Estimation of the offspring means for critical 2-type Galton–Watson processes with immigration

Ph.D. Thesis

Kristóf Körmendi

Supervisor: Professor Gyula Pap

Doctoral School of Mathematics and Computer Science
University of Szeged
2018

Contents

Acknowledgements

First of all I am truly grateful to my supervisor, Professor Gyula Pap, whose enthusiasm for open problems is boundless. I learned a great deal from him.

I'd like to express my gratitude to my fellow classmates, and the people of the Bolyai Institute for cultivating an atmosphere ideal for the pursuit of knowledge.

My high school mathematics teacher Vlagyimir Volosin also deserves a special mention, his excellent lectures put me on the path of becoming a mathematician.

Finally I'm thankful to my family, especially to my mother for their love and support.

Köszönetnyilvánítás

Köszönettel tartozom témavezetőmnek, Dr. Pap Gyulának, akinek a megoldatlan problémák iránti lelkesedése határtalan, rengeteget tanultam tőle.

Szeretném a hálámat kifejezni az egyetemi évfolyamtársaimnak és a Bolyai Intézet alkalmazottainak, amiért olyan légkört alakítottak ki melyben öröm volt a tanulás.

Külön említést érdemel Volosin Vlagyimir tanár úr, a középiskolai matematikatanárom, hiszen kitűnő matematikaórái tereltek matematikusi pályára.

Végezetül hálás vagyok a családomnak, különösképpen édesanyámnak a sok szeretetért és támogatásért, amit kaptam.

1 Introduction

Branching processes have a number of applications in biology, finance, economics, and queueing theory see Haccou et al. [7]. Many aspects of applications in epidemiology, genetics, and cell kinetics were presented at the 2009 Badajoz Workshop on Branching Processes, see Velasco et al. [23].

Statistical inference for Galton–Watson processes have an extensive literature in the single type case. The problem in its present form was described by Heyde and Seneta [5], [6]. In a series of articles the authors Wei and Winnicki (see [24], [25], [27]) described the asymptotic behaviour of the estimates gained by the conditional least squares method introduced by Klimko and Nelson [15] and its modification the weighted conditional least squares method by Nelson [18]. For a more in-depth discussion of the history of parameter estimation for branching process see the excellent survey article by Winnicki [26].

Multitype Galton-Watson processes are natural generalizations of the single-type case. For a textbook introduction to multitype branching processes one should check Athreya and Ney [1, Chapter V.] or Mode [17]. Statistical inference for these models are sparsely available. The asymptotic behaviour of the process itself was described by Quine [20] in the subcritical, by Ispány and Pap [12] in the critical and finally by Kesten and Stigum [14] in the supercritical case. Quine and Durham [21] described a strongly consistent and asymptotically normal estimator for the offspring mean matrix in the subcritical case, while Shete and Sriram [22] established similar results in the supercritical case, however in the supercritical case the estimator requires more information than just the generation sizes of each type of individual. Results in the critical case was first established by Ispány et al. [10] under heavy restrictions on the structure of the offspring mean matrix, then later Körmendi and Pap [16] lifted the restrictions and described the asymptotic behaviour of an estimator in both the critical and subcritical cases.

This thesis is devoted to developing a toolkit for asymptotic study of estimators in 2-type critical Galton–Watson processes. In Section 2 we introduce the basic notations for our model. We define a criticality parameter, namely the spectral radius of the offspring mean matrix, and describe the classification of 2-type Galton–Watson processes into subcritical, critical and supercritical cases based on its value. Then we state a functional limit theorem for the process by Ispány and Pap [12]. This limit is curious, because it is degenerate in the sense that it is concentrated on a single line whose direction is determined by the right Perron vector of the offspring mean matrix.

Section 3 contains the development of the toolkit. We define a decomposition of the process based on the phenomena observed at the end of the previous section. We want to use these random variables as building blocks of any estimator whose asymptotic properties we want to investigate. In order to do that we need a firm understanding of their behaviour, so we estimate their growth as the number of generations in the underlying process tends to infinity. Our first upper bounds are too big, so in few select cases we refine them. Then we use a theorem by Ispány and Pap [11] to prove a joint limit theorem for these building blocks.

We demonstrate the applicability of this method in Section 4. First we reproduce the results in the special doubly symmetric model described in Ispány et al. [10, Theorem 3.1]. This was our first parameter estimation result for 2-type Galton–Watson processes and as such we chose a special model with heavy restrictions on the structure of the offspring mean matrix, where everything is relatively easy to calculate. By developing a better understanding of the ideas related to this decomposition we managed to tackle the general case, where we only assume that the offspring mean matrix is positively regular. These results can be found in Körmendi and Pap [16, Theorem 3.1] and are also reproduced in Subsection 4.2. We finish this section with a new result: We examine the asymptotic properties of a joint estimator of both the offspring mean matrix and the immigration mean.

2 Preliminaries

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 k^{th} generation will be denoted by $X_{k,i}$, the number of type ℓ offsprings produced by the j^{th} individual who is of type i belonging to the $(k-1)^{\mathrm{th}}$ generation will be denoted by $\xi_{k,j,i,\ell}$, and the number of type i

immigrants in the k^{th} generation will be denoted by $\varepsilon_{k,i}$. Then we have

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

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

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

and

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

Note that many authors define the offspring mean matrix as $\boldsymbol{m}_{\boldsymbol{\xi}}^\top$.

2.1 Eigenvectors of the offspring mean matrix

Our ultimate goal is to estimate the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$. In order to emphasize its importance we show how it plays a role in the asymptotic behaviour of the process. For $k \in \mathbb{Z}_+$, let $\mathcal{F}_k := \sigma\big(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\big)$. By (1),

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

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

$$\mathbb{E}(\boldsymbol{X}_k) = \boldsymbol{m}_{\boldsymbol{\xi}}^k\,\mathbb{E}(\boldsymbol{X}_0) + \sum_{j=0}^{k-1} \boldsymbol{m}_{\boldsymbol{\xi}}^j\,\boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}. \tag{3}$$

Hence, the asymptotic behaviour of the sequence $(\mathbb{E}(\boldsymbol{X}_k))_{k\in\mathbb{Z}_+}$ depends on the asymptotic behaviour 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 Frobenius–Perron theorem, e.g., Horn and Johnson [8, 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 [20]). We will write the offspring mean matrix of a 2-type Galton–Watson process with immigration in the form

$$\boldsymbol{m}_{\boldsymbol{\xi}} := \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}. \tag{4}$$

We will focus only on *positively regular* 2-type Galton–Watson processes with immigration, i.e., when there is a positive integer $k \in \mathbb{N}$ such that the entries of $\boldsymbol{m}_{\boldsymbol{\xi}}^k$ are positive (see Kesten and Stigum [14]), which is equivalent to $\beta, \gamma \in (0, \infty)$, $\alpha, \delta \in \mathbb{R}_+$ with $\alpha + \delta > 0$. Then the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ has eigenvalues

$$\lambda_+ := \frac{\alpha + \delta + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2},$$
$$\lambda_- := \frac{\alpha + \delta - \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2},$$

satisfying $\lambda_+ > 0$ and $-\lambda_+ < \lambda_- < \lambda_+$, hence the spectral radius of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is

$$\varrho = r(\boldsymbol{m}_{\boldsymbol{\xi}}) = \lambda_+ = \frac{\alpha + \delta + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2}. \tag{5}$$

By the Frobenius–Perron theorem (see, e.g., Horn and Johnson [8, Theorems 8.2.11 and 8.5.1]),

$$\lambda_+^{-k} \boldsymbol{m}_{\boldsymbol{\xi}}^k \to \boldsymbol{u}_{\mathrm{right}} \boldsymbol{u}_{\mathrm{left}}^\top \qquad \text{as} \quad k \to \infty,$$

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

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

Using the so-called Putzer's spectral formula (see, e.g., Putzer [19]), the powers of $\boldsymbol{m}_{\boldsymbol{\xi}}$ can be written in the form

$$\begin{aligned}\boldsymbol{m}_{\boldsymbol{\xi}}^k &= \frac{\lambda_+^k}{\lambda_+ - \lambda_-}\begin{bmatrix}\lambda_+ - \delta & \beta \\ \gamma & \lambda_+ - \alpha\end{bmatrix} + \frac{\lambda_-^k}{\lambda_+ - \lambda_-}\begin{bmatrix}\lambda_+ - \alpha & -\beta \\ -\gamma & \lambda_+ - \delta\end{bmatrix} \\ &= \lambda_+^k \boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top + \lambda_-^k \boldsymbol{v}_{\mathrm{right}}\boldsymbol{v}_{\mathrm{left}}^\top, \qquad k \in \mathbb{N},\end{aligned} \tag{6}$$

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

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

The process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ is critical and positively regular if and only if $\alpha, \delta \in [0, 1)$ and $\beta, \gamma \in (0, \infty)$ with $\alpha + \delta > 0$ and $\beta\gamma = (1-\alpha)(1-\delta)$, then the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ has eigenvalues $\lambda_+ = 1$ and

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

Now we explore the relations of the eigenvectors of $\boldsymbol{m}_{\boldsymbol{\xi}}$ with each other. Our first result is also known as the principle of biorthogonality, we prove it nonetheless to show its proof is different from that of Lemma 2.3.

Lemma 2.1. *Suppose that the process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ is critical and positively regular, then we have $\langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle = \langle \boldsymbol{v}_{\mathrm{right}}, \boldsymbol{u}_{\mathrm{left}}\rangle = 0$.*

Proof. Using that $\boldsymbol{u}_{\mathrm{right}}$ and $\boldsymbol{v}_{\mathrm{left}}$ are eigenvectors of the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ we get

$$\langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle = \langle \boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{u}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle = \langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{m}_{\boldsymbol{\xi}}^\top\boldsymbol{v}_{\mathrm{left}}\rangle = \lambda_-\langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle.$$

Since $\lambda_- \neq 1$ this concludes $\langle \boldsymbol{u}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle = 0$. The proof of $\langle \boldsymbol{v}_{\mathrm{right}}, \boldsymbol{u}_{\mathrm{left}}\rangle = 0$ can be carried out similarly. □

Lemma 2.2. *Suppose that the process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ is critical and positively regular, then we have*

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

Proof. We can calculate the determinant the following way

$$\det\left(\begin{bmatrix}\boldsymbol{u}_{\mathrm{right}} & \boldsymbol{v}_{\mathrm{right}}\end{bmatrix}\right) = \boldsymbol{u}_{\mathrm{right}}^\top\begin{bmatrix}0 & 1 \\ -1 & 0\end{bmatrix}\boldsymbol{v}_{\mathrm{right}} = \boldsymbol{u}_{\mathrm{right}}^\top\boldsymbol{u}_{\mathrm{left}} = 1.$$

□

Lemma 2.3. *Suppose that the process* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *is critical and positively regular, then we have* $\langle \boldsymbol{v}_{\mathrm{right}}, \boldsymbol{v}_{\mathrm{left}}\rangle = 1$.

Proof. Using that $\beta\gamma = (1-\alpha)(1-\delta)$ we get

$$\begin{aligned}
\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{right}}\rangle &= \left\langle \frac{1}{2-\alpha-\delta}\begin{bmatrix} -\beta-1+\alpha \\ \gamma+1-\delta \end{bmatrix}, \frac{1}{\beta+1-\alpha}\begin{bmatrix} -1+\alpha \\ \beta \end{bmatrix}\right\rangle \\
&= \frac{(-\beta-1+\alpha)(-1+\alpha)+(\gamma+1-\delta)\beta}{(2-\alpha-\delta)(\beta+1-\alpha)} \\
&= \frac{(\beta+1-\alpha)(1-\alpha)+\beta\gamma+(1-\delta)\beta}{(2-\alpha-\delta)(\beta+1-\alpha)} \\
&= \frac{(\beta+1-\alpha)(1-\alpha)+(1-\alpha)(1-\delta)+(1-\delta)\beta}{(2-\alpha-\delta)(\beta+1-\alpha)} = 1.
\end{aligned}$$

□

2.2 A limit theorem for the process

Next we will recall a convergence result for critical and positively regular 2-type Galton–Watson processes with immigration. A function $f : \mathbb{R}_+ \to \mathbb{R}^d$ is called *càdlàg* if it is right continuous with left limits. Let $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathbb{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg and continuous functions on $\mathbb{R}_+$, respectively. Let $\mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^d)$ denote the Borel σ-field in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ for the metric characterized by Jacod and Shiryaev [13, VI.1.15] (with this metric $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ is a complete and separable metric space). For $\mathbb{R}^d$-valued stochastic processes $(\boldsymbol{\mathcal{Y}}_t)_{t\in\mathbb{R}_+}$ and $(\boldsymbol{\mathcal{Y}}_t^{(n)})_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, with càdlàg paths we write $\boldsymbol{\mathcal{Y}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{Y}}$ as $n \to \infty$ if the distribution of $\boldsymbol{\mathcal{Y}}^{(n)}$ on the space $(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^d))$ converges weakly to the distribution of $\boldsymbol{\mathcal{Y}}$ on the space $(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^d))$ as $n \to \infty$. Concerning the notation $\stackrel{\mathcal{D}}{\longrightarrow}$ we note that if ξ and ξ_n, $n \in \mathbb{N}$, are random elements with values in a metric space (E, ρ), then we also denote by $\xi_n \stackrel{\mathcal{D}}{\longrightarrow} \xi$ the weak convergence of the distributions of ξ_n on the space $(E, \mathcal{B}(E))$ towards the distribution of ξ on the space $(E, \mathcal{B}(E))$ as $n \to \infty$, where $\mathcal{B}(E)$ denotes the Borel σ-algebra on E induced by the given metric ρ.

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 [12, Theorem 3.1].

Theorem. 2.4. *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*

$$\left(\boldsymbol{\mathcal{X}}_t^{(n)}\right)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{D}} (\boldsymbol{\mathcal{X}}_t)_{t\in\mathbb{R}_+} := (\mathcal{Z}_t \boldsymbol{u}_{\mathrm{right}})_{t\in\mathbb{R}_+} \tag{7}$$

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

$$\begin{aligned} \mathrm{d}\mathcal{Z}_t &= \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \,\mathrm{d}t + \sqrt{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{u}_{\mathrm{left}} \rangle \mathcal{Z}_t^+}\,\mathrm{d}\mathcal{W}_t, \quad t\in\mathbb{R}_+, \\ \mathcal{Z}_0 &= 0, \end{aligned} \tag{8}$$

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

$$\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} := \sum_{i=1}^2 \langle \boldsymbol{e}_i, \boldsymbol{u}_{\mathrm{right}} \rangle \boldsymbol{V}_{\boldsymbol{\xi}_i} = \frac{\beta \boldsymbol{V}_{\boldsymbol{\xi}_1} + (1-\alpha)\boldsymbol{V}_{\boldsymbol{\xi}_2}}{\beta+1-\alpha} \tag{9}$$

is a mixed offspring variance matrix.

In fact, in Ispány and Pap [12, Theorem 3.1], the above result has been proved under the higher moment assumptions

$$\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^4) < \infty, \qquad \mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^4) < \infty, \qquad \mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^4) < \infty,$$

which have been relaxed in Danka and Pap [4, Theorem 3.1].

Remark 2.5. *The SDE* (8) *has a unique strong solution* $(\mathcal{Z}_t^{(z)})_{t\in\mathbb{R}_+}$ *for all initial values* $\mathcal{Z}_0^{(z)} = z \in \mathbb{R}$*, and if* $z \geq 0$*, then* $\mathcal{Z}_t^{(z)}$ *is nonnegative for all* $t\in\mathbb{R}_+$ *with probability one, hence* $\mathcal{Z}_t^+$ *may be replaced by* $\mathcal{Z}_t$ *under the square root in* (8)*, see, for example, Ikeda and Watanbe [9, Chapter IV, Example 8.2].*

In this section we have introduced a number of assumptions on the process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ For the sake of easier reference we collect those assumptions here. First a condition that guarantees that our process is critical and positively regular. The process satisfies the *criticality condition* if

$$\alpha, \delta \in [0,1), \ \beta, \gamma \in (0,\infty), \ \alpha+\delta > 0, \ \beta\gamma = (1-\alpha)(1-\delta). \tag{CPR}$$

Then we have a condition that we start from an empty initial population, that is $\boldsymbol{X}_0 = \boldsymbol{0}$. If we don't want to be stuck in $\boldsymbol{0}$ we have to assume that the immigration distribution isn't degenerate $\mathbf{0}$, it is sufficient to assume $\boldsymbol{m}_{\varepsilon} \neq \mathbf{0}$ for this. The process satisfies the *zero start condition* if

$$\boldsymbol{X}_0 = \mathbf{0}, \qquad \boldsymbol{m}_{\varepsilon} \neq \mathbf{0}. \tag{ZS}$$

Next we have a condition on the moments of the process, where we assume the finiteness of ℓ^{th} moments of both the offspring and immigration distributions. This in terms implies the finiteness of the ℓ^{th} moment of the process itself. The process satisfies the *moment condition* for some $\ell \in \mathbb{N}$ if

$$\mathbb{E}\left(\|\boldsymbol{\xi}_{1,1,1}\|^{\ell}\right) < \infty, \qquad \mathbb{E}\left(\|\boldsymbol{\xi}_{1,1,2}\|^{\ell}\right) < \infty, \qquad \mathbb{E}\left(\|\boldsymbol{\varepsilon}_{1}\|^{\ell}\right) < \infty. \tag{M}$$

Finally we have a condition that doesn't appear in this section however it will be necessary later. The process satisfies the *non degeneracy condition* if

$$\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle \neq 0. \tag{ND}$$

The reason for this condition can be understood if one looks at Lemma 3.11, as that describes a relation between the two parts of the upcoming decomposition.

By the Frobenius–Perron theorem $\boldsymbol{u}_{\text{right}}$ is a vector whose coordinates are all positive and add up to 1, hence $\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}$ defined in (9) is a convex combination of the offspring covariance matrices and as such it is a positive semidefinite matrix. Therefore

$$\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle \geq 0,$$

so when (ND) doesn't hold, we have

$$\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle = 0.$$

One can easily check the following

$$\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle = \sum_{i=1}^{2} \langle \boldsymbol{e}_1, \boldsymbol{u}_{\text{right}} \rangle \operatorname{Var}\left(\langle \boldsymbol{v}_{\text{left}}, \boldsymbol{\xi}_{1,1,i} \rangle\right),$$

consequently

$$\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle = 0 \quad \Longrightarrow \quad \operatorname{Var}\left(\langle \boldsymbol{v}_{\text{left}}, \boldsymbol{\xi}_{1,1,i} \rangle\right) = 0, \quad i = 1, 2.$$

So when the non-degeneracy condition fails, then both offspring distributions are degenerate. In this thesis we prove results under (ND), however we note, that Körmendi and Pap [16] contains some results under the degenerate case as well.

2.3 Insight from a failed attempt

With Theorem 2.4 in hand we will try to use the continuous mapping theorem to analyse an estimator. This attempt will fail, but understanding why it fails will give us clues to which direction should we continue our research.

