\ell, i, j) = (4, 0, 2)$, we have

$$
n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k|^3 \overset{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \overset{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty,
$$

hence (5.14) will follow from

$$
n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{\langle \overline{\boldsymbol{V_\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle}{1 - \lambda^2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right\| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty \tag{5.16}
$$

for all $T > 0$. The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2$ as a sum of a martingale and some other terms. Using recursions (4.7), (4.4) and formulas (A.1) and (A.2), we obtain

$$
\mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) = \mathbb{E}\Big(\big(U_{k-2} + \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_\varepsilon \rangle\big)\big(\lambda V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_\varepsilon \rangle\big)^2 \,\Big|\, \mathcal{F}_{k-2} \Big)
$$

$$
= \lambda^2 U_{k-2} V_{k-2}^2 + \boldsymbol{v}_{\text{left}}^\top \mathbb{E}(\boldsymbol{M}_{k-1} \boldsymbol{M}_{k-1}^\top \,|\, \mathcal{F}_{k-2}) \boldsymbol{v}_{\text{left}} \, U_{k-2}
$$

$$
+ \text{constant} + \text{linear combination of } U_{k-2} V_{k-2}, \ V_{k-2}^2, \ U_{k-2} \ \text{and} \ V_{k-2}
$$

$$
= \lambda^2 U_{k-2} V_{k-2}^2 + \langle \overline{\boldsymbol{V_\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle U_{k-2}^2 + \text{constant}
$$

$$
+ \text{linear combination of } U_{k-2} V_{k-2}, \ V_{k-2}^2, \ U_{k-2} \ \text{and} \ V_{k-2}.
$$

Thus

$$
\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 = \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) \big] + \sum_{k=2}^{\lfloor nt \rfloor} \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2})
$$

$$
= \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) \big] + \lambda^2 \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}^2 + \langle \overline{\boldsymbol{V_\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2
$$

$$
+ \mathrm{O}(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}, \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}^2, \ \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \ \text{and} \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.
$$

Consequently,

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 = \frac{1}{1-\lambda^2} \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \mid \mathcal{F}_{k-2}) \right]$$

$$+ \frac{\langle \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda^2} \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2 - \frac{\lambda^2}{1-\lambda^2} U_{\lfloor nt \rfloor - 1} V_{\lfloor nt \rfloor - 1}^2 + \mathrm{O}(n)$$

$$+ \text{ linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}, \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}^2, \ \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Using (A.5) with $(\ell, i, j) = (8, 1, 2)$ we have

$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \mid \mathcal{F}_{k-2}) \right] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (5.16), it suffices to prove

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0, \tag{5.17}$$

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.18}$$

$$n^{-3} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} V_{\lfloor nt \rfloor}^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3/2} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0 \tag{5.19}$$

as $n \to \infty$. Using (A.3) with $(\ell, i, j) = (2, 1, 1)$; $(\ell, i, j) = (4, 0, 2)$; $(\ell, i, j) = (2, 1, 0)$, and $(\ell, i, j) = (2, 0, 1)$, we have (5.17) and (5.18). By (A.4) with $(\ell, i, j) = (4, 1, 2)$, and $(\ell, i, j) = (4, 1, 0)$, we have (5.19). Thus we conclude (5.16), and hence (5.14). By Lemma A.1 and (A.3) with $(\ell, i, j) = (2, 1, 1)$ and $(\ell, i, j) = (2, 1, 0)$, we get

$$n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} - \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\boldsymbol{V}_{\boldsymbol{\xi}}} \right\| \xrightarrow{\mathbb{P}} 0 \tag{5.20}$$

as $n \to \infty$ for all $T > 0$. As a last step, using (5.7), we obtain (5.9).