As a method of estimation we are going to use the conditional least squares method, however we will not go into much detail here, since Section 4.2 contains not only the construction of the estimator but also a successful attempt at describing its limiting behaviour. The difference of the least squares estimate and the real offspring mean matrix can be written as

$$\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} = \boldsymbol{C}_n \boldsymbol{A}_n^{-1},$$

where

$$\boldsymbol{A}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{X}_{k-1} \boldsymbol{X}_{k-1}^\top,$$
$$\boldsymbol{C}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{M}_k \boldsymbol{X}_{k-1}^\top.$$

For a detailed proof see Lemma 4.5 and Corollary 4.6. We will focus on the asymptotic behaviour of the matrix $\boldsymbol{A}_n$. By Theorem 2.4 and the continuous mapping theorem we have

$$n^{-3}\boldsymbol{A}_n = \int_0^1 \boldsymbol{\mathcal{X}}_t^{(n)} \left(\boldsymbol{\mathcal{X}}_t^{(n)}\right)^\top dt \xrightarrow{\mathcal{D}} \int_0^1 \mathcal{Z}_t^2 \,\mathrm{d}t\, \boldsymbol{u}_{\mathrm{right}} \boldsymbol{u}_{\mathrm{right}}^\top =: \boldsymbol{\mathcal{A}}.$$

Unfortunately the matrix $\boldsymbol{\mathcal{A}}$ is non-invertible, since $\det(\boldsymbol{\mathcal{A}}) = 0$. So a straight up application of the continuous mapping theorem fails to find the non-zero limit that we are looking for. The correct way to look at this is that Theorem 2.4 is incomplete, we need something more. In the following section we are introducing a decomposition of the process and based on that, we prove a limit theorem that is an extension of Theorem 2.4 allowing us to examine the estimators successfully.

3 A toolkit for asymptotic study of estimates

3.1 A decomposition of 2-type Galton–Watson processes

In the previous section we saw that the eigenvectors of the matrix $\boldsymbol{m_\xi}$ play an important role in the asymptotic behaviour of the process itself. It is curious in Theorem 2.4 that the limit of a 2-dimensional process is degenerate in the sense that it is concentrated on a single line whose direction is determined by $\boldsymbol{u}_{\mathrm{right}}$. In this section we define a decomposition of the process based on the eigenvectors of $\boldsymbol{m_\xi}$.

Applying (2), let us introduce the sequence

$$\boldsymbol{M}_k := \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \mid \mathcal{F}_{k-1}) = \boldsymbol{X}_k - \boldsymbol{m_\xi}\boldsymbol{X}_{k-1} - \boldsymbol{m_\varepsilon}, \qquad k \in \mathbb{N}, \tag{10}$$

of martingale differences with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$. By (10), the process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ satisfies the recursion

$$\boldsymbol{X}_k = \boldsymbol{m_\xi}\boldsymbol{X}_{k-1} + \boldsymbol{m_\varepsilon} + \boldsymbol{M}_k, \qquad k \in \mathbb{N}. \tag{11}$$

We derive a useful decomposition for X_k, $k \in \mathbb{N}$. Let us introduce the sequence

$$U_k := \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{X}_k \rangle = \frac{(\gamma + 1 - \delta)X_{k,1} + (\beta + 1 - \alpha)X_{k,2}}{1 - \lambda_-}, \qquad k \in \mathbb{Z}_+. \tag{12}$$

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

$$U_k = U_{k-1} + \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m_\varepsilon} \rangle + \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{M}_k \rangle, \qquad k \in \mathbb{N}, \tag{13}$$

since $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m_\xi}\boldsymbol{X}_{k-1} \rangle = \boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{m_\xi}\boldsymbol{X}_{k-1} = \boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{X}_{k-1} = U_{k-1}$, because $\boldsymbol{u}_{\mathrm{left}}$ is a left eigenvector of the mean matrix $\boldsymbol{m_\xi}$ belonging to the eigenvalue 1. Hence $(U_k)_{k\in\mathbb{Z}_+}$ is a nonnegative unstable AR(1) process with positive drift $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m_\varepsilon} \rangle$ and with heteroscedastic innovation $(\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{M}_k \rangle)_{k\in\mathbb{N}}$. Note that the solution of the recursion (13) is

$$U_k = \sum_{j=1}^{k} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{M}_j + \boldsymbol{m_\varepsilon} \rangle, \qquad k \in \mathbb{N}, \tag{14}$$

and applying the continuous mapping theorem to Theorem 2.4 yields

$$(n^{-1}U_{\lfloor nt \rfloor})_{t\in\mathbb{R}_+} = (\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{X}}_t^{(n)} \rangle)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{D}} (\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{X}}_t \rangle)_{t\in\mathbb{R}_+} = (\mathcal{Z}_t)_{t\in\mathbb{R}_+} \tag{15}$$

as $n \to \infty$, where $(\mathcal{Z}_t)_{t\in\mathbb{R}_+}$ is the pathwise unique strong solution of the SDE (8). We could think of the variables $(U_k)_{k\in\mathbb{Z}_+}$ as the *well behaved* part of our decomposition, because they allow us to get the underlying 1-dimensional stochastic process in Theorem 2.4 . Moreover, let

$$V_k := \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{X}_k \rangle = \frac{-(1-\alpha)X_{k,1} + \beta X_{k,2}}{\beta + 1 - \alpha}, \qquad k \in \mathbb{Z}_+. \tag{16}$$

Note that we have

$$V_k = \lambda_- V_{k-1} + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_k \rangle, \qquad k \in \mathbb{N}, \tag{17}$$

since $\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} \rangle = \boldsymbol{v}_{\mathrm{left}}^\top \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} = \lambda_- \boldsymbol{v}_{\mathrm{left}}^\top \boldsymbol{X}_{k-1} = \lambda_- V_{k-1}$, because $\boldsymbol{v}_{\mathrm{left}}$ 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 $\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$ and with heteroscedastic innovation $(\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_k \rangle)_{k\in\mathbb{N}}$. Note that the solution of the recursion (17) is

$$V_k = \sum_{j=1}^{k} \lambda_-^{k-j} \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_j + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle, \qquad k \in \mathbb{N}, \tag{18}$$

and applying the continuous mapping theorem to Theorem 2.4 yields

$$(n^{-1} V_{\lfloor nt \rfloor})_{t\in\mathbb{R}_+} = (\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{\mathcal{X}}_t^{(n)} \rangle)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{D}} (\langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{\mathcal{X}}_t \rangle)_{t\in\mathbb{R}_+} = 0.$$

We could think of the variables $(V_k)_{k\in\mathbb{Z}_+}$ as the *problematic* part of our decomposition, because the continuous mapping theorem does not find the nonzero limit of them, since the scaling is incorrect. By (1) and (10), we obtain the decomposition

$$\boldsymbol{M}_k = \sum_{j=1}^{X_{k-1,1}} \big(\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1})\big) + \sum_{j=1}^{X_{k-1,2}} \big(\boldsymbol{\xi}_{k,j,2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2})\big) + \big(\boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k)\big), \tag{19}$$

for all $k \in \mathbb{N}$. The recursion (11) 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 (6),

$$\begin{aligned}
\boldsymbol{X}_k &= \sum_{j=1}^{k} \left(\boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top + \lambda_-^{k-j}\boldsymbol{v}_{\mathrm{right}}\boldsymbol{v}_{\mathrm{left}}^\top\right)(\boldsymbol{m}_{\boldsymbol{\varepsilon}} + \boldsymbol{M}_j) \\
&= \boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top \sum_{j=1}^{k}(\boldsymbol{X}_j - \boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{j-1}) + \boldsymbol{v}_{\mathrm{right}}\boldsymbol{v}_{\mathrm{left}}^\top \sum_{j=1}^{k} \lambda_-^{k-j}(\boldsymbol{X}_j - \boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{j-1}) \\
&= \boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top \sum_{j=1}^{k}(\boldsymbol{X}_j - \boldsymbol{X}_{j-1}) + \boldsymbol{v}_{\mathrm{right}}\boldsymbol{v}_{\mathrm{left}}^\top \sum_{j=1}^{k} \left[\lambda_-^{k-j}\boldsymbol{X}_j - \lambda_-^{k-j+1}\boldsymbol{X}_{j-1}\right] \\
&= \boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top\boldsymbol{X}_k + \boldsymbol{v}_{\mathrm{right}}\boldsymbol{v}_{\mathrm{left}}^\top\boldsymbol{X}_k = U_k\boldsymbol{u}_{\mathrm{right}} + V_k\boldsymbol{v}_{\mathrm{right}},
\end{aligned}$$

hence

$$\boldsymbol{X}_k = \begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}} & \boldsymbol{v}_{\mathrm{right}} \end{bmatrix} \begin{bmatrix} U_k \\ V_k \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\beta+1-\alpha}U_k - \frac{\beta+1-\alpha}{1-\lambda_-}V_k \\ \frac{1-\alpha}{\beta+1-\alpha}U_k + \frac{\gamma+1-\delta}{1-\lambda_-}V_k \end{bmatrix}, \tag{20}$$

for all $k \in \mathbb{Z}_+$.

We want to use this decomposition as a tool to investigate asymptotic properties of various estimators of the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$. Any estimator based on the sample $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ can be rewritten in terms of the variables $U_1, \ldots, U_n, V_1, \ldots, V_n$, thus a good understanding of their behaviour can give us insight into the behaviour of the estimator itself. We note that this reformulation of an estimator is strictly a theoretical tool to prove theorems about it, as without knowing $\boldsymbol{m}_{\boldsymbol{\xi}}$ we also don't know $\boldsymbol{u}_{\mathrm{left}}$ and $\boldsymbol{v}_{\mathrm{left}}$, therefore we are unable to calculate U_k and V_k.

3.2 An estimation of moments

We want to bound the growth of $(\boldsymbol{M}_k)_{k\in\mathbb{Z}_+}$, $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$, $(U_k)_{k\in\mathbb{Z}_+}$ and $(V_k)_{k\in\mathbb{Z}_+}$ and some related expressions as $k \to \infty$. The reader will find statements in this section like this:

$$n^{-7/2} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

What these allows us to do is identify negligible terms in an expression, that is terms that with the right scaling disappear in the limit. We will establish nonzero limits for some of these expression in the next section.

First note that, for all $k \in \mathbb{N}$, $\mathbb{E}(\boldsymbol{M}_k \mid \mathcal{F}_{k-1}) = \boldsymbol{0}$ and thus $\mathbb{E}(\boldsymbol{M}_k) = \boldsymbol{0}$, since $\boldsymbol{M}_k = \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \mid \mathcal{F}_{k-1})$.

Lemma 3.1. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration satisfying* (M) *with* $\ell = 2$ *and* $\boldsymbol{X}_0 = \boldsymbol{0}$ *. Then*

$$\begin{aligned} \operatorname{Var}(\boldsymbol{M}_k \mid \mathcal{F}_{k-1}) &= X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}} \\ &= U_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + V_{k-1}\widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \boldsymbol{V}_{\boldsymbol{\varepsilon}} \end{aligned} \tag{21}$$

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

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

Proof. Using the decomposition (19), where, for all $k \in \mathbb{N}$, the random vectors $\big\{\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1}),\, \boldsymbol{\xi}_{k,j.2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2}),\, \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k) : j \in \mathbb{N}\big\}$ are independent of each other, also independent of $\mathcal{F}_{k-1}$, and have zero mean vector, we conclude (21). □

We will make good use of the following lemma on sums of i.i.d. random variables and its generalization.

Lemma 3.2. *Let* $(\zeta_k)_{k\in\mathbb{N}}$ *be independent and identically distributed random variables such that* $\mathbb{E}(\|\zeta_1\|^k) < \infty$ *with some* $k \in \mathbb{N}$.

(i) *Then there exists a polynomial* $Q : \mathbb{R} \to \mathbb{R}$, *with degree at most* k *such that*

$$\mathbb{E}\big((\zeta_1 + \cdots + \zeta_N)^k\big) = Q(N), \qquad N \in \mathbb{N}.$$

(ii) *If* $\mathbb{E}(\zeta_1) = 0$ *then there exists a polynomial* $R : \mathbb{R} \to \mathbb{R}$, *with degree at most* $\lfloor k/2 \rfloor$ *such that*

$$\mathbb{E}\big((\zeta_1 + \cdots + \zeta_N)^k\big) = R(N), \qquad N \in \mathbb{N}.$$

The coefficients of the polynomials Q *and* R *depend on the moments* $\mathbb{E}(\zeta_1^i)$, $i \in \{1, \ldots, N\}$.

Proof. (i) By the multinomial theorem we have

$$(\zeta_1 + \cdots + \zeta_N)^k = \left(\sum_{i=1}^N \zeta_i\right)^k = \sum_{\substack{i_1,\ldots,i_N \in \mathbb{Z}_+ \\ i_1+\cdots+i_N = k}} \frac{k!}{i_1! \ldots i_N!} \prod_{j=1}^N \zeta_j^{i_j}.$$

Taking expectation of both sides and using the independence of the variables yields

$$\mathbb{E}\left(\left(\sum_{i=1}^N \zeta_i\right)^k\right) = \sum_{\substack{i_1,\ldots,i_N \in \mathbb{Z}_+ \\ i_1+\cdots+i_N = k}} \frac{k!}{i_1! \ldots i_N!} \prod_{j=1}^N \mathbb{E}\left(\zeta_j^{i_j}\right).$$

Since the variables are also identically distributed, we can regroup this sum by introducing

$$k_s := \big|\{j \in \{1, \ldots, N\} : i_j = s\}\big|, \quad s \in \{1, \ldots, k\},$$

leading to

$$\mathbb{E}\left(\left(\sum_{i=1}^{N} \zeta_i\right)^k\right) = \sum_{\substack{s\in\{1,\ldots,k\},\\ (k_1,\ldots,k_s)\in H_s}} \binom{N}{k_1} \cdots \binom{N-k_1-\cdots-k_{s-1}}{k_s} \prod_{i=1}^{s} \left[\mathbb{E}\left(\zeta_1^i\right)\right]^{k_i},$$

where

$$H_s = \{(k_1, \ldots, k_s) \in \mathbb{Z}_+^s \colon k_s \neq 0,\ k_1 + 2k_2 + 3k_3 + \cdots + sk_s = k\}.$$

Since

$$\binom{N}{k_1}\binom{N-k_1}{k_2} \cdots \binom{N-k_1-\cdots-k_{s-1}}{k_s} = \frac{N(N-1)\cdots(N-k_1-k_2-\cdots-k_s+1)}{k_1!k_2!\cdots k_s!}$$

is a polynomial of the variable N having degree $k_1+\cdots+k_s \le k$, we have shown the existence of Q.

(ii) Using the same decomposition, we have

$$\mathbb{E}\left(\left(\sum_{i=1}^{N} \zeta_i\right)^k\right) = \sum_{\substack{s\in\{2,\ldots,k\},\\ (0,k_2,\ldots,k_s)\in H_s}} \binom{N}{k_2} \cdots \binom{N-k_2-\cdots-k_{s-1}}{k_s} \prod_{i=2}^{s} \left[\mathbb{E}\left(\zeta_1^i\right)\right]^{k_i}.$$

Here

$$\binom{N}{k_2}\binom{N-k_2}{k_3} \cdots \binom{N-k_2-\cdots-k_{s-1}}{k_s} = \frac{N(N-1)\cdots(N-k_2-k_3-\cdots-k_s+1)}{k_2!k_3!\cdots k_s!}$$

is a polynomial of the variable N having degree $k_2+\cdots+k_s$. Since

$$k = 2k_2 + 3k_3 + \cdots + sk_s \ge 2(k_2 + k_3 + \cdots + k_s),$$

we have $k_2 + \cdots + k_s \le k/2$ yielding part (ii). □

Lemma 3.2 can be generalized in the following way.

Lemma 3.3. *Let* $(\boldsymbol{\zeta}_i)_{i\in\mathbb{N}}$ *be independent and identically distributed random vectors with values in* $\mathbb{R}^2$ *such that* $\mathbb{E}\left(\|\boldsymbol{\zeta}_i\|^{k+\ell}\right) < \infty$ *with some* $k, \ell \in \mathbb{N}$.

(i) *There exists a polynomial* $Q : \mathbb{R} \to \mathbb{R}$*, with degree at most* $k+\ell$ *such that*

$$\mathbb{E}\left((\zeta_{1,1} + \cdots + \zeta_{N,1})^k(\zeta_{1,2} + \cdots + \zeta_{N,2})^\ell\right) = Q(N), \qquad N \in \mathbb{N}.$$

(ii) *If* $\mathbb{E}(\boldsymbol{\zeta}_1) = \mathbf{0}$ *then there exists a polynomial* $Q : \mathbb{R} \to \mathbb{R}$*, with degree at most* $\lfloor (k+\ell)/2 \rfloor$ *such that*

$$\mathbb{E}\left((\zeta_{1,1} + \cdots + \zeta_{N,1})^k(\zeta_{1,2} + \cdots + \zeta_{N,2})^\ell\right) = R(N), \qquad N \in \mathbb{N}.$$

The coefficients of the polynomials Q *and* R *depend on the moments*

$$\mathbb{E}\left(\zeta_{1,1}^i \zeta_{1,2}^j\right), \qquad i \in \{1, \ldots, k\}, \quad j \in \{1, \ldots, \ell\}.$$

We can use Lemma 3.3 to express the moments of $\boldsymbol{M}_k$ with the help of $X_{k-1,1}$ and $X_{k-1,2}$.

Corollary 3.4. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (M) *with some* $\ell \in \mathbb{N}$ *and* $\boldsymbol{X}_0 = \mathbf{0}$. *Then for all* $s, t \in \mathbb{Z}_+$, $s + t \leq \ell$, *there exists a polynomial* $R_{s,t} : \mathbb{R}^2 \to \mathbb{R}$ *having degree at most* $\lfloor (s+t)/2 \rfloor$ *such that*

$$\mathbb{E}\left(M_{k,1}^s M_{k,2}^t \mid \mathcal{F}_{k-1}\right) = R_{s,t}\left(X_{k-1,1}, X_{k-1,2}\right).$$

Proof. By (10) we have

$$\begin{aligned}
M_{k,1} = \boldsymbol{e}_1^\top \boldsymbol{M}_k &= X_{k,1} - \alpha X_{k-1,1} - \beta X_{k-1,2} - m_{\varepsilon,1} \\
&= \sum_{j=1}^{X_{k-1,1}} [\xi_{k,j,1,1} - \alpha] + \sum_{j=1}^{X_{k-1,2}} [\xi_{k,j,2,1} - \beta] + [\varepsilon_{k,1} - m_{\varepsilon,1}] \\
&= \sum_{j=1}^{X_{k-1,1}} [\xi_{k,j,1,1} - \mathbb{E}(\xi_{k,j,1,1})] + \sum_{j=1}^{X_{k-1,2}} [\xi_{k,j,2,1} - \mathbb{E}(\xi_{k,j,2,1})] \\
&\quad + [\varepsilon_{k,1} - \mathbb{E}(\varepsilon_{k,1})],
\end{aligned}$$

and

$$\begin{aligned}
M_{k,2} = &\sum_{j=1}^{X_{k-1,1}} [\xi_{k,j,1,2} - \mathbb{E}(\xi_{k,j,1,2})] + \sum_{j=1}^{X_{k-1,2}} [\xi_{k,j,2,2} - \mathbb{E}(\xi_{k,j,2,2})] \\
&+ [\varepsilon_{k,2} - \mathbb{E}(\varepsilon_{k,2})].
\end{aligned}$$

Introduce the notations

$$\boldsymbol{\zeta}_j := \boldsymbol{\xi}_{k,j,1} - \mathbb{E}\left(\boldsymbol{\xi}_{k,j,1}\right), \quad \boldsymbol{\eta}_j := \boldsymbol{\xi}_{k,j,2} - \mathbb{E}\left(\boldsymbol{\xi}_{k,j,2}\right), \quad \boldsymbol{\theta} := \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k),$$

then the random vectors $\left\{\boldsymbol{\zeta}_j, \boldsymbol{\eta}_j, \boldsymbol{\theta} : j \in \mathbb{N}\right\}$ are independent and have zero mean vector. Using the multinomial theorem twice we get

$$\begin{aligned}
&\mathbb{E}\left(M_{k,1}^s M_{k,2}^t \mid \mathcal{F}_{k-1}\right) \\
&= \mathbb{E}\left(\left(\sum_{j=1}^N \zeta_{j,1} + \sum_{j=1}^M \eta_{j,1} + \theta_1\right)^s \left(\sum_{j=1}^N \zeta_{j,2} + \sum_{j=1}^M \eta_{j,2} + \theta_2\right)^t\right)\Bigg|_{\substack{N=X_{k-1,1}\\ M=X_{k-1,2}}} \\
&= \sum_{\substack{i_1,i_2,j_1,j_2\in\mathbb{Z}_+\\ i_1+i_2\leq s\\ j_1+j_2\leq t}} \frac{s!t!}{i_1!i_2!(s-i_1-i_2)!j_1!j_2!(t-j_1-j_2)!}\mathbb{E}\left(\theta_1^{s-i_1-i_2}\theta_2^{t-j_1-j_2}\right) \\
&\qquad\qquad \times \mathbb{E}\left(\left(\sum_{j=1}^N \zeta_{j,1}\right)^{i_1}\left(\sum_{j=1}^N \zeta_{j,2}\right)^{j_1}\right)\Bigg|_{N=X_{k-1,1}} \\
&\qquad\qquad \times \mathbb{E}\left(\left(\sum_{j=1}^M \eta_{j,1}\right)^{i_2}\left(\sum_{j=1}^M \eta_{j,2}\right)^{j_2}\right)\Bigg|_{M=X_{k-1,2}}
\end{aligned}$$

By Lemma 3.3 there exist polynomials $R^{(1)}_{i_1,i_2}$ and $R^{(2)}_{j_1,j_2}$ having degrees at most $\lfloor (i_1+i_2)/2\rfloor$ and $\lfloor (j_1+j_2)/2\rfloor$ respectively such that

$$\begin{aligned}
&\mathbb{E}\left(M_{k,1}^s M_{k,2}^t \mid \mathcal{F}_{k-1}\right) \\
&= \sum_{\substack{i_1,i_2,j_1,j_2\in\mathbb{Z}_+\\ i_1+i_2\leq s\\ j_1+j_2\leq t}} \frac{s!t!\,\mathbb{E}\left(\theta_1^{s-i_1-i_2}\theta_2^{t-j_1-j_2}\right) R^{(1)}_{i_1,i_2}(X_{k-1,1}) R^{(2)}_{j_1,j_2}(X_{k-1,2})}{i_1!i_2!(s-i_1-i_2)!j_1!j_2!(t-j_1-j_2)!} \\
&=: R_{s,t}(X_{k-1,1}, X_{k-1,2}),
\end{aligned}$$

where $R_{s,t}$ is a polynomial with degree

$$\deg(R_{s,t}) \leq \max_{\substack{i_1,i_2,j_1,j_2\in\mathbb{Z}_+\\ i_1+i_2\leq s\\ j_1+j_2\leq t}} \left(\deg\left(R^{(1)}_{i_1,i_2}\right) + \deg\left(R^{(2)}_{j_1,j_2}\right)\right) \leq \frac{s+t}{2}.$$

□

Let $\otimes$ denote the Kronecker product of matrices, then we can state the following.

Lemma 3.5. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (M) *with some* $\ell \in \mathbb{N}$ *and* $\boldsymbol{X}_0 = \boldsymbol{0}$. *Then* $\mathbb{E}(\|\boldsymbol{X}_k\|^i) = \mathrm{O}(k^i)$ *and further*

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

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

Proof. For the first statement it is sufficient to show, that for any polynomial $P : \mathbb{R}^2 \to \mathbb{R}$ with degree ℓ there exists a constant $c_{P,\ell}$ such that for all $k \in \mathbb{N}$

$$\mathbb{E}\left(|P(X_{k,1}, X_{k,2})|\right) \le c_{P,\ell} k^\ell. \tag{22}$$

We'll show this by induction on ℓ.

If $\ell = 1$, then

$$\begin{aligned}
\mathbb{E}\left(X_{k,1}\right) &= \mathbb{E}\left(\boldsymbol{e}_1^\top \boldsymbol{X}_k\right) = \boldsymbol{e}_1^\top \mathbb{E}\left(\boldsymbol{X}_k\right) = \boldsymbol{e}_1^\top \sum_{j=1}^{k} \boldsymbol{m}_{\boldsymbol{\xi}}^{k-j} \boldsymbol{m}_{\boldsymbol{\varepsilon}} \\
&= \boldsymbol{e}_1^\top \left(\frac{k}{1-\lambda_-} \begin{bmatrix} 1-\delta & \beta \\ \gamma & 1-\alpha \end{bmatrix} + \frac{1-\lambda_-^k}{(1-\lambda_-)^2} \begin{bmatrix} 1-\alpha & -\beta \\ -\gamma & 1-\delta \end{bmatrix} \right) \boldsymbol{m}_{\boldsymbol{\varepsilon}},
\end{aligned}$$

similarly

$$\mathbb{E}\left(X_{k,2}\right) = \boldsymbol{e}_2^\top \left(\frac{k}{1-\lambda_-} \begin{bmatrix} 1-\delta & \beta \\ \gamma & 1-\alpha \end{bmatrix} + \frac{1-\lambda_-^k}{(1-\lambda_-)^2} \begin{bmatrix} 1-\alpha & -\beta \\ -\gamma & 1-\delta \end{bmatrix} \right) \boldsymbol{m}_{\boldsymbol{\varepsilon}}.$$

This proves the statement for $\ell = 1$.

Now fix some $\ell \in \mathbb{N}$ and suppose, that (22) holds for any polynomial P with degree at most $\ell - 1$. Since every polynomial is a linear combination of monomials all we have to prove is that for all $s, t \in \mathbb{Z}_+$, $s + t = \ell$ there exists $c_{s,t} \in \mathbb{R}_+$ such that for all $k \in \mathbb{N}$

$$\mathbb{E}\big(X_{k,1}^s X_{k,2}^t\big) \le c_{s,t}\, k^{s+t}.$$

By (11) we have

$$\begin{aligned}
X_{k,1}^s X_{k,2}^t &= \left(\boldsymbol{e}_1^\top \left(\boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} + \boldsymbol{M}_{k-1}\right)\right)^s \\
&\qquad \times \left(\boldsymbol{e}_2^\top \left(\boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} + \boldsymbol{M}_{k-1}\right)\right)^t \\
&= \left(\alpha X_{k-1,1} + \beta X_{k-1,2} + m_{\varepsilon,1} + M_{k-1,1}\right)^s \\
&\qquad \times \left(\delta X_{k-1,1} + \gamma X_{k-1,2} + m_{\varepsilon,2} + M_{k-1,2}\right)^t.
\end{aligned}$$

By Corollary 3.4 one can show, that there exists a polynomial $Q_{s,t}$ having degree at most $s+t-1=\ell-1$ such that

$$\begin{aligned}\mathbb{E}\left(X_{k,1}^{s}X_{k,2}^{t} \mid \mathcal{F}_{k-1}\right) = &\left(\alpha X_{k-1,1}+\beta X_{k-1,2}\right)^{s}\left(\delta X_{k-1,1}+\gamma X_{k-1,2}\right)^{t}\\ &+ Q_{s,t}\left(X_{k-1,1},X_{k-1,2}\right).\end{aligned}$$

Using the binomial theorem we get

$$\begin{aligned}&\left(\alpha X_{k-1,1}+\beta X_{k-1,2}\right)^{s}\left(\delta X_{k-1,1}+\gamma X_{k-1,2}\right)^{t}\\ &=\left(\sum_{i=0}^{s}\binom{s}{i}\beta^{i}\alpha^{s-i}X_{k-1,1}^{s-i}X_{k-1,2}^{i}\right)\left(\sum_{j=0}^{t}\binom{t}{j}\delta^{j}\gamma^{t-j}X_{k-1,1}^{t-j}X_{k-1,2}^{j}\right)\\ &=\sum_{i=0}^{s}\sum_{j=0}^{t}X_{k-1,1}^{s+t-i-j}X_{k-1,2}^{i+j}\binom{s}{i}\binom{t}{j}\alpha^{s-i}\beta^{i}\gamma^{t-j}\delta^{j}\\ &=\sum_{m=0}^{s+t}X_{k-1,1}^{s+t-m}X_{k-1,2}^{m}h_{t,m}\end{aligned}$$

where

$$h_{t,m}=\sum_{i=0}^{m}\binom{\ell-t}{i}\binom{t}{m-i}\alpha^{\ell-t-i}\beta^{i}\gamma^{t+i-m}\delta^{m-i}$$

Consequently

$$\mathbb{E}\left(\begin{bmatrix}X_{k,1}^{\ell}\\ X_{k,1}^{\ell-1}X_{k,2}\\ \vdots\\ X_{k,1}X_{k,2}^{\ell-1}\\ X_{k,2}^{\ell}\end{bmatrix} \mid \mathcal{F}_{k-1}\right)=\boldsymbol{H}_{\ell}\begin{bmatrix}X_{k-1,1}^{\ell}\\ X_{k-1,1}^{\ell-1}X_{k-1,2}\\ \vdots\\ X_{k-1,1}X_{k-1,2}^{\ell-1}\\ X_{k-1,2}^{\ell}\end{bmatrix}+\begin{bmatrix}Q_{\ell,0}(X_{k-1,1},X_{k-1,2})\\ Q_{\ell-1,1}(X_{k-1,1},X_{k-1,2})\\ \vdots\\ Q_{1,\ell-1}(X_{k-1,1},X_{k-1,2})\\ Q_{0,\ell}(X_{k-1,1},X_{k-1,2})\end{bmatrix}$$

with $\boldsymbol{H}_{\ell}=\left(h_{t,m}\right)_{t,m=0}^{\ell}\in\mathbb{R}^{(\ell+1)\times(\ell+1)}$. Iterating this recursion, taking expectation of both sides and applying the tower-rule yields

$$\mathbb{E}\left(\begin{bmatrix}X_{k,1}^{\ell}\\ X_{k,1}^{\ell-1}X_{k,2}\\ \vdots\\ X_{k,1}X_{k,2}^{\ell-1}\\ X_{k,2}^{\ell}\end{bmatrix}\right)=\sum_{i=0}^{k-1}\boldsymbol{H}_{\ell}^{i}\begin{bmatrix}\mathbb{E}\left(Q_{\ell,0}(X_{k-1-i,1},X_{k-1-i,2})\right)\\ \mathbb{E}\left(Q_{\ell-1,1}(X_{k-1-i,1},X_{k-1-i,2})\right)\\ \vdots\\ \mathbb{E}\left(Q_{1,\ell-1}(X_{k-1-i,1},X_{k-1-i,2})\right)\\ \mathbb{E}\left(Q_{0,\ell}(X_{k-1-i,1},X_{k-1-i,2})\right)\end{bmatrix}.$$

We now show that the matrix $\boldsymbol{H}_{\ell}$ has spectral radius 1, thus $\|H_{\ell}^{i}\|=\mathrm{O}(1)$. Let us denote the coordinates of vector $\boldsymbol{u}_{\text{right}}$ with u_1 and u_2 respectively, then introduce the vector

$$\boldsymbol{u}_{\mathrm{right}}^{(\ell)} = \begin{bmatrix} u_1^\ell & u_1^{\ell-1}u_2 & \dots & u_1 u_2^{\ell-1} & u_2^\ell \end{bmatrix}^\top .$$

Since $\boldsymbol{u}_{\mathrm{right}}$ is a right eigenvector of the matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue 1 we get

$$\boldsymbol{H}_\ell\, \boldsymbol{u}_{\mathrm{right}}^{(\ell)} = \begin{bmatrix} \sum\limits_{m=0}^{\ell} u_1^{\ell-m} u_2^m h_{0,m} \\ \vdots \\ \sum\limits_{m=0}^{\ell} u_1^{\ell-m} u_2^m h_{\ell,m} \end{bmatrix} = \begin{bmatrix} (\alpha u_1 + \beta u_2)^\ell \\ (\alpha u_1 + \beta u_2)^{\ell-1}(\gamma u_1 + \delta u_2) \\ \vdots \\ (\gamma u_1 + \delta u_2)^\ell \end{bmatrix} = \boldsymbol{u}_{\mathrm{right}}^{(\ell)}.$$

Therefore $\boldsymbol{u}_{\mathrm{right}}^{(\ell)}$ is an eigenvector of the matrix $\boldsymbol{H}_\ell$ belonging to the eigenvalue 1 with all positive components, thus by the Perron–Frobenius theorem we have that the spectral radius of $\boldsymbol{H}_\ell$ is 1. Consequently there exists $h_\ell \in \mathbb{R}_+$ such that for all $i \in \mathbb{N}$ we have $\|H_\ell^i\| \leq h_\ell$. Then by the induction hypothesis there exist constants $c_{Q_{s,t},\ell}$ such that for all $k \in \mathbb{N}$

$$\mathbb{E}\left(|Q_{s,t}(X_{k,1}, X_{k,2})|\right) \leq c_{Q_{s,t},\ell} k^\ell$$

Putting it all together we have

$$\mathbb{E}\left(X_{k,1}^s X_{k,2}^t\right) \leq h_\ell \max\{c_{Q_{s,t},\ell} : s+t=\ell\} \sum_{i=0}^{k-1} (k-1-i)^{\ell-1} = \mathrm{O}(k^\ell).$$

This concludes the statement for $\boldsymbol{X}_k$, and by Corollary 3.4 for $\boldsymbol{M}_k$. By (12) we have

$$\mathbb{E}(U_k^i) = \mathbb{E}\left[\left(\frac{(\gamma+1-\delta)X_{k,1} + (\beta+1-\alpha)X_{k,2}}{1-\lambda_-}\right)^i\right] = \mathrm{O}\left(k^i\right).$$

Next, for $j \in \mathbb{Z}_+$ with $2j \leq \ell$, we prove $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$ using induction in k. By the recursion $V_k = (\alpha+\delta-1)V_{k-1} + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_k + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $k \in \mathbb{N}$, we have $\mathbb{E}(V_k) = (\alpha+\delta-1)\mathbb{E}(V_{k-1}) + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $k \in \mathbb{N}$, with initial value $\mathbb{E}(V_0) = 0$, hence

$$\mathbb{E}(V_k) = \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \sum_{i=0}^{k-1} (\alpha+\delta-1)^i, \qquad k \in \mathbb{N},$$

which yields $|\mathbb{E}(V_k)| = \mathrm{O}(1)$. Indeed, for all $k \in \mathbb{N}$,

$$\left|\sum_{i=0}^{k-1} (\alpha + \delta - 1)^i\right| \leq \frac{1}{1 - |\alpha + \delta - 1|}.$$

The rest of the proof of $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$ can be carried out as in Corollary 9.1 of Barczy et al. [3]. □

The next corollary can be derived exactly as Corollary 9.2 of Barczy et al. [3].

Corollary 3.6. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (M) *with some* $\ell \in \mathbb{N}$ *and* $\boldsymbol{X}_0 = \boldsymbol{0}$. *Then*

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

$$n^{-\kappa} \sum_{k=1}^{n} \left|U_k^i V_k^j\right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty, \tag{23}$$

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

$$n^{-\kappa} \sup_{t\in[0,T]} \left|U_{\lfloor nt\rfloor}^i V_{\lfloor nt\rfloor}^j\right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty, \tag{24}$$

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

$$n^{-\kappa} \sup_{t\in[0,T]} \left|\sum_{k=1}^{\lfloor nt\rfloor} [U_k^i V_k^j - \mathbb{E}(U_k^i V_k^j \mid \mathcal{F}_{k-1})]\right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty. \tag{25}$$

Unfortunately the above corollary doesn't always give good enough bounds. In a few select cases we provide sharper bounds on the growth of these variables.

Remark 3.7. *In the special case* $(\ell, i, j) = (2, 1, 0)$, *one can improve* (24), *namely, one can show*

$$n^{-\kappa} \sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty \text{ for } \kappa > 1, \tag{26}$$

see Barczy et al. [3].