Finally, we check the conditional Lindeberg condition

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1} \right) \xrightarrow{\mathbb{P}} 0 \qquad \text{for all } \theta > 0 \ \text{ and } \ T > 0. \tag{5.21}$$

We have $\mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1} \right) \leqslant \theta^{-2} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^4 \mid \mathcal{F}_{k-1} \right)$ and

$$\|\boldsymbol{Z}_k^{(n)}\|^4 \leqslant 3 \left(n^{-4} + n^{-8} U_{k-1}^4 + n^{-6} V_{k-1}^4 \right) \|\boldsymbol{M}_{k-1}\|^4.$$

Hence

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \right) \to 0 \qquad \text{as } n \to \infty \ \text{ for all } \theta > 0 \ \text{ and } \ T > 0,$$

since $\mathbb{E}(\|\boldsymbol{M}_k\|^4) = \mathrm{O}(k^2)$, $\mathbb{E}(\|\boldsymbol{M}_k\|^4 U_{k-1}^4) \leqslant \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8) \mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)$ and $\mathbb{E}(\|\boldsymbol{M}_k\|^4 V_{k-1}^4) \leqslant \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8) \mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)$ by Corollary A.3. This yields (5.21). $\qquad \square$

We call the attention to the fact that our eighth order moment conditions $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8) < \infty$ are used for applying Corollary A.3.

Appendices

A Estimations of moments

In the proof of Theorem 3.1, good bounds for moments of the random vectors and variables $(\boldsymbol{M}_k)_{k\in\mathbb{Z}_+}$, $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$, $(U_k)_{k\in\mathbb{Z}_+}$ and $(V_k)_{k\in\mathbb{Z}_+}$ are extensively used. First note that, for all $k \in \mathbb{N}$, $\mathbb{E}(\boldsymbol{M}_k \,|\, \mathcal{F}_{k-1}) = \boldsymbol{0}$ and $\mathbb{E}(\boldsymbol{M}_k) = \boldsymbol{0}$, since $\boldsymbol{M}_k = \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1})$. We present these results without proofs as they can be proven the same way as in Appendix B of [9].

A.1 Lemma. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration and with* $\boldsymbol{X}_0 = \boldsymbol{0}$. *If* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2) < \infty$ *and* $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2) < \infty$ *then*

$$\mathrm{Var}(\boldsymbol{M}_k \,|\, \mathcal{F}_{k-1}) = X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}} = U_{k-1}\overline{\boldsymbol{V}_{\boldsymbol{\xi}}} + V_{k-1}\widetilde{\boldsymbol{V}_{\boldsymbol{\xi}}} + \boldsymbol{V}_{\boldsymbol{\varepsilon}} \tag{A.1}$$

for all $k \in \mathbb{N}$, *where*

$$\widetilde{\boldsymbol{V}_{\boldsymbol{\xi}}} := \sum_{i=1}^{2}\langle e_i, \boldsymbol{v}_{\mathrm{right}}\rangle \boldsymbol{V}_{\boldsymbol{\xi}_i} = \frac{\beta\boldsymbol{V}_{\boldsymbol{\xi}_1} - (1-\delta)\boldsymbol{V}_{\boldsymbol{\xi}_2}}{\beta + 1 - \delta}.$$

If $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^3) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^3) < \infty$ *and* $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^3) < \infty$ *then, for all* $k \in \mathbb{N}$,

$$\mathbb{E}(\boldsymbol{M}_k^{\otimes 3} \,|\, \mathcal{F}_{k-1}) = X_{k-1,1}\,\mathbb{E}[(\boldsymbol{\xi}_{1,1,1} - \mathbb{E}(\boldsymbol{\xi}_{1,1,1})^{\otimes 3}]$$
$$+ X_{k-1,2}\,\mathbb{E}[(\boldsymbol{\xi}_{1,1,2} - \mathbb{E}(\boldsymbol{\xi}_{1,1,2})^{\otimes 3}] + \mathbb{E}[(\boldsymbol{\varepsilon}_1 - \mathbb{E}(\boldsymbol{\varepsilon}_1)^{\otimes 3}]. \tag{A.2}$$

A.2 Lemma. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration such that* $\alpha, \delta \in [0,1)$ *and* $\beta, \gamma \in (0, \infty)$ *with* $\alpha + \delta > 0$ *and* $\beta\gamma = (1-\alpha)(1-\delta)$ *(hence it is critical and and positively regular). Suppose* $\boldsymbol{X}_0 = \boldsymbol{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then* $\mathbb{E}(\|\boldsymbol{X}_k\|^\ell) = \mathrm{O}(k^\ell)$, *i.e.*, $\sup_{k\in\mathbb{N}} k^{-\ell}\,\mathbb{E}(\|\boldsymbol{X}_k\|^\ell) < \infty$.