Lemma 3.8. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with* $\ell = 4$. *Then for each* $T > 0$,

$$n^{-3/2} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1} \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as} \;\; n \to \infty.$$

Proof. In order to prove the statement we derive a decomposition of $\sum_{k=1}^{\lfloor nt\rfloor} V_k$ as a martingale and some other terms. Using the recursion (17), we obtain

$$\begin{aligned}\mathbb{E}(V_k \mid \mathcal{F}_{k-1}) &= \mathbb{E}\left(\lambda_- V_{k-1} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{M}_k + \boldsymbol{m}_\varepsilon\rangle \mid \mathcal{F}_{k-1}\right)\\ &= \lambda_- V_{k-1} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{m}_\varepsilon\rangle\end{aligned}$$

Thus

$$\sum_{k=1}^{\lfloor nt\rfloor} V_k = \sum_{k=1}^{\lfloor nt\rfloor} [V_k - \mathbb{E}(V_k \mid \mathcal{F}_{k-1})] + \lambda_- \sum_{k=2}^{\lfloor nt\rfloor} V_{k-1} + \mathrm{O}(n)$$

Consequently

$$\sum_{k=1}^{\lfloor nt\rfloor} V_k = \frac{1}{1-\lambda_-} \sum_{k=1}^{\lfloor nt\rfloor} [V_k - \mathbb{E}(V_k \mid \mathcal{F}_{k-1})] - \frac{\lambda_-}{1-\lambda_-} V_{\lfloor nt\rfloor-1} + \mathrm{O}(n)$$

Using (25) with $(\ell, i, j) = (4, 0, 1)$ and (24) with $(\ell, i, j) = (2, 0, 1)$ we conclude the statement. □

Lemma 3.9. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with* $\ell = 4$. *Then for each* $T > 0$,

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

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 (17), (13) and Lemma 3.1, we obtain

$$\begin{aligned}&\mathbb{E}(U_{k-1}V_{k-1} \mid \mathcal{F}_{k-2}) =\\ &\quad = \mathbb{E}\Big(\big(U_{k-2} + \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_\varepsilon\rangle\big)\big(\lambda_- V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_\varepsilon\rangle\big) \,\Big|\, \mathcal{F}_{k-2}\Big)\\ &\quad = \lambda_- U_{k-2}V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{m}_\varepsilon\rangle U_{k-2} + \lambda_- \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{m}_\varepsilon\rangle V_{k-2}\\ &\qquad\quad + \boldsymbol{u}_{\text{left}}^\top \boldsymbol{m}_\varepsilon \boldsymbol{m}_\varepsilon^\top \boldsymbol{v}_{\text{left}} + \boldsymbol{u}_{\text{left}}^\top \mathbb{E}(\boldsymbol{M}_{k-1}\boldsymbol{M}_{k-1}^\top \mid \mathcal{F}_{k-2}) \boldsymbol{v}_{\text{left}}\\ &\quad = \lambda_- U_{k-2}V_{k-2} + \text{constant} + \text{linear combination of } \; U_{k-2} \;\text{ and }\; V_{k-2}.\end{aligned}$$

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

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

$$- \frac{\lambda_-}{1-\lambda_-} U_{\lfloor nt\rfloor-1}V_{\lfloor nt\rfloor-1} + \mathrm{O}(n)$$

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

Using (25) with $(\ell, i, j) = (4, 1, 1)$ we have

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

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

$$n^{-5/2} \sum_{k=1}^{\lfloor nT\rfloor} U_k \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad n^{-5/2} \sum_{k=1}^{\lfloor nT\rfloor} |V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{27}$$

$$n^{-5/2} \sup_{t\in[0,T]} |U_{\lfloor nt\rfloor}V_{\lfloor nt\rfloor}| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{28}$$

as $n \to \infty$. Using (23) with $(\ell, i, j) = (2, 1, 0)$ and $(\ell, i, j) = (2, 0, 1)$ we have (27), and by (24) with $(\ell, i, j) = (3, 1, 1)$ we have (28), thus we conclude the statement. □

Lemma 3.10. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with* $\ell = 8$. *Then for each* $T > 0$,

$$n^{-7/2} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 V_{k-1} \right| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as} \quad n \to \infty.$$

Proof. The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 V_{k-1}$ as a sum of a martingale and some other terms. Using recursions (13), (17) and Lemma 3.1, we obtain

$$\begin{aligned}
&\mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2}) \\
&= \mathbb{E}\left(\big(U_{k-2} + \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\varepsilon}\rangle\big)^2 \big(\lambda_- V_{k-2} + \langle \boldsymbol{v}_{\text{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\varepsilon}\rangle\big) \,\Big|\, \mathcal{F}_{k-2} \right) \\
&= \lambda_- U_{k-2}^2 V_{k-2} + \text{constant} \\
&\qquad + \text{linear combination of } U_{k-2}, V_{k-2}, U_{k-2}^2, V_{k-2}^2 \text{ and } U_{k-2}V_{k-2}.
\end{aligned}$$

Thus

$$\begin{aligned}
&\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 V_{k-1} = \\
&= \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\big] + \sum_{k=2}^{\lfloor nt\rfloor} \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2}) \\
&= \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\big] + \lambda_- \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}^2 V_{k-2} + \mathrm{O}(n) \\
&\quad + \text{linear combination of } \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}, \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}, \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}^2, \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}^2 \\
&\qquad \text{and} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2} V_{k-2}.
\end{aligned}$$

Consequently

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

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

$$\text{and } \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}, \ \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}^2, \ \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}^2, \ \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2} V_{k-2}.$$

Using (25) with $(\ell, i, j) = (8, 2, 1)$ we have

$$n^{-7/2} \sup_{t\in[0,T]} \left| \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\big] \right| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to prove the lemma, it suffices to show that

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} U_k \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} U_k^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} |V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0,$$

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} V_k^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} |U_k V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0,$$

$$n^{-7/2} \sup_{t\in[0,T]} |U_{\lfloor nt\rfloor}^2 V_{\lfloor nt\rfloor}| \stackrel{\mathbb{P}}{\longrightarrow} 0$$

as $n \to \infty$. These follow from a straightforward application of (23) and (24). □

3.3 Limit theorems for building blocks

Up to this point we have defined a decomposition of the process and proven some zero limit theorems about a few expression related to it. We will use these results to find nonzero limits.

First we relate the sums of squares of the variables V_k to the well-behaved part of our decomposition, the variables U_k. If the process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ satisfies the condition (ND), then this can be used to find the nonzero limit of the aforementioned sum.

Lemma 3.11. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with* $\ell = 8$. *Then for each* $T > 0$, *we have*

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

Proof. In order to prove the statement, we derive a decomposition of $\sum_{k=1}^{\lfloor nt\rfloor} V_k^2$ as a sum of a martingale and some other terms. Using recursion (17) and Lemma 3.1, we obtain

$$\begin{aligned}
\mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) &= \mathbb{E}\left[(\lambda_- V_{k-1} + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_k + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle)^2 \,\Big|\, \mathcal{F}_{k-1}\right] \\
&= \lambda_-^2 V_{k-1}^2 + 2\lambda_- \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle V_{k-1} + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \\
&\quad + \boldsymbol{v}_{\mathrm{left}}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \boldsymbol{v}_{\mathrm{left}} \\
&= \lambda_-^2 V_{k-1}^2 + \boldsymbol{v}_{\mathrm{left}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}} U_{k-1} + \text{constant} + \text{constant} \times V_{k-1}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{k=1}^{\lfloor nt\rfloor} V_k^2 &= \sum_{k=1}^{\lfloor nt\rfloor} \left[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\right] + \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) \\
&= \sum_{k=1}^{\lfloor nt\rfloor} \left[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\right] + \lambda_-^2 \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 + \boldsymbol{v}_{\mathrm{left}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} \\
&\quad + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\sum_{k=1}^{\lfloor nt\rfloor} V_k^2 &= \frac{1}{1-\lambda_-^2} \sum_{k=1}^{\lfloor nt\rfloor} \left[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\right] + \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\, \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} \\
&\quad - \frac{\lambda_-^2}{1-\lambda_-^2} V_{\lfloor nt\rfloor}^2 + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}.
\end{aligned}$$

Using (25) with $(\ell, i, j) = (8, 0, 2)$ we obtain

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

Using (24) with $(\ell, i, j) = (3, 0, 2)$ we obtain

$$n^{-2} \sup_{t\in[0,T]} V_{\lfloor nt\rfloor}^2 \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as} \ \ n \to \infty.$$

Moreover,

$$n^{-2} \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1} \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as} \ \ n \to \infty.$$

follows by (23) with the choice $(\ell, i, j) = (4, 0, 1)$. Consequently, we obtain the statement. □

We recall a result about convergence of random step processes towards a diffusion process, see Ispány and Pap [11].

Theorem. 3.12. *Let* $\boldsymbol{\gamma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ *be a continuous function. Assume that uniqueness in the sense of probability law holds for the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{U}}_t = \gamma(t, \boldsymbol{\mathcal{U}}_t)\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \tag{29}$$

with initial value $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{u}_0$ *for all* $\boldsymbol{u}_0 \in \mathbb{R}^d$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *is an* r*-dimensional standard Wiener process. Let* $(\boldsymbol{\mathcal{U}}_t)_{t\in\mathbb{R}_+}$ *be a solution of* (29) *with initial value* $\boldsymbol{\mathcal{U}}_0 = \mathbf{0} \in \mathbb{R}^d$.

For each $n \in \mathbb{N}$, *let* $(\boldsymbol{U}_k^{(n)})_{k\in\mathbb{N}}$ *be a sequence of* d*-dimensional martingale differences with respect to a filtration* $(\mathcal{F}_k^{(n)})_{k\in\mathbb{Z}_+}$, *i.e.,* $\mathbb{E}(\boldsymbol{U}_k^{(n)} \mid \mathcal{F}_{k-1}^{(n)}) = 0$, $n \in \mathbb{N}$, $k \in \mathbb{N}$. *Let*

$$\boldsymbol{\mathcal{U}}_t^{(n)} := \sum_{k=1}^{\lfloor nt\rfloor} \boldsymbol{U}_k^{(n)}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Suppose that $\mathbb{E}\big(\|\boldsymbol{U}_k^{(n)}\|^2\big) < \infty$ *for all* $n, k \in \mathbb{N}$. *Suppose that for each* $T > 0$,

(i) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} \operatorname{Var}\big(\boldsymbol{U}_k^{(n)} \mid \mathcal{F}_{k-1}^{(n)}\big) - \int_0^t \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})\boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})^\top \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0,$

(ii) $\displaystyle\sum_{k=1}^{\lfloor nT\rfloor} \mathbb{E}\big(\|\boldsymbol{U}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{U}_k^{(n)}\|>\theta\}} \,\big|\, \mathcal{F}_{k-1}^{(n)}\big) \stackrel{\mathbb{P}}{\longrightarrow} 0$ *for all* $\theta > 0$,

where $\stackrel{\mathbb{P}}{\longrightarrow}$ *denotes convergence in probability. Then* $\boldsymbol{\mathcal{U}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{U}}$ *as* $n \to \infty$.

Note that in (i) of Theorem 3.12, $\|\cdot\|$ denotes a matrix norm, while in (ii) it denotes a vector norm.

We will use the above theorem to prove limit theorems on our building blocks. However this theorem only applies to martingale differences, so we must restrict ourself to using those. Consider the sequence of stochastic processes

$$\boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1}\boldsymbol{M}_k \\ n^{-2}\boldsymbol{M}_k U_{k-1} \\ n^{-3/2}\boldsymbol{M}_k V_{k-1} \end{bmatrix} = \begin{bmatrix} n^{-1} \\ n^{-2}U_{k-1} \\ n^{-3/2}V_{k-1} \end{bmatrix} \otimes \boldsymbol{M}_k$$

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \boldsymbol{\mathcal{M}}_t^{(n)} \\ \boldsymbol{\mathcal{N}}_t^{(n)} \\ \boldsymbol{\mathcal{P}}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt\rfloor} \boldsymbol{Z}_k^{(n)}$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$.

Theorem. 3.13. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with* $\ell = 8$. *Then we have*

$$\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}} \qquad \text{as } n \to \infty, \tag{30}$$

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*

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t) \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{31}$$

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* $\gamma : \mathbb{R}_+ \times (\mathbb{R}^2)^3 \to (\mathbb{R}^{2\times 2})^{3\times 2}$ *is defined by*

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

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

Proof. In order to show convergence $\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}}$, we apply Theorem 3.12 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 3.13.

We start by showing that the SDE (31) 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 SDE (31) has the form

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \mathrm{d}\boldsymbol{\mathcal{N}}_t \\ \mathrm{d}\boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} (\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^+)^{1/2}\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ (\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^+)^{3/2}\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad (32)$$

for all $t \in \mathbb{R}_+$. One can prove that the first 2-dimensional equation of the SDE (32) has a pathwise unique strong solution $(\boldsymbol{\mathcal{M}}_t^{(\boldsymbol{y}_0)})_{t\in\mathbb{R}_+}$ with arbitrary initial value $\boldsymbol{\mathcal{M}}_0^{(\boldsymbol{y}_0)} = \boldsymbol{y}_0 \in \mathbb{R}^2$. Indeed, it is equivalent to the existence of a pathwise unique strong solution of the SDE

$$\begin{cases} \mathrm{d}\mathcal{S}_t = \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\,\mathrm{d}t + (\mathcal{S}_t^+)^{1/2}\, \boldsymbol{u}_{\mathrm{left}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\boldsymbol{\mathcal{W}}_t, \\ \mathrm{d}\boldsymbol{\mathcal{Q}}_t = -\boldsymbol{\Pi}\boldsymbol{m}_{\boldsymbol{\varepsilon}}\,\mathrm{d}t + (\mathcal{S}_t^+)^{1/2}\, \big(\boldsymbol{I}_2 - \boldsymbol{\Pi}\big)\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\boldsymbol{\mathcal{W}}_t, \end{cases} \qquad t \in \mathbb{R}_+, \qquad (33)$$

with initial value $\big(\mathcal{S}_0^{(\boldsymbol{y}_0)}, \boldsymbol{\mathcal{Q}}_0^{(\boldsymbol{y}_0)}\big) = \big(\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{y}_0\rangle, (\boldsymbol{I}_2 - \boldsymbol{\Pi})\boldsymbol{y}_0\big) \in \mathbb{R}\times\mathbb{R}^2$, where $\boldsymbol{I}_2$ denotes the 2-dimensional unit matrix and $\boldsymbol{\Pi} := \boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{left}}^\top$, since we have the correspondences

$$\mathcal{S}_t^{(\boldsymbol{y}_0)} = \boldsymbol{u}_{\mathrm{left}}^\top(\boldsymbol{\mathcal{M}}_t^{(\boldsymbol{y}_0)} + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}), \qquad \boldsymbol{\mathcal{Q}}_t^{(\boldsymbol{y}_0)} = \boldsymbol{\mathcal{M}}_t^{(\boldsymbol{y}_0)} - \mathcal{S}_t^{(\boldsymbol{y}_0)}\boldsymbol{u}_{\mathrm{right}}$$
$$\boldsymbol{\mathcal{M}}_t^{(\boldsymbol{y}_0)} = \boldsymbol{\mathcal{Q}}_t^{(\boldsymbol{y}_0)} + \mathcal{S}_t^{(\boldsymbol{y}_0)}\boldsymbol{u}_{\mathrm{right}},$$

see the proof of Ispány and Pap [12, Theorem 3.1]. By Remark 2.5, $\mathcal{S}_t^+$ may be replaced by $\mathcal{S}_t$ for all $t \in \mathbb{R}_+$ in the first equation of (33) provided that $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{y}_0\rangle \in \mathbb{R}_+$, hence $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^+$ may be replaced by $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle$ for all $t \in \mathbb{R}_+$ in (32). Thus the SDE (31) has a pathwise unique strong solution with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$.

Now we show that conditions (i) and (ii) of Theorem 3.12 hold. The conditional variance has the form

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

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

$$\begin{bmatrix} \mathcal{R}_{n,s} & \mathcal{R}_{n,s}^2 & \boldsymbol{0} \\ \mathcal{R}_{n,s}^2 & \mathcal{R}_{n,s}^3 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\mathcal{R}_{n,s}^2 \end{bmatrix} \otimes \overline{\boldsymbol{V}}_{\xi}$$

for $s \in \mathbb{R}_+$, where $\mathcal{R}_{n,s} := \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, and we used that $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^+ = \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$. Indeed, by (10), we get

$$\begin{aligned} \mathcal{R}_{n,s} &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \boldsymbol{u}_{\mathrm{left}}, s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{X}_k - \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + s \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{X}_{\lfloor ns \rfloor} \rangle + \frac{ns - \lfloor ns \rfloor}{n} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n} \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \in \mathbb{R}_+ \end{aligned} \tag{34}$$

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

In order to check condition (i) of Theorem 3.12, we need to prove that for each $T > 0$,

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{V}_{\boldsymbol{M}_k} - \int_0^t \mathcal{R}_{n,s} \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{35}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} - \int_0^t \mathcal{R}_{n,s}^2 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{36}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} - \int_0^t \mathcal{R}_{n,s}^3 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{37}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1 - \lambda_-^2} \int_0^t \mathcal{R}_{n,s}^2 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{38}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{39}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{40}$$

as $n \to \infty$.

First we show (35). By (34),

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

Using Lemma 3.1, we have $\boldsymbol{V}_{\boldsymbol{M}_k} = U_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + V_{k-1}\widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \boldsymbol{V}_{\boldsymbol{\varepsilon}}$, thus, in order to show (35), it suffices to prove

$$n^{-2}\sum_{k=1}^{\lfloor nT\rfloor}|V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-2}\sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{41}$$

$$n^{-2}\sup_{t\in[0,T]}\left[\lfloor nt\rfloor + (nt-\lfloor nt\rfloor)^2\right] \to 0 \tag{42}$$

as $n\to\infty$. Using (23) with $(\ell,i,j)=(2,0,1)$ and (24) with $(\ell,i,j)=(2,1,0)$, we have (41). Clearly, (42) follows from $|nt-\lfloor nt\rfloor|\leq 1$, $n\in\mathbb{N}$, $t\in\mathbb{R}_+$, thus we conclude (35).

Next we turn to prove (36). By (34),

$$\begin{aligned}\int_0^t \mathcal{R}_{n,s}^2\,\mathrm{d}s = {}& \frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor-1} U_k^2 + \frac{1}{n^3}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\sum_{k=1}^{\lfloor nt\rfloor-1} U_k + \frac{nt-\lfloor nt\rfloor}{n^3}U_{\lfloor nt\rfloor}^2\\ &+ \frac{(nt-\lfloor nt\rfloor)^2}{n^3}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle U_{\lfloor nt\rfloor}\\ &+ \frac{\lfloor nt\rfloor + (nt-\lfloor nt\rfloor)^3}{3n^3}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2.\end{aligned}$$

Using Lemma 3.1, we obtain

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}\boldsymbol{V}_{\boldsymbol{\varepsilon}}. \tag{43}$$

Thus, in order to show (36), it suffices to prove

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}|U_kV_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}U_k \stackrel{\mathbb{P}}{\longrightarrow} 0, \quad n^{-3/2}\sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{44}$$

$$n^{-3}\sup_{t\in[0,T]}\left[\lfloor nt\rfloor + (nt-\lfloor nt\rfloor)^3\right] \to 0 \tag{45}$$

as $n\to\infty$. By (23) with $(\ell,i,j)=(2,1,1)$ and $(\ell,i,j)=(2,1,0)$, and by (26), we have (44). Clearly, (45) follows from $|nt-\lfloor nt\rfloor|\leq 1$, $n\in\mathbb{N}$, $t\in\mathbb{R}_+$, thus we conclude (36).

Now we turn to check (37). Again by (34), we have

$$\begin{aligned}\int_0^t \mathcal{R}_{n,s}^3 \,\mathrm{d}s &= \frac{1}{n^4}\sum_{k=1}^{\lfloor nt\rfloor-1} U_k^3 + \frac{3}{2n^4}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \sum_{k=1}^{\lfloor nt\rfloor-1} U_k^2 + \frac{1}{n^4}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \sum_{k=1}^{\lfloor nt\rfloor-1} U_k \\ &+ \frac{nt-\lfloor nt\rfloor}{n^4} U_{\lfloor nt\rfloor}^3 + \frac{3(nt-\lfloor nt\rfloor)^2}{2n^4}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle U_{\lfloor nt\rfloor}^2 \\ &+ \frac{(nt-\lfloor nt\rfloor)^3}{n^4}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 U_{\lfloor nt\rfloor} + \frac{\lfloor nt\rfloor + (nt-\lfloor nt\rfloor)^4}{4n^4}\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^3.\end{aligned}$$

Using Lemma 3.1, we obtain

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^3 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 V_{k-1} \widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \boldsymbol{V}_{\boldsymbol{\varepsilon}}. \tag{46}$$

Thus, in order to show (37), it suffices to prove

$$n^{-4}\sum_{k=1}^{\lfloor nT\rfloor} |U_k^2 V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-4}\sum_{k=1}^{\lfloor nT\rfloor} U_k^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{47}$$

$$n^{-4}\sum_{k=1}^{\lfloor nT\rfloor} U_k \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-4/3}\sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{48}$$

$$n^{-4}\sup_{t\in[0,T]}\left[\lfloor nt\rfloor + (nt-\lfloor nt\rfloor)^4\right] \to 0 \tag{49}$$

as $n\to\infty$. By (23) with $(\ell,i,j)=(4,2,1)$, $(\ell,i,j)=(4,2,0)$, and $(\ell,i,j)=(2,1,0)$, and by (26), we have (47) and (48). Clearly, (49) follows again from $|nt-\lfloor nt\rfloor|\leq 1$, $n\in\mathbb{N}$, $t\in\mathbb{R}_+$, thus we conclude (37).

Next we turn to prove (38). First we show that

$$n^{-3}\sup_{t\in[0,T]}\left\|\sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\right\| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{50}$$

as $n\to\infty$ for all $T>0$. Using Lemma 3.1, we obtain

$$\sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} V_{k-1}^2 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^3 \widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{\varepsilon}}. \tag{51}$$

Using (23) with $(\ell,i,j)=(6,0,3)$ and $(\ell,i,j)=(4,0,2)$, we have

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

hence (50) will follow from

$$n^{-3} \sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \right\| \xrightarrow{\mathbb{P}} 0 \qquad \text{as} \ \ n\to\infty \tag{52}$$

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 (13), (17) and formulas (21), we obtain

$$\begin{aligned}
&\mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2}) \\
&\quad = \mathbb{E}\Big((U_{k-2} + \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle) \big(\lambda_- V_{k-2} + \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\big)^2 \,\Big|\, \mathcal{F}_{k-2}\Big) \\
&\quad = \lambda_-^2 U_{k-2}V_{k-2}^2 + \boldsymbol{v}_{\mathrm{left}}^\top \mathbb{E}(\boldsymbol{M}_{k-1}\boldsymbol{M}_{k-1}^\top \mid \mathcal{F}_{k-2}) \boldsymbol{v}_{\mathrm{left}}\, U_{k-2} \\
&\qquad + \text{constant} + \text{linear combination of} \ \ U_{k-2}V_{k-2}, \ \ V_{k-2}^2, \ \ U_{k-2} \ \ \text{and} \ \ V_{k-2} \\
&\quad = \lambda_-^2 U_{k-2}V_{k-2}^2 + \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle\, U_{k-2}^2 + \text{constant} \\
&\qquad + \text{linear combination of} \ \ U_{k-2}V_{k-2}, \ \ V_{k-2}^2, \ \ U_{k-2} \ \ \text{and} \ \ V_{k-2}.
\end{aligned}$$

Thus

$$\begin{aligned}
&\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 \\
&\quad = \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2})\big] + \sum_{k=2}^{\lfloor nt\rfloor} \mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2}) \\
&\quad = \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2})\big] \\
&\qquad + \lambda_-^2 \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}^2 + \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}^2 + \mathrm{O}(n) \\
&\qquad + \ \text{linear combination of} \ \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}, \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}^2, \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}, \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}.
\end{aligned}$$

Consequently,

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

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

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

Using (25) with $(\ell, i, j) = (8, 1, 2)$ we have

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

Thus, in order to show (52), it suffices to prove

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}|U_kV_k| \xrightarrow{\mathbb{P}} 0, \qquad n^{-3}\sum_{k=1}^{\lfloor nT\rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0, \tag{53}$$

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor} U_k \xrightarrow{\mathbb{P}} 0, \qquad n^{-3}\sum_{k=1}^{\lfloor nT\rfloor} |V_k| \xrightarrow{\mathbb{P}} 0, \tag{54}$$

$$n^{-3}\sup_{t\in[0,T]} U_{\lfloor nt\rfloor}V_{\lfloor nt\rfloor}^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3/2}\sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \xrightarrow{\mathbb{P}} 0 \tag{55}$$

as $n\to\infty$. Using (23) 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 (53) and (54). By (24) with $(\ell, i, j) = (4, 1, 2)$, and by (26), we have (55). Thus we conclude (52), and hence (50). By Lemma 3.1 and (23) with $(\ell, i, j) = (2, 1, 1)$ and $(\ell, i, j) = (2, 1, 0)$, we get

$$n^{-3}\sup_{t\in[0,T]}\left\|\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k} - \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\right\| \xrightarrow{\mathbb{P}} 0 \tag{56}$$

as $n\to\infty$ for all $T>0$. As a last step, using (36), we obtain (38).

For (39), consider

$$\sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2\widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}\boldsymbol{V}_{\boldsymbol{\varepsilon}}, \tag{57}$$

where we used Lemma 3.1. Using (23) with $(\ell, i, j) = (4, 0, 2)$, and $(\ell, i, j) = (2, 0, 1)$, we have

$$n^{-5/2} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-5/2} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty,$$

hence (39) follows from Lemma 3.9.

Convergence (40) can be handled in the same way as (39). For completeness we present all of the details. By Lemma 3.1, we have

$$\begin{aligned} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 \widetilde{\boldsymbol{V}}_{\boldsymbol{\xi}} \\ + \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \boldsymbol{V}_{\boldsymbol{\varepsilon}}. \end{aligned} \tag{58}$$

Using (23) with $(\ell, i, j) = (4, 1, 2)$, and $(\ell, i, j) = (2, 1, 1)$, we have

$$n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} U_{k-1} V_{k-1}^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} |U_{k-1} V_{k-1}| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty,$$

hence (40) follows from Lemma 3.10.

Finally, we check condition (ii) of Theorem 3.12, i.e., the conditional Lindeberg condition

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\big(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \,\big|\, \mathcal{F}_{k-1}\big) \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{59}$$

for all $\theta > 0$ and $T > 0$. We have

$$\mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \,\middle|\, \mathcal{F}_{k-1}\right) \leq \theta^{-2} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^4 \,\middle|\, \mathcal{F}_{k-1}\right)$$

and

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

Hence

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

since $\mathbb{E}(\|\boldsymbol{M}_k\|^4) = \mathrm{O}(k^2)$, and further

$$\begin{aligned}\mathbb{E}(\|\boldsymbol{M}_k\|^4 U_{k-1}^4) &\leq \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8)\mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)\\ \mathbb{E}(\|\boldsymbol{M}_k\|^4 V_{k-1}^4) &\leq \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8)\mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)\end{aligned} \tag{60}$$

by Corollary 3.5. This yields (59). □

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 3.5.

Remark 3.14. *Let us introduce the process $\mathcal{Y}_t := \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle$, $t \in \mathbb{R}_+$, where $(\boldsymbol{\mathcal{M}}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the first 2-dimensional equation of SDE* (32)*. Then $\mathcal{Y}_0 = 0$ and by Itô's formula we obtain*

$$\mathrm{d}\mathcal{Y}_t = \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \,\mathrm{d}t + \sqrt{\mathcal{Y}_t^+}\boldsymbol{u}_{\mathrm{left}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \quad t \in \mathbb{R}_+. \tag{61}$$

If $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{u}_{\mathrm{left}}, \boldsymbol{u}_{\mathrm{left}}\rangle \neq 0$, then the process

$$\widetilde{\mathcal{W}}_t := \frac{\boldsymbol{u}_{\mathrm{left}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\boldsymbol{\mathcal{W}}_t}{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{u}_{\mathrm{left}}, \boldsymbol{u}_{\mathrm{left}}\rangle^{1/2}}$$

is a (one-dimensional) standard Wiener process, hence $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ satisfies SDE (8)*. Consequently $(\mathcal{Y}_t)_{t\in\mathbb{R}_+} \stackrel{\mathcal{D}}{=} (\mathcal{Z}_t)_{t\in\mathbb{R}_+}$ and by Theorem 2.4,*

$$\left(\boldsymbol{\mathcal{X}}_t^{(n)}\right)_{t\in\mathbb{R}_+} \stackrel{\mathcal{D}}{\longrightarrow} (\boldsymbol{\mathcal{X}}_t)_{t\in\mathbb{R}_+} \stackrel{\mathcal{D}}{=} (\mathcal{Y}_t\boldsymbol{u}_{\mathrm{right}})_{t\in\mathbb{R}_+}$$

as $n \to \infty$.

The following corollary is the essential piece of our toolkit. We will make heavy use of this statement in the following section.

Corollary 3.15. *Let $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with $\ell = 8$. Then we have*

$$\sum_{k=1}^n \begin{bmatrix} n^{-3}U_{k-1}^2\\ n^{-2}V_{k-1}^2\\ n^{-1}\boldsymbol{M}_k\\ n^{-2}\boldsymbol{M}_kU_{k-1}\\ n^{-3/2}\boldsymbol{M}_kV_{k-1}\end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\\ \boldsymbol{\mathcal{M}}_1\\ \int_0^1 \mathcal{Y}_t\,\mathrm{d}\boldsymbol{\mathcal{M}}_t\\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}}{(1-\lambda_-^2)^{1/2}}\int_0^1 \mathcal{Y}_t\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t\end{bmatrix}$$

as $n \to \infty$.

Remark 3.16. *If the process satisfies* (ND)*, then the above Corollary shows the nonzero limits of the "building blocks" used to construct our estimates. However if the process doesn't satisfy the condition* (ND) *then* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle = 0$ *and the second and fifth coordinates tend to 0.*

Proof. We can write the solution for SDE (32) in the following way

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

for all $t \in \mathbb{R}_+$. By the method of the proof of $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ in Theorem 3.1 in Barczy et al. [2], applying the continuous mapping theorem, one can easily derive

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{X}}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{62}$$

where

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad \widetilde{\boldsymbol{\mathcal{X}}}_t := \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \boldsymbol{u}_{\mathrm{right}},$$

for all $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. More precisely, using that

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

we have

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} = \psi_n(\boldsymbol{\mathcal{Z}}^{(n)}), \qquad n \in \mathbb{N},$$

where the mapping $\psi_n : \mathbb{D}(\mathbb{R}_+, (\mathbb{R}^2)^3) \to \mathbb{D}(\mathbb{R}_+, (\mathbb{R}^2)^4)$ is given by

$$\psi_n(f_1, f_2, f_3)(t) := \begin{bmatrix} \sum_{j=1}^{\lfloor nt \rfloor} \boldsymbol{m}_{\boldsymbol{\xi}}^{\lfloor nt \rfloor - j} \left(f_1\left(\frac{j}{n}\right) - f_1\left(\frac{j-1}{n}\right) + \frac{\boldsymbol{m}_{\boldsymbol{\varepsilon}}}{n} \right) \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$. Further, we have

$$\begin{bmatrix} \widetilde{\boldsymbol{\mathcal{X}}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} = \psi(\boldsymbol{\mathcal{Z}}),$$

where the mapping $\psi : \mathbb{D}(\mathbb{R}_+, (\mathbb{R}^2)^3) \to \mathbb{D}(\mathbb{R}_+, (\mathbb{R}^2)^4)$ is given by

$$\psi(f_1, f_2, f_3)(t) := \begin{bmatrix} \langle \boldsymbol{u}_{\mathrm{left}}, f_1(t) + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \boldsymbol{u}_{\mathrm{right}} \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$ and $t \in \mathbb{R}_+$. By page 603 in Barczy et al. [2], the mappings ψ_n, $n \in \mathbb{N}$, and ψ are measurable (the latter one is continuous too), since the coordinate functions are measurable. Hence, by (30) and the continuous mapping theorem, we have

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} = \psi_n(\boldsymbol{\mathcal{Z}}^{(n)}) \stackrel{\mathcal{D}}{\longrightarrow} \psi(\boldsymbol{\mathcal{Z}}) = \begin{bmatrix} \widetilde{\boldsymbol{\mathcal{X}}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty,$$

as desired. Next, by Lemma 3.11 we get

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-1} \boldsymbol{M}_k \\ n^{-2} \boldsymbol{M}_k U_{k-1} \\ n^{-3/2} \boldsymbol{M}_k V_{k-1} \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \langle \boldsymbol{u}_{\mathrm{left}}, \widetilde{\boldsymbol{\mathcal{X}}}_t \rangle^2 \, \mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda_-^2} \int_0^1 \langle \boldsymbol{u}_{\mathrm{left}}, \widetilde{\boldsymbol{\mathcal{X}}}_t \rangle \, \mathrm{d}t \\ \boldsymbol{\mathcal{M}}_1 \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}$$

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

4 Estimates for the offspring mean matrix

Here is a showcase of the usefulness of the toolkit developed in the previous section. We derive a limit theorem for the estimation of the offspring mean matrix, $\boldsymbol{m}_{\boldsymbol{\xi}}$ in three different settings. The notations introduced in each subsection are unique to that subsection, for example the matrix $\boldsymbol{A}_n$ has a different meaning in each of the following subsections.

4.1 The doubly symmetric process

The aim of this section is to reproduce the results of [10, Theorem 3.1.]. We call a 2-type Galton–Watson process *doubly symmetric* if its offspring mean matrix has the form

$$\boldsymbol{m}_{\boldsymbol{\xi}} = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}.$$

In this case $\gamma = \beta,\ \delta = \alpha$ and condition (CPR) takes the form

$$\alpha \in (0,1), \qquad \beta = 1 - \alpha \in (0,1) \tag{CPR*}$$

We have

$$\begin{aligned}
&\lambda_+ = 1, \quad \lambda_- = 1 - 2\beta, \\
&\boldsymbol{u}_{\mathrm{right}} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{u}_{\mathrm{left}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_{\mathrm{right}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_{\mathrm{left}} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \\
&U_k = \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{X}_k \rangle = X_{k,1} + X_{k,2}, \quad V_k = \langle \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{X}_k \rangle = \frac{1}{2}\left(X_{k,2} - X_{k,1}\right)
\end{aligned}$$

Lemma 4.1. *The joint CLS estimator for α and β has the form*

$$\begin{bmatrix} \widehat{\alpha}_n \\ \widehat{\beta}_n \end{bmatrix} = \boldsymbol{A}_n^{-1} \boldsymbol{B}_n,$$

on the set $\Omega_n := \{\omega \in \Omega \colon \det(\boldsymbol{A}_n) > 0\}$, where

$$\begin{aligned}
\boldsymbol{A}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) &= \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1} \end{bmatrix}^2 \\
\boldsymbol{B}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) &= \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1} \end{bmatrix} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}\right).
\end{aligned}$$

Proof. Define the function $Q_n : \left(\mathbb{R}^2\right)^{n+1} \to \mathbb{R}$ as

$$Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \begin{bmatrix} a \\ b \end{bmatrix}\right) = \sum_{k=1}^{n} \left\| \boldsymbol{x}_k - \begin{bmatrix} a & b \\ b & a \end{bmatrix} \boldsymbol{x}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

with the convention that $\boldsymbol{x}_0 = \boldsymbol{0}$. A CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is a measurable function $F_n : \left(\mathbb{R}^2\right)^n \to \mathbb{R}^2$ such that

$$Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, F_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right)\right) = \inf_{a,b\in\mathbb{R}} Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \begin{bmatrix} a \\ b \end{bmatrix}\right)$$

for all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$. We need to show that

$$F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \boldsymbol{A}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^{-1} \boldsymbol{B}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$$

on the set

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

Fix $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and find the critical points (where all partial derivatives vanishes). The function Q_n can be written in the form

$$\begin{aligned} Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \begin{bmatrix} a \\ b \end{bmatrix}\right) &= \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right)^2 \\ &\quad + \sum_{k=1}^{n} \left(x_{k,2} - b x_{k-1,1} - a x_{k-1,2} - m_{\varepsilon,2}\right)^2. \end{aligned}$$

To find the critical points we have to solve

$$\begin{cases} \frac{\partial}{\partial a} Q_n = 0 \\ \frac{\partial}{\partial b} Q_n = 0 \end{cases} \implies \begin{cases} \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right) x_{k-1,1} \\ \quad + \sum_{k=1}^{n} \left(x_{k,2} - b x_{k-1,1} - a x_{k-1,2} - m_{\varepsilon,2}\right) x_{k-1,2} = 0 \\ \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right) x_{k-1,2} \\ \quad + \sum_{k=1}^{n} \left(x_{k,2} - b x_{k-1,1} - a x_{k-1,2} - m_{\varepsilon,2}\right) x_{k-1,1} = 0 \end{cases}$$

Rearranging gives us the equation

$$\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1}^2 + x_{k-1,2}^2 & 2x_{k-1,1}x_{k-1,2} \\ 2x_{k-1,1}x_{k-1,2} & x_{k-1,1}^2 + x_{k-1,2}^2 \end{bmatrix}\right)\begin{bmatrix} a \\ b \end{bmatrix}$$

$$= \sum_{k=1}^{n}\begin{bmatrix} (x_{k,1} - m_{\varepsilon,1})x_{k-1,1} + (x_{k,2} - m_{\varepsilon,2})x_{k-1,2} \\ (x_{k,1} - m_{\varepsilon,1})x_{k-1,2} + (x_{k,2} - m_{\varepsilon,2})x_{k-1,1} \end{bmatrix},$$

which can be rewritten as

$$\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix}^2\right)\begin{bmatrix} a \\ b \end{bmatrix} = \sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix}(\boldsymbol{x}_k - \boldsymbol{m}_\varepsilon),$$

or using the notation for $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$

$$\boldsymbol{A}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\begin{bmatrix} a \\ b \end{bmatrix} = \boldsymbol{B}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n).$$

So $F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is the only critical point if $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in D(F_n)$. However we still have to prove that it is in fact a minimum, we will use the second order derivatives. The Hessian matrix of Q_n is

$$\boldsymbol{H}_n = 2\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1}^2 + x_{k-1,2}^2 & 2x_{k-1,1}x_{k-1,2} \\ 2x_{k-1,1}x_{k-1,2} & x_{k-1,1}^2 + x_{k-1,2}^2 \end{bmatrix} = 2\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix}^2,$$

as we can see it does not depend on the parameters a, b. We are going to show that $\boldsymbol{H}_n$ is positive definite, we are going to do this by the equivalent condition that its leading principal minors are all positive. Remember we are working on the set $D(F_n)$, thus

$$\begin{aligned} 0 < \det(\boldsymbol{A}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) &= \det\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix}^2\right) \\ &= \left(\sum_{k=1}^{n} x_{k-1,1}^2 + x_{k-1,2}^2\right)^2 - 4\left(\sum_{k=1}^{n} x_{k-1,1}x_{k-1,2}\right)^2 \\ &< \left(\sum_{k=1}^{n} x_{k-1,1}^2 + x_{k-1,2}^2\right)^2. \end{aligned}$$

Consequently $\sum_{k=1}^{n} x_{k-1,1}^2 + x_{k-1,2}^2 > 0$, the matrix $\boldsymbol{H}_n$ is positive definite, and $F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is the minimum of Q_n. This proves the formula for $\begin{bmatrix} \widehat{\alpha}_n \\ \widehat{\beta}_n \end{bmatrix}$.