A.3 Corollary. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a critical, positively regular 2-type Galton–Watson process. Suppose* $\boldsymbol{X}_0 = \boldsymbol{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then*

$$\mathbb{E}(\|\boldsymbol{X}_k\|^i) = \mathrm{O}(k^i), \qquad \mathbb{E}(\boldsymbol{M}_k^{\otimes i}) = \mathrm{O}(k^{\lfloor i/2 \rfloor}), \qquad \mathbb{E}(U_k^i) = \mathrm{O}(k^i), \qquad \mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$$

for $i, j \in \mathbb{Z}_+$ *with* $i \leqslant \ell$ *and* $2j \leqslant \ell$.

A.4 Corollary. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a critical, positively regular 2-type Galton–Watson process. Suppose* $\boldsymbol{X}_0 = \boldsymbol{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then*

(i) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \lfloor \ell/2 \rfloor$, *and for all* $\kappa > i + \frac{j}{2} + 1$, *we have*

$$n^{-\kappa} \sum_{k=1}^{n} |U_k^i V_k^j| \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty, \tag{A.3}$$

(ii) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \ell$, *for all* $T > 0$, *and for all* $\kappa > i + \frac{j}{2} + \frac{i+j}{\ell}$, *we have*

$$n^{-\kappa} \sup_{t\in[0,T]} |U_{\lfloor nt \rfloor}^i V_{\lfloor nt \rfloor}^j| \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty, \tag{A.4}$$

(iii) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \lfloor \ell/4 \rfloor$, *for all* $T > 0$, *and for all* $\kappa > i + \frac{j}{2} + \frac{1}{2}$, *we have*

$$n^{-\kappa} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} [U_k^i V_k^j - \mathbb{E}(U_k^i V_k^j \,|\, \mathcal{F}_{k-1})] \right| \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty. \tag{A.5}$$

A.5 Remark. In some parts of this paper we need the above statements with a smaller κ than it is provided by Corollary A.4. Fortunately we can use a decomposition argument in those cases to sharpen the statements, for example see the proof of (5.9) and Lemma B.2.

B CLS estimators

For each $n \in \mathbb{N}$, a CLS estimator $\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)}$ of $\boldsymbol{m}_{\boldsymbol{\xi}}$ based on a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ can be obtained by minimizing the sum of squares

$$\sum_{k=1}^{n} \left\| \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \mid \mathcal{F}_{k-1}) \right\|^2 = \sum_{k=1}^{n} \left\| \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

with respect to $\boldsymbol{m}_{\boldsymbol{\xi}}$ over $\mathbb{R}^{2 \times 2}$. In what follows, we use the notation $\boldsymbol{x}_0 := \boldsymbol{0}$. For all $n \in \mathbb{N}$, we define the function $Q_n : (\mathbb{R}^2)^n \times \mathbb{R}^{2 \times 2} \to \mathbb{R}$ by

$$Q_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n; \boldsymbol{m}'_{\boldsymbol{\xi}}) := \sum_{k=1}^{n} \left\| \boldsymbol{x}_k - \boldsymbol{m}'_{\boldsymbol{\xi}} \boldsymbol{x}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

for all $\boldsymbol{m}'_{\boldsymbol{\xi}} \in \mathbb{R}^{2 \times 2}$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$. By definition, for all $n \in \mathbb{N}$, a CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is a measurable function $F_n : (\mathbb{R}^2)^n \to \mathbb{R}^{2 \times 2}$ such that

$$Q_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n; F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) = \inf_{\boldsymbol{m}'_{\boldsymbol{\xi}} \in \mathbb{R}^{2 \times 2}} Q_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n; \boldsymbol{m}'_{\boldsymbol{\xi}})$$

for all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$. Next we give the solutions of this extremum problem.