□

We have the formula for the estimators, but in order to prove limit theorems we also need a formula for the difference from the real parameters.

Corollary 4.2. *The difference of the CLS estimator and the real parameter values can be written as*

$$\begin{bmatrix}\widehat{\alpha}_n - \alpha \\ \widehat{\beta}_n - \beta\end{bmatrix} = \boldsymbol{A}_n^{-1}\boldsymbol{C}_n,$$

where

$$\boldsymbol{C}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix} \boldsymbol{M}_k.$$

Proof. To get the formula for the difference, one can write

$$\begin{aligned}
\begin{bmatrix}\widehat{\alpha}_n - \alpha \\ \widehat{\beta}_n - \beta\end{bmatrix} &= \boldsymbol{A}_n^{-1}\boldsymbol{B}_n - \begin{bmatrix}\alpha \\ \beta\end{bmatrix} = \boldsymbol{A}_n^{-1}\left(\boldsymbol{B}_n - \boldsymbol{A}_n \begin{bmatrix}\alpha \\ \beta\end{bmatrix}\right) \\
&= \boldsymbol{A}_n^{-1}\left(\sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix}(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}) - \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix}^2 \begin{bmatrix}\alpha \\ \beta\end{bmatrix}\right) \\
&= \boldsymbol{A}_n^{-1}\left(\sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix}(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}})\right) \\
&= \boldsymbol{A}_n^{-1}\left(\sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix}\boldsymbol{M}_k\right).
\end{aligned}$$

□

Theorem. 4.3. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type doubly symmetric Galton–Watson process with immigration satisfying conditions* (CPR*), (ZS) *and* (M) *with* $\ell = 8$*. If the process satisfies* (ND) *as well, then the probability of the existence of the estimators* $\widehat{\alpha}_n$ *and* $\widehat{\beta}_n$ *tends to 1 as* $n \to \infty$*, and further*

$$n^{1/2}\begin{bmatrix}\widehat{\alpha}_n - \alpha \\ \widehat{\beta}_n - \beta\end{bmatrix} \xrightarrow{\mathcal{D}} \sqrt{\alpha\beta}\,\frac{\int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{Y}_t \,\mathrm{d}t}\begin{bmatrix}1 \\ -1\end{bmatrix} \tag{63}$$

as $n \to \infty$, *where* $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ *is defined in Remark 3.14.*

Proof. We start by finding the nonzero limit of $\det(\boldsymbol{A}_n)$. This will allow us to prove the asymptotic existence of the estimator. We have

$$\begin{aligned}
\det(\boldsymbol{A}_n) &= \det\left(\sum_{k=1}^{n}\begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1}\end{bmatrix}^2\right) \\
&= \det\left(\sum_{k=1}^{n}\begin{bmatrix} X_{k-1,1}^2 + X_{k-1,2}^2 & 2X_{k-1,1}X_{k-1,2} \\ 2X_{k-1,1}X_{k-1,2} & X_{k-1,1}^2 + X_{k-1,2}^2\end{bmatrix}\right) \\
&= \left(\sum_{k=1}^{n} X_{k-1,1}^2 + X_{k-1,2}^2\right)^2 - \left(\sum_{k=1}^{n} 2X_{k-1,1}X_{k-1,2}\right)^2 \\
&= \left(\sum_{k=1}^{n}\left[X_{k-1,1} + X_{k-1,2}\right]^2\right)\left(\sum_{k=1}^{n}\left[X_{k-1,1} - X_{k-1,2}\right]^2\right).
\end{aligned}$$

The decomposition (20) yields

$$\det(\boldsymbol{A}_n) = 4\sum_{k=1}^{n} U_{k-1}^2 \sum_{k=1}^{n} V_{k-1}^2, \tag{64}$$

for all $n \in \mathbb{N}$. By Corollary 3.15 and the continuous mapping theorem we have

$$n^{-5}\det(\boldsymbol{A}_n) \xrightarrow{\mathcal{D}} \frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \tag{65}$$

as $n \to \infty$. The process satisfies (ZS), therefore $\boldsymbol{m}_{\boldsymbol{\varepsilon}} \neq \boldsymbol{0}$, consequently by the SDE (61), we have

$$\mathbb{P}(\mathcal{Y}_t = 0 \ \text{ for all } \ t \in [0,1]) = 0.$$

This implies

$$\mathbb{P}\left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) = 1.$$

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. Note that

$$\mathbb{P}(\Omega_n) = \mathbb{P}\left(\det(\boldsymbol{A}_n) > 0\right) = \mathbb{P}\left(n^{-5}\det(\boldsymbol{A}_n) > 0\right).$$

If the process satisfies (ND), then $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle > 0$, and by (65),

$$\begin{aligned}
\mathbb{P}(\Omega_n) &\to \mathbb{P}\left(\frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) \\
&= \mathbb{P}\left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) = 1
\end{aligned}$$

as $n \to \infty$. This proves the asymptotic existence of the estimators.

Next we turn to prove convergence (63). We do this by finding stochastic expansions for the product $\widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n$. We will use Corollary 3.15 again, so in fact we are proving a joint convergence of the sequence $(\det(\boldsymbol{A}_n), \widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n)_{n \in \mathbb{N}}$. We have

$$\begin{aligned}\widetilde{\boldsymbol{A}}_n &= \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1}^2 + X_{k-1,2}^2 & -2X_{k-1,1}X_{k-1,2} \\ -2X_{k-1,1}X_{k-1,2} & X_{k-1,1}^2 + X_{k-1,2}^2 \end{bmatrix} \\ &= \frac{1}{2} \sum_{k=1}^{n} \begin{bmatrix} 4V_{k-1}^2 + U_{k-1}^2 & 4V_{k-1}^2 - U_{k-1}^2 \\ 4V_{k-1}^2 - U_{k-1}^2 & 4V_{k-1}^2 + U_{k-1}^2 \end{bmatrix} \\ &= \frac{1}{2} \sum_{k=1}^{n} U_{k-1}^2 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + 2 \sum_{k=1}^{n} V_{k-1}^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{C}_n &= \sum_{k=1}^{n} \begin{bmatrix} X_{k-1,1} & X_{k-1,2} \\ X_{k-1,2} & X_{k-1,1} \end{bmatrix} \boldsymbol{M}_k \\ &= \frac{1}{2} \sum_{k=1}^{n} \begin{bmatrix} U_{k-1} - 2V_{k-1} & U_{k-1} + 2V_{k-1} \\ U_{k-1} + 2V_{k-1} & U_{k-1} - 2V_{k-1} \end{bmatrix} \boldsymbol{M}_k \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \sum_{k=1}^{n} \boldsymbol{M}_k U_{k-1} + \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \sum_{k=1}^{n} \boldsymbol{M}_k V_{k-1}.\end{aligned}$$

These reformulations in terms of U_{k-1} and V_{k-1} along with Corollary 3.15 imply stochastic expansions

$$\begin{aligned}\widetilde{\boldsymbol{A}}_n &= n^3 A_{n,1} \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} + n^2 A_{n,2} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \\ \boldsymbol{C}_n &= \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} n^2 \boldsymbol{C}_{n,1} + \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} n^{3/2} \boldsymbol{C}_{n,2},\end{aligned}$$

where

$$
\begin{aligned}
A_{n,1} &= n^{-3} \sum_{k=1}^{n} U_{k-1}^2 \xrightarrow{\mathcal{D}} \mathcal{A}_1 := \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t, \\
A_{n,2} &= n^{-2} \sum_{k=1}^{n} V_{k-1}^2 \xrightarrow{\mathcal{D}} \mathcal{A}_2 := \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1 - \lambda_-^2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}t, \\
\boldsymbol{C}_{n,1} &= n^{-2} \sum_{k=1}^{n} \boldsymbol{M}_k U_{k-1} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_1 := \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t, \\
\boldsymbol{C}_{n,2} &= n^{-3/2} \sum_{k=1}^{n} \boldsymbol{M}_k V_{k-1} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 := \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1 - \lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t,
\end{aligned}
$$

jointly as $n \to \infty$. Multiplying these together we get

$$
\widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n = n^5 \boldsymbol{D}_{n,1} + n^{9/2} \boldsymbol{D}_{n,2} + n^4 \boldsymbol{D}_{n,3} + n^{7/2} \boldsymbol{D}_{n,4}, \tag{66}
$$

where

$$
\boldsymbol{D}_{n,1} := A_{n,1} \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \boldsymbol{C}_{n,1} = \boldsymbol{0}
$$

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

$$
\begin{aligned}
\boldsymbol{D}_{n,2} &:= A_{n,1} \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{C}_{n,2} \xrightarrow{\mathcal{D}} \mathcal{A}_1 \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{\mathcal{C}}_2, \\
\boldsymbol{D}_{n,3} &:= A_{n,2} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \boldsymbol{C}_{n,1} \xrightarrow{\mathcal{D}} \mathcal{A}_2 \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \boldsymbol{\mathcal{C}}_1, \\
\boldsymbol{D}_{n,4} &:= A_{n,2} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{C}_{n,2} = \boldsymbol{0}
\end{aligned}
$$

as $n \to \infty$. Putting it all together we get

$$
n^{-9/2} \widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n \xrightarrow{\mathcal{D}} \mathcal{A}_1 \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{\mathcal{C}}_2
$$

as $n \to \infty$, where

$$
\begin{aligned}
\mathcal{A}_1 \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{\mathcal{C}}_2 &= \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1 - \lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \\
&\stackrel{\mathcal{D}}{=} \frac{2 \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{(1 - \lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t \begin{bmatrix} 1 \\ -1 \end{bmatrix},
\end{aligned}
$$

since

$$\begin{bmatrix}-1 & 1\\ 1 & -1\end{bmatrix} \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\widetilde{\boldsymbol{\mathcal{W}}}_t = 2\begin{bmatrix}\boldsymbol{v}_{\mathrm{left}}^\top\\ -\boldsymbol{v}_{\mathrm{left}}^\top\end{bmatrix} \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\widetilde{\boldsymbol{\mathcal{W}}}_t = 2\left\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\boldsymbol{v}_{\mathrm{left}}, \widetilde{\boldsymbol{\mathcal{W}}}_t\right\rangle \begin{bmatrix}1\\ -1\end{bmatrix}$$
$$\stackrel{\mathcal{D}}{=} 2\left\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\right\rangle^{1/2} \widetilde{\mathcal{W}}_t \begin{bmatrix}1\\ -1\end{bmatrix}.$$

We have shown the joint convergence

$$\begin{bmatrix} n^{-5}\det(\boldsymbol{A}_n)\\ n^{-9/2}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n\end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\\ \frac{2\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle}{(1-\lambda_-^2)^{1/2}}\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t \begin{bmatrix}1\\ -1\end{bmatrix}\end{bmatrix}$$

as $n\to\infty$. Using the continuous mapping theorem on this result gives us the desired convergence

$$n^{1/2}\begin{bmatrix}\widehat{\alpha}_n-\alpha\\ \widehat{\beta}_n-\beta\end{bmatrix} = \frac{n^{-9/2}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n}{n^{-5}\det(\boldsymbol{A}_n)} \stackrel{\mathcal{D}}{\longrightarrow} \sqrt{\alpha\beta}\,\frac{\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{Y}_t\,\mathrm{d}t}\begin{bmatrix}1\\ -1\end{bmatrix},$$

since $1-\lambda_-^2 = 1-(1-2\beta)^2 = 4\alpha\beta$. □

In the critical, doubly symmetric case the spectral radius of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is

$$\varrho = \lambda_+ = \alpha+\beta,$$

so we can define a natural estimator for ϱ by $\widehat{\varrho}_n := \widehat{\alpha}_n + \widehat{\beta}_n$.

Theorem. 4.4. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type doubly symmetric Galton–Watson process with immigration satisfying conditions* (CPR*), (ZS) *and* (M) *with* $\ell = 8$. *If the process satisfies* (ND) *as well, then the probability of the existence of the estimator* $\widehat{\varrho}_n$ *tends to 1 as* $n\to\infty$, *and further*

$$n\left(\widehat{\varrho}_n - 1\right) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\int_0^1 \mathcal{Y}_t\,\mathrm{d}\left(\mathcal{Y}_t - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle t\right)}{\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t} \tag{67}$$

as $n\to\infty$.

Proof. Using Lemma 4.1 we can write

$$\widehat{\varrho}_n - 1 = \widehat{\alpha}_n - \alpha + \widehat{\beta}_n - \beta = \begin{bmatrix}1\\ 1\end{bmatrix}^\top \begin{bmatrix}\widehat{\alpha}_n-\alpha\\ \widehat{\beta}_n-\beta\end{bmatrix} = \frac{\boldsymbol{u}_{\mathrm{left}}^\top \widetilde{\boldsymbol{A}}_n\boldsymbol{C}_n}{\det(\boldsymbol{A}_n)}.$$

By stochastic expansion (66) we get

$$\boldsymbol{u}_{\mathrm{left}}^\top \widetilde{\boldsymbol{A}}_n\boldsymbol{C}_n = n^{9/2}\boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{D}_{n,2} + n^4\boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{D}_{n,3},$$

where

$$\boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{D}_{n,2} = A_{n,1} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{C}_{n,2} = \boldsymbol{0},$$

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

$$\begin{aligned}\boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{D}_{n,3} = A_{n,2} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \boldsymbol{C}_{n,1}, &= 4A_{n,2} \boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{C}_{n,1} \\ &\xrightarrow{\mathcal{D}} 4\mathcal{A}_2 \boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{\mathcal{C}}_1 = \frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1 - \lambda_-^2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \boldsymbol{u}_{\mathrm{left}}^\top \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t\end{aligned}$$

as $n \to \infty$. By (65) we have

$$\begin{bmatrix} n^{-5} \det(\boldsymbol{A}_n) \\ n^{-4} \boldsymbol{u}_{\mathrm{left}}^\top \widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \\ \frac{4\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \, \boldsymbol{u}_{\mathrm{left}}^\top \int_0^1 \mathcal{Y}_t \, \mathrm{d}\boldsymbol{\mathcal{M}}_t \end{bmatrix}$$

as $n \to \infty$. Using the continuous mapping theorem on this result gives us the desired convergence

$$n\,(\widehat{\varrho}_n - 1) = \frac{n^{-4} \boldsymbol{u}_{\mathrm{left}}^\top \widetilde{\boldsymbol{A}}_n \boldsymbol{C}_n}{n^{-5} \det(\boldsymbol{A}_n)} \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{Y}_t \, \mathrm{d}\,(\mathcal{Y}_t - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle t)}{\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t}$$

as $n \to \infty$, since $\boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{\mathcal{M}}_t = \mathcal{Y}_t - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle t$. □

4.2 The general process with known immigration mean

Next up we reproduce the results in the general case [16, Theorem 3.1.]. There are two main differences. The first is the lighter assumptions on the structure of the offspring mean matrix is, we only assume that the matrix is positively regular. Second we needed a more convoluted approach in treating the estimator for the criticality parameter as it is now a non-linear function of the matrix elements.

We structured the proofs in a similar way to the previous section, the reason for this is to show how the method can be applied in a streamlined fashion.

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 $\mathbb{R}^{2\times 2}$.

Lemma 4.5. *The CLS estimator of* $\boldsymbol{m}_{\boldsymbol{\xi}}$ *has the form* $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} = \boldsymbol{B}_n \boldsymbol{A}_n^{-1}$ *on the set* $\Omega_n := \{\omega \in \Omega \colon \det(\boldsymbol{A}_n) > 0\}$, *where*

$$\boldsymbol{A}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{X}_{k-1} \boldsymbol{X}_{k-1}^\top,$$

$$\boldsymbol{B}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}\right) \boldsymbol{X}_{k-1}^\top.$$

Proof. Define the function $Q_n : \left(\mathbb{R}^2\right)^n \times \mathbb{R}^{2\times 2} \to \mathbb{R}$ as

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

with the convention that $\boldsymbol{x}_0 = \boldsymbol{0}$. A CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ is a measurable function $F_n : \left(\mathbb{R}^2\right)^n \to \mathbb{R}^{2\times 2}$ such that

$$Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, F_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right)\right) = \inf_{\boldsymbol{m} \in \mathbb{R}^{2\times 2}} Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{m}\right)$$

for all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$. We need to show that

$$F_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) = \boldsymbol{B}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) \boldsymbol{A}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right)^{-1}$$

on the set

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

Fix $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and find the critical points (where all partial derivatives vanishes). Let

$$\boldsymbol{m} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$\begin{aligned} Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{m}\right) = & \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right)^2 \\ & + \sum_{k=1}^{n} \left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - m_{\varepsilon,2}\right)^2. \end{aligned}$$

To find the critical points we have to solve

$$\begin{cases} \frac{\partial}{\partial a} Q_n = 0 \\ \frac{\partial}{\partial b} Q_n = 0 \\ \frac{\partial}{\partial c} Q_n = 0 \\ \frac{\partial}{\partial d} Q_n = 0 \end{cases} \implies \begin{cases} -2 \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right) x_{k-1,1} = 0 \\ -2 \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - m_{\varepsilon,1}\right) x_{k-1,2} = 0 \\ -2 \sum_{k=1}^{n} \left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - m_{\varepsilon,2}\right) x_{k-1,1} = 0 \\ -2 \sum_{k=1}^{n} \left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - m_{\varepsilon,2}\right) x_{k-1,2} = 0 \end{cases}$$

Rearranging gives us the equation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\sum_{k=1}^{n} \boldsymbol{x}_{k-1} \boldsymbol{x}_{k-1}^{\top}\right) = \sum_{k=1}^{n} \left(\boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}\right) \boldsymbol{x}_{k-1}^{\top},$$

which using the notation for $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ can be written in the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \boldsymbol{A}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) = \boldsymbol{B}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right).$$

So $F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is the only critical point if $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in D(F_n)$. However we still have to prove that it is in fact a minimum, we will use the second order derivatives. The Hessian matrix of Q_n is

$$\boldsymbol{H}_n = 2 \sum_{k=1}^{n} \begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1} x_{k-1,2} & 0 & 0 \\ x_{k-1,1} x_{k-1,2} & x_{k-1,2}^2 & 0 & 0 \\ 0 & 0 & x_{k-1,1}^2 & x_{k-1,1} x_{k-1,2} \\ 0 & 0 & x_{k-1,1} x_{k-1,2} & x_{k-1,2}^2 \end{bmatrix},$$

as we can see it does not depend on the parameters a, b, c, d. We are going to show that $\boldsymbol{H}_n$ is positive definite. Note that for any vector $\boldsymbol{v} \in \mathbb{R}^4$

$$\begin{aligned}
\boldsymbol{v}^\top \boldsymbol{H}_n \boldsymbol{v} &= 2 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}^\top \left(\sum_{k=1}^n \boldsymbol{x}_{k-1} \boldsymbol{x}_{k-1}^\top \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + 2 \begin{bmatrix} v_3 \\ v_4 \end{bmatrix}^\top \left(\sum_{k=1}^n \boldsymbol{x}_{k-1} \boldsymbol{x}_{k-1}^\top \right) \begin{bmatrix} v_3 \\ v_4 \end{bmatrix} \\
&= 2 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}^\top \boldsymbol{A}_n (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + 2 \begin{bmatrix} v_3 \\ v_4 \end{bmatrix}^\top \boldsymbol{A}_n (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \begin{bmatrix} v_3 \\ v_4 \end{bmatrix}.
\end{aligned}$$

Therefore it is sufficient to show that the matrix $\boldsymbol{A}_n (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is positive definite, we are going to do this by the equivalent condition that its leading principal minors are all positive. Remember we are working on the set $D(F_n)$, thus

$$\begin{aligned}
0 < \det (\boldsymbol{A}_n (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)) &= \det \left(\sum_{k=1}^n \boldsymbol{x}_{k-1} \boldsymbol{x}_{k-1}^\top \right) \\
&= \sum_{k=1}^n x_{k-1,1}^2 \sum_{k=1}^n x_{k-1,2}^2 - \left(\sum_{k=1}^n x_{k-1,1} x_{k-1,2} \right)^2 \\
&< \sum_{k=1}^n x_{k-1,1}^2 \sum_{k=1}^n x_{k-1,2}^2 .
\end{aligned}$$

Consequently $\sum_{k=1}^n x_{k-1,1}^2 > 0$, the matrices $\boldsymbol{A}_n (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ and $\boldsymbol{H}_n$ are positive definite, and $F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is the minimum of Q_n. This proves the formula for $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)}$. □

We have the formula for the estimator. To prove limit theorems we have to handle the difference between the estimator and the real parameter value.

Corollary 4.6. *The difference of the CLS estimator and the real parameter value can be expressed as*

$$\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} = \boldsymbol{C}_n \boldsymbol{A}_n^{-1},$$

where

$$\boldsymbol{C}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{k=1}^n \boldsymbol{M}_k \boldsymbol{X}_{k-1}^\top .$$

Proof. We can get the formula for the difference in the following way

$$
\begin{aligned}
\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} &= \boldsymbol{B}_n \boldsymbol{A}_n^{-1} - \boldsymbol{m}_{\boldsymbol{\xi}} = \left(\boldsymbol{B}_n - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{A}_n\right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}\right) \boldsymbol{X}_{k-1}^{\top} - \boldsymbol{m}_{\boldsymbol{\xi}} \sum_{k=1}^{n} \boldsymbol{X}_{k-1} \boldsymbol{X}_{k-1}^{\top} \right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}}\right) \boldsymbol{X}_{k-1}^{\top} \right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \boldsymbol{M}_k \boldsymbol{X}_{k-1}^{\top} \right) \boldsymbol{A}_n^{-1}.
\end{aligned}
$$

□

In the critical case, by (7) and the continuous mapping theorem, one can derive

$$
n^{-3} \boldsymbol{A}_n \xrightarrow{\mathcal{D}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \boldsymbol{u}_{\mathrm{right}} \boldsymbol{u}_{\mathrm{right}}^{\top} =: \boldsymbol{\mathcal{A}}
$$

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

$$
\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} = \boldsymbol{C}_n \boldsymbol{A}_n^{-1} = \frac{1}{\det(\boldsymbol{A}_n)} \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n, \qquad n \in \mathbb{N}, \tag{68}
$$

on the set Ω_n, where $\widetilde{\boldsymbol{A}}_n$ denotes the adjugate matrix of $\boldsymbol{A}_n$ (i.e., the matrix of cofactors) given by

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

We can find the limit for the difference $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}$ by describing the asymptotic behaviour of the sequence $(\det(\boldsymbol{A}_n), \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n)_{n \in \mathbb{N}}$.

Theorem. 4.7. *Let $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ be a 2-type Galton–Watson process with immigration satisfying conditions* (CPR), (ZS) *and* (M) *with $\ell = 8$. If the process satisfies* (ND) *as well, then the probability of the existence of the estimator $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)}$ tends to 1 as $n \to \infty$, and further*

$$
n^{1/2} (\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}) \xrightarrow{\mathcal{D}} \frac{(1 - \lambda_-^2)^{1/2}}{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}} \frac{\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d} \widetilde{\boldsymbol{\mathcal{W}}}_t}{\int_0^1 \mathcal{Y}_t \, \mathrm{d}t} \boldsymbol{v}_{\mathrm{left}}^{\top} \tag{69}
$$

as $n \to \infty$, *with* $\mathcal{Y}_t := \langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_\varepsilon \rangle$, $t \in \mathbb{R}_+$, *where* $(\boldsymbol{\mathcal{M}}_t)_{t\in\mathbb{R}_+}$ *is the unique strong solution of the SDE*

$$\begin{aligned} \mathrm{d}\boldsymbol{\mathcal{M}}_t &= (\langle \boldsymbol{u}_{\text{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_\varepsilon \rangle^+)^{1/2} \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \\ \boldsymbol{\mathcal{M}}_0 &= \boldsymbol{0}, \end{aligned}$$

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.*

Proof. We start by finding the nonzero limit of $\det(\boldsymbol{A}_n)$. This will allow us to prove the asymptotic existence of the estimator. The decomposition (20) yields

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

Using Lemma 2.2 we get

$$\det(\boldsymbol{A}_n) = \left(\sum_{k=1}^{n-1} U_k^2\right)\left(\sum_{k=1}^{n-1} V_k^2\right) - \left(\sum_{k=1}^{n-1} U_k V_k\right)^2, \tag{70}$$

for all $n \in \mathbb{N}$. By Corollary 3.15, Lemma 3.9 and the continuous mapping theorem we have

$$n^{-5}\det(\boldsymbol{A}_n) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}}\rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \tag{71}$$

as $n \to \infty$. The process satisfies (ZS), therefore $\boldsymbol{m}_\varepsilon \neq \boldsymbol{0}$, consequently by the SDE (61), we have

$$\mathbb{P}(\mathcal{Y}_t = 0 \ \text{ for all } \ t \in [0,1]) = 0.$$

This implies

$$\mathbb{P}\left(\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t > 0\right) = 1.$$

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. Note that

$$\mathbb{P}(\Omega_n) = \mathbb{P}\left(\det(\boldsymbol{A}_n) > 0\right) = \mathbb{P}\left(n^{-5}\det(\boldsymbol{A}_n) > 0\right).$$

If the process satisfies (ND), then $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle > 0$, and by (71),

$$\begin{aligned}\mathbb{P}(\Omega_n) &\to \mathbb{P}\left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) \\ &= \mathbb{P}\left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) = 1\end{aligned}$$

as $n \to \infty$. This proves the asymptotic existence of the estimator $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}$.

Next we turn to prove convergence (69). We do this by finding stochastic expansions for the product $\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n$. We will use Corollary 3.15 again, so in fact we are proving a joint convergence of the sequence $(\det(\boldsymbol{A}_n), \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n)_{n\in\mathbb{N}}$.

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 (20), we can write

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

Similarly to Lemmas 2.2 and 2.3 one can show

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

therefore

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

Denote the standard base of $\mathbb{R}^2$ by

$$\boldsymbol{e}_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{e}_2 := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Corollary 3.15 implies stochastic expansions

$$\sum_{k=1}^{n} \boldsymbol{M}_k \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^\top = n^2 \boldsymbol{C}_{n,1} + n^{3/2} \boldsymbol{C}_{n,2}, \tag{72}$$

$$\sum_{\ell=1}^{n} \begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix} \begin{bmatrix} U_{\ell-1} \\ V_{\ell-1} \end{bmatrix}^\top = n^3 \boldsymbol{A}_{n,1} + n^{5/2} \boldsymbol{A}_{n,2} + n^2 \boldsymbol{A}_{n,3}, \tag{73}$$

where

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

$$\boldsymbol{C}_{n,2} := n^{-3/2} \sum_{k=1}^{n} \boldsymbol{M}_k V_{k-1} \boldsymbol{e}_2^\top$$

$$\stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{C}}_2 := \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{e}_2^\top,$$

and

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

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

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

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

$$\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n = (n^5 \boldsymbol{D}_{n,1} + n^{9/2} \boldsymbol{D}_{n,2} + n^4 \boldsymbol{D}_{n,3} + n^{7/2} \boldsymbol{D}_{n,4}) \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}, \tag{74}$$

where

$$\begin{aligned} \boldsymbol{D}_{n,1} :&= \boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \\ &= n^{-5} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{M}_k U_{k-1} U_{\ell-1}^2 \boldsymbol{e}_1^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0} \end{aligned} \tag{75}$$

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

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

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

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

as $n \to \infty$. Finally putting it all together we get

$$n^{-9/2} \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}$$

as $n \to \infty$, where

$$\begin{aligned} &\boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix} \\ &= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{e}_2^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix} \\ &= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{v}_{\mathrm{left}}^\top. \end{aligned}$$

We have shown the joint convergence

$$\begin{bmatrix} n^{-5} \det(\boldsymbol{A}_n) \\ n^{-9/2} \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{v}_{\mathrm{left}}^\top \end{bmatrix}$$

as $n \to \infty$. Using the continuous mapping theorem on this result gives us the desired convergence

$$n^{1/2}(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}) = \frac{n^{-9/2} \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n}{n^{-5} \det(\boldsymbol{A}_n)} \xrightarrow{\mathcal{D}} \frac{(1-\lambda_-^2)^{1/2}}{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}} \frac{\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t}{\int_0^1 \mathcal{Y}_t \, \mathrm{d}t} \boldsymbol{v}_{\mathrm{left}}^\top.$$

□

Theorem. 4.8. *Let* $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration satisfying conditions* (CPR), (ZS) *and* (M) *with* $\ell = 8$. *If the process satisfies* (ND) *as well, then the probability of the existence of the estimator* $\widehat{\varrho}_n$ *tends to 1 as* $n \to \infty$, *and further*

$$n(\widehat{\varrho}_n - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{Y}_t \, \mathrm{d}(\mathcal{Y}_t - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle t)}{\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t} \tag{76}$$

as $n \to \infty$.

Proof. Using the estimates

$$\widehat{\alpha}_n := \boldsymbol{e}_1^\top \widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \boldsymbol{e}_1, \qquad \widehat{\beta}_n := \boldsymbol{e}_1^\top \widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \boldsymbol{e}_2,$$
$$\widehat{\gamma}_n := \boldsymbol{e}_2^\top \widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \boldsymbol{e}_1, \qquad \widehat{\delta}_n := \boldsymbol{e}_2^\top \widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \boldsymbol{e}_2,$$

we can write $\widehat{\varrho}_n - 1$ in the form

$$\begin{aligned}
&\widehat{\varrho}_n - 1 = \widehat{\varrho}_n - \varrho \\
&= \frac{(\widehat{\alpha}_n - \alpha) + \left(\widehat{\delta}_n - \delta\right) + \sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} - \sqrt{(\alpha-\delta)^2 + 4\beta\gamma}}{2} \\
&= \frac{\widehat{\alpha}_n - \alpha + \widehat{\delta}_n - \delta}{2} + \frac{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 - (\alpha-\delta)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n - 4\beta\gamma}{2\left(\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + \sqrt{(\alpha-\delta)^2 + 4\beta\gamma}\right)} \\
&= \frac{c_n}{2\left(\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + (1-\lambda_-)\right)},
\end{aligned}$$

where

$$\begin{aligned}
c_n := (\widehat{\alpha}_n - \alpha) &\left[\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + (1-\lambda_-) + (\widehat{\alpha}_n + \alpha) - \left(\widehat{\delta}_n + \delta\right)\right] \\
+ \left(\widehat{\delta}_n - \delta\right) &\left[\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + (1-\lambda_-) - (\widehat{\alpha}_n + \alpha) + \left(\widehat{\delta}_n + \delta\right)\right] \\
+ 4\left(\widehat{\beta}_n - \beta\right)&\widehat{\gamma}_n + 4(\widehat{\gamma}_n - \gamma)\beta.
\end{aligned}$$

We handle c_n by replacing some terms with their limits, then we prove that the difference tends to 0 with the right scaling. Finally we show that the resulting simpler expression has the right limit with the same scaling. Slutsky's lemma and (69) imply $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \xrightarrow{\mathcal{D}} \boldsymbol{0}$, and hence $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \xrightarrow{\mathbb{P}} \boldsymbol{0}$ as $n \to \infty$. Thus $\widehat{\gamma}_n \xrightarrow{\mathbb{P}} \gamma$, and

$$\begin{aligned}
&(\widehat{\alpha}_n + \alpha) - \left(\widehat{\delta}_n + \delta\right) \xrightarrow{\mathbb{P}} 2(\alpha - \delta), \\
&\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n \xrightarrow{\mathbb{P}} (\alpha-\delta)^2 + 4\beta\gamma = (2-\alpha-\delta)^2 = (1-\lambda_-)^2
\end{aligned}$$

as $n \to \infty$. The aim of the following discussion is to show $n(c_n - d_n) \xrightarrow{\mathcal{D}} 0$ as $n \to \infty$, where

$$\begin{aligned} d_n &:= (\widehat{\alpha}_n - \alpha)\left[2(1-\lambda_-) + 2\alpha - 2\delta\right] + \left(\widehat{\delta}_n - \delta\right)\left[2(1-\lambda_-) - 2\alpha + 2\delta\right] \\ &\quad + 4\left(\widehat{\beta}_n - \beta\right)\gamma + 4\left(\widehat{\gamma}_n - \gamma\right)\beta \\ &= 4(1-\delta)(\widehat{\alpha}_n - \alpha) + 4\gamma(\widehat{\beta}_n - \beta) + 4\beta(\widehat{\gamma}_n - \gamma) + 4(1-\alpha)(\widehat{\delta}_n - \delta). \end{aligned}$$

We have

$$\begin{aligned} n\left(c_n - d_n\right) = n &\left[\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} - (1-\lambda_-)\right]\left[(\widehat{\alpha}_n - \alpha) + \left(\widehat{\delta}_n - \delta\right)\right] \\ &+ n\left[(\widehat{\alpha}_n - \alpha) - \left(\widehat{\delta}_n - \delta\right)\right]^2 + 4n\left(\widehat{\beta}_n - \beta\right)\left(\widehat{\gamma}_n - \gamma\right). \end{aligned}$$

We can write the difference $\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} - (1-\lambda_-)$ as

$$\frac{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 - (\alpha-\delta)^2 + 4\left(\widehat{\beta}_n - \beta\right)\widehat{\gamma}_n + 4\left(\widehat{\gamma}_n - \gamma\right)\beta}{\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + (1-\lambda_-)},$$

while expanding $\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 - (\alpha-\delta)^2$ yields

$$\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 - (\alpha-\delta)^2 = \left[(\widehat{\alpha}_n - \alpha) - \left(\widehat{\delta}_n - \delta\right)\right]\left[(\widehat{\alpha}_n + \alpha) - \left(\widehat{\delta}_n + \delta\right)\right].$$

Hence $n(c_n - d_n) \xrightarrow{\mathcal{D}} 0$ as $n \to \infty$ will follow from $f_n \xrightarrow{\mathcal{D}} 0$ as $n \to \infty$,

where

$$
\begin{aligned}
f_n :=& n\left(\widehat{\alpha}_n-\alpha\right)^2\left[\sqrt{\left(\widehat{\alpha}_n-\widehat{\delta}_n\right)^2+4\widehat{\beta}_n\widehat{\gamma}_n}+1-\lambda_-+\widehat{\alpha}_n+\alpha-\widehat{\delta}_n-\delta\right]\\
&+4n\left(\widehat{\alpha}_n-\alpha\right)\left(\widehat{\beta}_n-\beta\right)\widehat{\gamma}_n+4n\left(\widehat{\alpha}_n-\alpha\right)\left(\widehat{\gamma}_n-\gamma\right)\beta\\
&-2n\left(\widehat{\alpha}_n-\alpha\right)\left(\delta_n-\delta\right)\left[\sqrt{\left(\widehat{\alpha}_n-\widehat{\delta}_n\right)^2+4\widehat{\beta}_n\widehat{\gamma}_n}+1-\lambda_-\right]\\
&+4n\left(\widehat{\beta}_n-\beta\right)\left(\widehat{\gamma}_n-\gamma\right)\left[\sqrt{\left(\widehat{\alpha}_n-\widehat{\delta}_n\right)^2+4\widehat{\beta}_n\widehat{\gamma}_n}+1-\lambda_-\right]\\
&+4n\left(\widehat{\beta}_n-\beta\right)\left(\delta_n-\delta\right)\widehat{\gamma}_n+4n\left(\widehat{\gamma}_n-\gamma\right)\left(\delta_n-\delta\right)\beta\\
&+n\left(\delta_n-\delta\right)^2\left[\sqrt{\left(\widehat{\alpha}_n-\widehat{\delta}_n\right)^2+4\widehat{\beta}_n\widehat{\gamma}_n}+1-\lambda_--\widehat{\alpha}_n-\alpha+\widehat{\delta}_n+\delta\right]
\end{aligned}
$$

as $n\to\infty$. By (69) we have

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

as $n\to\infty$. Consequently,

$$
n^{1/2}\begin{bmatrix}\widehat{\alpha}_n-\alpha\\ \widehat{\beta}_n-\beta\\ \widehat{\gamma}_n-\gamma\\ \widehat{\delta}_n-\delta\end{bmatrix} \xrightarrow{\mathcal{D}} \frac{\mathcal{C}}{\beta+1-\alpha}\begin{bmatrix}-(1-\alpha)\boldsymbol{e}_1^\top\boldsymbol{\mathcal{I}}\\ \beta\boldsymbol{e}_1^\top\boldsymbol{\mathcal{I}}\\ -(1-\alpha)\boldsymbol{e}_2^\top\boldsymbol{\mathcal{I}}\\ \beta\boldsymbol{e}_2^\top\boldsymbol{\mathcal{I}}\end{bmatrix}
$$

as $n\to\infty$, with

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

By the continuous mapping theorem,

$$n \begin{bmatrix} (\widehat{\alpha}_n - \alpha)^2 \\ (\widehat{\alpha}_n - \alpha)(\widehat{\beta}_n - \beta) \\ (\widehat{\alpha}_n - \alpha)(\widehat{\gamma}_n - \gamma) \\ (\widehat{\alpha}_n - \alpha)(\widehat{\delta}_n - \delta) \\ (\widehat{\beta}_n - \beta)(\widehat{\gamma}_n - \gamma) \\ (\widehat{\beta}_n - \beta)(\widehat{\delta}_n - \delta) \\ (\widehat{\gamma}_n - \gamma)(\widehat{\delta}_n - \delta) \\ (\widehat{\delta}_n - \delta)^2 \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \frac{\mathcal{C}^2}{(\beta+1-\alpha)^2} \begin{bmatrix} (1-\alpha)^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_1^\top \boldsymbol{\mathcal{I}} \\ -(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_1^\top \boldsymbol{\mathcal{I}} \\ (1-\alpha)^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \\ -(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \\ -(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \\ \beta^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \\ -(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_2 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \\ \beta^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_2 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \end{bmatrix}$$

as $n \to \infty$, and using the continuous mapping theorem, Slutsky's lemma and $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{m}_{\boldsymbol{\xi}}$ as $n \to \infty$, we get

$$\begin{aligned} f_n \stackrel{\mathcal{D}}{\longrightarrow} & \, 2(1-\alpha)^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_1^\top \boldsymbol{\mathcal{I}} \big[1 - \lambda_- + \alpha - \delta\big] - 4(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_1^\top \boldsymbol{\mathcal{I}} \gamma \\ & + 4(1-\alpha)^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \beta + 4(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \, (1 - \lambda_-) \\ & - 8(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \, (1 - \lambda_-) + 4\beta^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_1 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \gamma \\ & - 4(1-\alpha)\beta \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_2 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \beta + 2\beta^2 \boldsymbol{\mathcal{I}}^\top \boldsymbol{e}_2 \boldsymbol{e}_2^\top \boldsymbol{\mathcal{I}} \big[1 - \lambda_- - \alpha + \delta\big] \\ = & \, \boldsymbol{\mathcal{I}}^\top \left(f^{(1)} \boldsymbol{e}_1 \boldsymbol{e}_1^\top + f^{(2)} \boldsymbol{e}_1 \boldsymbol{e}_2^\top + f^{(3)} \boldsymbol{e}_2 \boldsymbol{e}_2^\top \right) \boldsymbol{\mathcal{I}} = 0, \end{aligned}$$

since

$$f^{(1)} = 2(1-\alpha)^2 \big[1 - \lambda_- + \alpha - \delta\big] - 4(1-\alpha)\beta\gamma = 0,$$
$$f^{(2)} = 4(1-\alpha)^2 \beta + 4(1-\alpha)\beta \, (1 - \lambda_-) - 8(1-\alpha)\beta \, (1 - \lambda_-) + 4\beta^2 \gamma = 0,$$
$$f^{(3)} = -4(1-\alpha)\beta^2 + 2\beta^2 \big[1 - \lambda_- - \alpha + \delta\big] = 0.$$