B.1 Lemma. *For each $n \in \mathbb{N}$, any CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is a measurable function $F_n : (\mathbb{R}^2)^n \to \mathbb{R}^{2 \times 2}$ for which*

$$F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = H_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) G_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^{-1} \tag{B.1}$$

on the set

$$\left\{ (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in (\mathbb{R}^2)^n : \det(G_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) > 0 \right\},$$

where

$$G_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) := \sum_{k=1}^{n} \boldsymbol{x}_{k-1} \boldsymbol{x}_{k-1}^{\top}, \qquad H_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) := \sum_{k=1}^{n} (\boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}) \boldsymbol{x}_{k-1}^{\top}.$$

For the existence of these CLS estimators in case of a critical symmetric 2-type Galton–Watson process, i.e., when $\varrho = 1$, we need the following approximations.

B.2 Lemma. *Suppose that the assumptions of Theorem 3.1 hold. Then for each $T > 0$,*

$$n^{-5/2} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \right| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as} \ \ n \to \infty.$$

Proof. The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}$ as a sum of a martingale and some other terms. Using the recursions (4.7), (4.4) and Lemma A.1, we obtain

$$\mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2}) = \mathbb{E}\Big(\big(U_{k-2} + \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \big) \big(\lambda V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \big) \Big| \mathcal{F}_{k-2} \Big)$$

$$= \lambda U_{k-2} V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle U_{k-2} + \lambda \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle V_{k-2} + \boldsymbol{u}_{\text{left}}^{\top} \boldsymbol{m}_{\boldsymbol{\varepsilon}} \boldsymbol{m}_{\boldsymbol{\varepsilon}}^{\top} \boldsymbol{v}_{\text{left}}$$

$$+ \boldsymbol{u}^{\top} \mathbb{E}(\boldsymbol{M}_{k-1} \boldsymbol{M}_{k-1}^{\top} \mid \mathcal{F}_{k-2}) \boldsymbol{v}$$

$$= \lambda U_{k-2} V_{k-2} + \text{constant} + \text{linear combination of } U_{k-2} \text{ and } V_{k-2}.$$

Thus

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} = \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1} V_{k-1} - \mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2}) \big] + \sum_{k=2}^{\lfloor nt \rfloor} \mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2})$$

$$= \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1} V_{k-1} - \mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2}) \big] + \lambda \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}$$

$$+ \mathrm{O}(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Consequently

$$\sum_{k=2}^{\lfloor nt \rfloor} U_{k-1}V_{k-1} = \frac{1}{1-\lambda}\sum_{k=2}^{\lfloor nt \rfloor}\big[U_{k-1}V_{k-1} - \mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]$$

$$- \frac{\lambda}{1-\lambda}U_{\lfloor nt \rfloor -1}V_{\lfloor nt \rfloor -1} + \mathrm{O}(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Using (A.5) with $(\ell, i, j) = (4, 1, 1)$ we have

$$n^{-5/2}\sup_{t\in[0,T]}\left|\sum_{k=2}^{\lfloor nt \rfloor}\big[U_{k-1}V_{k-1} - \mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]\right| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show the statement, it suffices to prove

$$n^{-5/2}\sum_{k=1}^{\lfloor nT \rfloor} U_k \overset{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-5/2}\sum_{k=1}^{\lfloor nT \rfloor} |V_k| \overset{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-5/2}\sup_{t\in[0,T]}|U_{\lfloor nt \rfloor}V_{\lfloor nt \rfloor}| \overset{\mathbb{P}}{\longrightarrow} 0 \tag{B.2}$$

as $n \to \infty$. Using (A.3) with $(\ell, i, j) = (2, 1, 0)$ and $(\ell, i, j) = (2, 0, 1)$, and (A.4) with $(\ell, i, j) = (3, 1, 1)$ we have (B.2), thus we conclude the statement. $\qquad\square$

Using the same ideas as above one can prove the following.

B.3 Lemma. *Suppose that the assumptions of Theorem 3.1 hold. For each $T > 0$, we have*

$$n^{-2}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt \rfloor} V_k^2 - \frac{\langle \overline{V}_{\xi}\,v_{\text{left}}, v_{\text{left}}\rangle}{1-\lambda^2}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}\right| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad as \ \ n \to \infty.$$

Now we can prove asymptotic existence and uniqueness of CLS estimators of the offspring mean matrix and of the criticality parameter.

B.4 Proposition. *Suppose that the assumptions of Theorem 3.1 hold, and $\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle + \langle V_{\varepsilon}v_{\text{left}}, v_{\text{left}}\rangle + \langle v_{\text{left}}, m_{\varepsilon}\rangle^2 > 0$. Then $\lim_{n\to\infty}\mathbb{P}(\Omega_n) = 1$, where Ω_n is defined in (3.2), and hence the probability of the existence of a unique CLS estimator $\widehat{m}_{\xi}^{(n)}$ converges to 1 as $n \to \infty$, and this CLS estimator has form given in (3.1) on the set Ω_n. If $\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle > 0$ then $\lim_{n\to\infty}\mathbb{P}(\widetilde{\Omega}_n) = 1$, where $\widetilde{\Omega}_n$ is defined in (3.5), and hence the probability of the existence of the estimator $\widehat{\varrho}_n$ converges to 1 as $n \to \infty$.*

Proof. Recall convergence $(n^{-1}U_{\lfloor nt \rfloor})_{t\in\mathbb{R}_+} \overset{\mathcal{D}}{\longrightarrow} (\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ from (4.6). Using Lemmas B.3, B.2, and a version of the continuous mapping theorem (see [10], Lemma 3.1), one can show

$$\sum_{k=1}^{n}\begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-5/2}U_{k-1}V_{k-1} \\ n^{-2}V_{k-1}^2 \end{bmatrix} \overset{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t \\ 0 \\ \frac{\langle v_{\text{left}}, m_{\varepsilon}\rangle}{1-\lambda}\int_0^1 \mathcal{Y}_t\,\mathrm{d}t \end{bmatrix} \qquad \text{as } n \to \infty.$$

By (4.11) and continuous mapping theorem,

$$n^{-5}\det(A_n) \overset{\mathcal{D}}{\longrightarrow} \frac{\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle}{1-\lambda^2}\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t \qquad \text{as } n \to \infty. \tag{B.3}$$

Since $m_{\varepsilon} \neq 0$, by the SDE (2.8), we have $\mathbb{P}(\mathcal{Y}_t = 0 \text{ for all } t \in [0,1]) = 0$, which implies that $\mathbb{P}\big(\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t > 0\big) = 1$. Consequently, the distribution function of $\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t$ is continuous at 0.

If $\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle > 0$ then, by (B.3),

$$\mathbb{P}(\Omega_n) = \mathbb{P}\left(\det(A_n) > 0\right) = \mathbb{P}\left(n^{-5}\det(A_n) > 0\right)$$

$$\to \mathbb{P}\left(\frac{\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle}{1-\lambda^2}\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t > 0\right) = \mathbb{P}\left(\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t > 0\right) = 1$$

as $n \to \infty$.

If $\langle\overline{V}_{\xi}v_{\text{left}}, v_{\text{left}}\rangle > 0$, then (3.6) yields $\widehat{m}_{\xi}^{(n)} \overset{\mathcal{D}}{\longrightarrow} m_{\xi}$ as $n \to \infty$, and hence $\widehat{m}_{\xi}^{(n)} \overset{\mathbb{P}}{\longrightarrow} m_{\xi}$ as $n \to \infty$, thus $(\widehat{\alpha}_n - \widehat{\delta}_n)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n \overset{\mathbb{P}}{\longrightarrow} (\alpha - \delta)^2 + 4\beta\gamma = (1-\lambda)^2 > 0$, implying

$$\mathbb{P}(\widetilde{\Omega}_n) = \mathbb{P}\left((\widehat{\alpha}_n - \widehat{\delta}_n)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n \geqslant 0\right) \to 1 \qquad \text{as } n \to \infty,$$

hence we obtain the satement. $\qquad\square$

References

1. ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*, Springer-Verlag, New York-Heidelberg.
2. BARCZY, M., ISPÁNY, M. and PAP, G. (2011). Asymptotic behavior of unstable INAR(p) processes. *Stochastic Processes and their Applications* **121(3)** 583–608.
3. BARCZY, M., ISPÁNY, M. and PAP, G. (2014). Asymptotic behavior of conditional least squares estimators for unstable integer-valued autoregressive models of order 2. *Scandinavian Journal of Statistics* **41(4)** 866–892.
4. COX, J.C., J.E. INGERSOLL AND S.A. ROSS (1985). A theory of the term structure of interest rates. *Econometrica* **53(2)** 385-407.
5. DANKA, T. and PAP, G. (2016+). Asymptotic behavior of critical indecomposable multi-type branching processes with immigration. To appear in *European Series in Applied and Industrial Mathematics (ESAIM). Probability and Statistics.* Available on the ArXiv: `http://arxiv.org/abs/1401.3440`.
6. HACCOU, P., JAGERS, P. and VATUTIN, V. (2005). *Branching Processes*, Cambridge University Press, Cambridge.
7. HORN, R. A. and JOHNSON, CH. R. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.
8. IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo.
9. ISPÁNY, M., KÖRMENDI, K. and PAP, G. (2014). Asymptotic behavior of CLS estimators for 2-type doubly symmetric critical Galton–Watson processes with immigration. *Bernoulli* **20(4)** 2247–2277.
10. ISPÁNY, M. and PAP, G. (2010). A note on weak convergence of step processes. *Acta Mathematica Hungarica* **126(4)** 381–395.
11. ISPÁNY, M. and PAP, G. (2014). Asymptotic behavior of critical primitive multi-type branching processes with immigration. *Stochastic Analysis and Applications* **32(5)** 727–741.
12. ISPÁNY, M., PAP, G. and VAN ZUIJLEN, M. C. A. (2003). Asymptotic inference for nearly unstable INAR(1) models. *Journal of Applied Probability* **40(3)** 750–765.
13. KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *The Annals of Mathematical Statistics* **37(5)** 1211–1223.
14. KLIMKO, L. A. and NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* **6(3)** 629–642.
15. KÖRMENDI, K. and PAP, G. (2015). Statistical inference of 2-type critical Galton–Watson processes with immigration. Available on the ArXiv: `arxiv.org/abs/1502.04900`.
16. PUTZER, E. J. (1966). Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients. *The American Mathematical Monthly* **73(1)** 2–7.
17. QUINE, M. P. (1970). The multi-type Galton–Watson process with immigration. *Journal of Applied Probability* **7(2)** 411–422.
18. Velasco, M.G., Puerto, I.M., Martínez, R., Molina, M., Mota, M., and Ramos, A. 2010. *Workshop on Branching Processes and Their Applications*, Lecture Notes in Statistics, Proceedings 197, Springer-Verlag, Berlin, Heidelberg.
19. WEI, C. Z. and WINNICKI, J. (1989). Some asymptotic results for the branching process with immigration *Stochastic Processes and their Applications* **31(2)** 261–282.
20. WEI, C. Z. and WINNICKI, J. (1990). Estimation of the means in the branching process with immigration. *The Annals of Statistics* **18** 1757–1773.
21. WINNICKI, J. (1988). Estimation theory for the branching process with immigration. In: *Statistical inference from stochastic processes* (Ithaca, NY, 1987), 301–322, Contemp. Math., 80, Amer. Math. Soc., Providence, RI.