Consequently, (76) will follow from

$$\frac{n d_n}{2\left(\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n \widehat{\gamma}_n} + (1 - \lambda_-)\right)} \stackrel{\mathcal{D}}{\longrightarrow} \frac{\int_0^1 \mathcal{Y}_t \, \mathrm{d}(\mathcal{Y}_t - t\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle)}{\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t}$$

as $n \to \infty$. We can write

$$\begin{aligned} d_n = & \, 4\,(1-\delta)\, \boldsymbol{e}_1^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \boldsymbol{e}_1 + 4\gamma \boldsymbol{e}_1^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \boldsymbol{e}_2 \\ & + 4\beta \boldsymbol{e}_2^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \boldsymbol{e}_1 + 4(1-\alpha) \boldsymbol{e}_2^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \boldsymbol{e}_2 \\ = & \, 4\boldsymbol{e}_1^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \big[(1-\delta)\, \boldsymbol{e}_1 + \gamma \boldsymbol{e}_2\big] \\ & + 4\boldsymbol{e}_2^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \big[\beta \boldsymbol{e}_1 + (1-\alpha) \boldsymbol{e}_2\big]. \end{aligned}$$

Noticing

$$\beta \boldsymbol{e}_1 + (1-\alpha)\boldsymbol{e}_2 = \begin{bmatrix} \beta \\ 1-\alpha \end{bmatrix} = (\beta + 1 - \alpha)\, \boldsymbol{u}_{\text{right}},$$

$$(1-\delta)\, \boldsymbol{e}_1 + \gamma \boldsymbol{e}_2 = \begin{bmatrix} 1-\delta \\ \gamma \end{bmatrix} = \frac{\gamma}{1-\alpha} \begin{bmatrix} \beta \\ 1-\alpha \end{bmatrix} = (\gamma + 1 - \delta)\, \boldsymbol{u}_{\text{right}},$$

we get

$$\begin{aligned} d_n &= 4\left[(\gamma + 1 - \delta)\, \boldsymbol{e}_1^\top + (\beta + 1 - \alpha)\, \boldsymbol{e}_2^\top\right] \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}\right) \boldsymbol{u}_{\text{right}} \\ &= 4\,(1-\lambda_-)\, \boldsymbol{u}_{\text{left}}^\top \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}\right) \boldsymbol{u}_{\text{right}}. \end{aligned} \tag{77}$$

We use again the asymptotic expansion (74) of $\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n$. We have

$$\boldsymbol{D}_{n,2} = n^{-9/2} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \boldsymbol{M}_k \Big(V_{k-1} U_{\ell-1}^2 - U_{k-1} U_{\ell-1} V_{\ell-1} \Big) \boldsymbol{v}_{\text{left}}^\top,$$

implying $\boldsymbol{D}_{n,2} \boldsymbol{u}_{\text{right}} = \boldsymbol{0}$ for all $n \in \mathbb{N}$. Therefore

$$n d_n = \frac{4\,(1-\lambda_-)\, \boldsymbol{u}_{\text{left}}^\top \left(\boldsymbol{D}_{n,3} + n^{-1/2} \boldsymbol{D}_{n,4}\right) \boldsymbol{u}_{\text{right}}}{n^{-5} \det(A_n)},$$

where

$$\begin{aligned} \boldsymbol{u}_{\text{left}}^\top \Big(\boldsymbol{D}_{n,3} + n^{-1/2} \boldsymbol{D}_{n,4}\Big) \boldsymbol{u}_{\text{right}} &\xrightarrow{\mathcal{D}} \boldsymbol{u}_{\text{left}}^\top \left(\boldsymbol{\mathcal{C}}_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3 \begin{bmatrix} -\boldsymbol{v}_{\text{left}}^\top \\ \boldsymbol{u}_{\text{left}}^\top \end{bmatrix} \right) \boldsymbol{u}_{\text{right}} \\ &= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t\, \boldsymbol{u}_{\text{left}}^\top \int_0^1 \mathcal{Y}_t \,\mathrm{d}\boldsymbol{\mathcal{M}}_t\, \boldsymbol{u}_{\text{left}}^\top \boldsymbol{u}_{\text{right}}. \end{aligned}$$

Putting it all together we have

$$\frac{4\,(1-\lambda_-)}{2\left(\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n \widehat{\gamma}_n} + (1-\lambda_-)\right)} \xrightarrow{\mathbb{P}} 1,$$

and

$$\begin{bmatrix} \boldsymbol{u}_{\text{left}}^\top \left(\boldsymbol{D}_{n,3} + n^{-1/2} \boldsymbol{D}_{n,4}\right) \boldsymbol{u}_{\text{right}} \\ n^{-5} \det(A_n) \end{bmatrix} \xrightarrow{\mathcal{D}} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\text{left}}, \boldsymbol{v}_{\text{left}} \rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \begin{bmatrix} \boldsymbol{u}_{\text{left}}^\top \int_0^1 \mathcal{Y}_t \,\mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \end{bmatrix}.$$

Thus we can conclude (76). □

4.3 The general process with unknown immigration mean

In this section we offer a small extension on the previous results. So far we always assumed the knowledge of the immigration mean vector, in what follows we treat this parameter as unknown and prove results on the existence and limit distribution of a joint estimator of the two means. Unfortunately this does not offer a significant improvement as the results on the offspring mean estimator are unsatisfactory.

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

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

with respect to $\boldsymbol{m}_{\boldsymbol{\xi}}$, $\boldsymbol{m}_{\varepsilon}$ over $\mathbb{R}^{2\times 2} \times \mathbb{R}^2$.

Lemma 4.9. *The joint CLS estimator of* $\boldsymbol{m}_{\boldsymbol{\xi}}$ *and* $\boldsymbol{m}_{\varepsilon}$ *has the form*

$$\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} = \boldsymbol{B}_n \boldsymbol{A}_n^{-1}$$
$$\widehat{\boldsymbol{m}_{\varepsilon}}^{(n)} = \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{X}_k - \widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{X}_{k-1},$$

on the set $\Omega_n := \{\omega \in \Omega \colon \det(\boldsymbol{A}_n) > 0\}$, *where*

$$\boldsymbol{A}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{X}_{k-1} \boldsymbol{X}_{k-1}^{\top} - \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{X}_{k-1} \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^{\top},$$
$$\boldsymbol{B}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{X}_{k} \boldsymbol{X}_{k-1}^{\top} - \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{X}_{k} \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^{\top}.$$

Proof. The proof uses the same ideas as in Lemma 4.5, but for the sake of completeness we present it here. Define the function $Q_n : \left(\mathbb{R}^2\right)^n \times \mathbb{R}^{2\times 2} \times \mathbb{R}^2 \to \mathbb{R}$ as

$$Q_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \boldsymbol{m}, \boldsymbol{u}) = \sum_{k=1}^{n} \| \boldsymbol{x}_k - \boldsymbol{m} \boldsymbol{x}_{k-1} - \boldsymbol{u} \|^2$$

with the convention that $\boldsymbol{x}_0 = \boldsymbol{0}$. A joint CLS estimator of $\boldsymbol{m}_{\boldsymbol{\xi}}$ and $\boldsymbol{m}_{\varepsilon}$ is a measurable function $F_n : \left(\mathbb{R}^2\right)^n \to \mathbb{R}^{2\times 2} \times \mathbb{R}^2$ such that

$$Q_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n, F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)) = \inf_{\boldsymbol{m} \in \mathbb{R}^{2\times 2},\ \boldsymbol{u} \in \mathbb{R}^2} Q_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n, \boldsymbol{m}, \boldsymbol{u})$$

for all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$. We need to show that on the set

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

we have

$$F_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) = \begin{bmatrix} F_{n,\boldsymbol{m}_\xi}\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) \\ F_{n,\boldsymbol{m}_\varepsilon}\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) \end{bmatrix},$$

where

$$F_{n,\boldsymbol{m}_\xi}\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) = \boldsymbol{B}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) \boldsymbol{A}_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right)^{-1}$$
$$F_{n,\boldsymbol{m}_\varepsilon}\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right) = \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{x}_k - F_{n,\boldsymbol{m}_\xi}\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right)\frac{1}{n}\sum_{k=1}^{n} \boldsymbol{x}_{k-1}.$$

Fix $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and find the critical points (where all partial derivatives vanishes). Let

$$\boldsymbol{m} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

then

$$Q_n\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{m}, \boldsymbol{u}\right) = \sum_{k=1}^{n} \left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - u_1\right)^2 + \sum_{k=1}^{n} \left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - u_2\right)^2.$$

To find the critical points we have to solve

$$\begin{cases} \frac{\partial}{\partial a} Q_n = 0 \\ \frac{\partial}{\partial b} Q_n = 0 \\ \frac{\partial}{\partial c} Q_n = 0 \\ \frac{\partial}{\partial d} Q_n = 0 \\ \frac{\partial}{\partial u_1} Q_n = 0 \\ \frac{\partial}{\partial u_2} Q_n = 0 \end{cases} \implies \begin{cases} -2\sum_{k=1}^{n}\left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - u_1\right) x_{k-1,1} = 0 \\ -2\sum_{k=1}^{n}\left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - u_1\right) x_{k-1,2} = 0 \\ -2\sum_{k=1}^{n}\left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - u_2\right) x_{k-1,1} = 0 \\ -2\sum_{k=1}^{n}\left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - u_2\right) x_{k-1,2} = 0 \\ -2\sum_{k=1}^{n}\left(x_{k,1} - a x_{k-1,1} - b x_{k-1,2} - u_1\right) = 0 \\ -2\sum_{k=1}^{n}\left(x_{k,2} - c x_{k-1,1} - d x_{k-1,2} - u_2\right) = 0 \end{cases}$$

Rearranging gives us the equations

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \sum_{k=1}^{n} \boldsymbol{x}_{k-1}\boldsymbol{x}_{k-1}^\top + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sum_{k=1}^{n} \boldsymbol{x}_{k-1}^\top = \sum_{k=1}^{n} \boldsymbol{x}_k \boldsymbol{x}_{k-1}^\top,$$

$$n \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \sum_{k=1}^{n} \boldsymbol{x}_{k-1} = \sum_{k=1}^{n} \boldsymbol{x}_k.$$

So $F_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is the only critical point if $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in D(F_n)$. However we still have to prove that it is in fact a minimum, we will use the second order derivatives. The Hessian matrix of Q_n, $\boldsymbol{H}_n$ is

$$2\sum_{k=1}^{n} \begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} & 0 & 0 & x_{k-1,1} & 0 \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 & 0 & 0 & x_{k-1,2} & 0 \\ 0 & 0 & x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} & 0 & x_{k-1,1} \\ 0 & 0 & x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 & 0 & x_{k-1,2} \\ x_{k-1,1} & x_{k-1,2} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{k-1,1} & x_{k-1,2} & 0 & 1 \end{bmatrix},$$

as we can see it does not depend on the parameters a, b, c, d, u_1, u_2. Forgetting about the positive constant multiplier 2, and exchanging the third and fifth rows and the third and fifth columns (this does not change the positive definite property of a matrix) we get a block diagonal matrix

$$\sum_{k=1}^{n} \begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} & x_{k-1,1} & 0 & 0 & 0 \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 & x_{k-1,2} & 0 & 0 & 0 \\ x_{k-1,1} & x_{k-1,2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} & x_{k-1,1} \\ 0 & 0 & 0 & x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 & x_{k-1,2} \\ 0 & 0 & 0 & x_{k-1,1} & x_{k-1,2} & 1 \end{bmatrix}.$$

To show that this matrix (and consequently $\boldsymbol{H}_n$) is positive definite, it is sufficient to show that the 3×3 matrix in the upper right corner

$$\boldsymbol{H}_n^* := \sum_{k=1}^{n} \begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} & x_{k-1,1} \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 & x_{k-1,2} \\ x_{k-1,1} & x_{k-1,2} & 1 \end{bmatrix}$$

is positive definite. We are going to do this by showing, that its leading principal minors are all positive, that is

$$0 < \det(\boldsymbol{H}_n^*),$$

$$0 < \det\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 \end{bmatrix}\right),$$

$$0 < \sum_{k=1}^{n} x_{k-1,1}^2,$$

if $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in D(F_n)$. Surprisingly

$$\begin{aligned}
\det(\boldsymbol{H}_n^*) &= n\det(\boldsymbol{A}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) \\
&= n\sum_{k=1}^{n} x_{k-1,1}^2 \sum_{k=1}^{n} x_{k-1,2}^2 - n\left(\sum_{k=1}^{n} x_{k-1,1}x_{k-1,2}\right)^2 \\
&\quad + 2\sum_{k=1}^{n} x_{k-1,1}x_{k-1,2}\sum_{k=1}^{n} x_{k-1,1}\sum_{k=1}^{n} x_{k-1,2} \\
&\quad - \sum_{k=1}^{n} x_{k-1,1}^2\left(\sum_{k=1}^{n} x_{k-1,2}\right)^2 - \sum_{k=1}^{n} x_{k-1,2}^2\left(\sum_{k=1}^{n} x_{k-1,1}\right)^2 \\
&= n\det\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 \end{bmatrix}\right) \\
&\quad - \sum_{k=1}^{n}\left(x_{k-1,1}\sum_{\ell=1}^{n} x_{\ell-1,2} + x_{k-1,2}\sum_{\ell=1}^{n} x_{\ell-1,1}\right)^2.
\end{aligned}$$

Therefore on the set $D(F_n)$ we have

$$\begin{aligned}
0 < \det(\boldsymbol{A}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) &= \frac{1}{n}\det(\boldsymbol{H}_n^*) \\
&< \det\left(\sum_{k=1}^{n}\begin{bmatrix} x_{k-1,1}^2 & x_{k-1,1}x_{k-1,2} \\ x_{k-1,1}x_{k-1,2} & x_{k-1,2}^2 \end{bmatrix}\right) \\
&= \sum_{k=1}^{n} x_{k-1,1}^2 \sum_{k=1}^{n} x_{k-1,2}^2 - \left(\sum_{k=1}^{n} x_{k-1,1}x_{k-1,2}\right)^2 \\
&< \sum_{k=1}^{n} x_{k-1,1}^2 \sum_{k=1}^{n} x_{k-1,2}^2.
\end{aligned}$$

Consequently $\sum_{k=1}^{n} x_{k-1,1}^2 > 0$, the matrices $\boldsymbol{H}_n^*$ and $\boldsymbol{H}_n$ are positive definite, and $F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is the minimum of Q_n. This proves the formula for the joint estimator. □

We have the formula for the estimator, but in order to prove a limit theorem we have to express the difference from the real parameter values in a form that we can work with.

Corollary 4.10. *The difference of the CLS estimates and the real parameter values can be expressed as*

$$\begin{aligned}
\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} &= \boldsymbol{C}_n \boldsymbol{A}_n^{-1}, \\
\widehat{\boldsymbol{m}_{\varepsilon}}^{(n)} - \boldsymbol{m}_{\varepsilon} &= \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{M}_k - \left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}\right)\frac{1}{n}\sum_{k=1}^{n} \boldsymbol{X}_{k-1},
\end{aligned} \tag{78}$$

where

$$\boldsymbol{C}_n(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{k=1}^{n} \boldsymbol{M}_k \boldsymbol{X}_{k-1}^\top - \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{M}_k \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^\top.$$

Proof. We can find the formula for the difference with a straightforward calculation

$$\begin{aligned}
\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} &= \boldsymbol{B}_n \boldsymbol{A}_n^{-1} - \boldsymbol{m}_{\boldsymbol{\xi}} = \left(\boldsymbol{B}_n - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{A}_n\right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \boldsymbol{X}_k \boldsymbol{X}_{k-1}^\top - \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{X}_k \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^\top - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{A}_n\right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1}\right) \boldsymbol{X}_{k-1}^\top \right. \\
&\qquad \left. - \frac{1}{n}\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1}\right) \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^\top \right) \boldsymbol{A}_n^{-1} \\
&= \left(\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\varepsilon}\right) \boldsymbol{X}_{k-1}^\top \right. \\
&\qquad \left. - \frac{1}{n}\sum_{k=1}^{n} \left(\boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\varepsilon}\right) \sum_{k=1}^{n} \boldsymbol{X}_{k-1}^\top \right) \boldsymbol{A}_n^{-1} \\
&= \boldsymbol{C}_n \boldsymbol{A}_n^{1}.
\end{aligned}$$

□

In the critical case, by (7) and the continuous mapping theorem, one can derive

$$n^{-3}\boldsymbol{A}_n \stackrel{\mathcal{D}}{\longrightarrow} \left(\int_0^1 \mathcal{Y}_t^2\,dt - \left(\int_0^1 \mathcal{Y}_t\,dt\right)^2\right)\boldsymbol{u}_{\mathrm{right}}\boldsymbol{u}_{\mathrm{right}}^\top =: \boldsymbol{\mathcal{A}}$$

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

$$\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} = \boldsymbol{C}_n\boldsymbol{A}_n^{-1} = \frac{1}{\det(\boldsymbol{A}_n)}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n, \qquad n \in \mathbb{N}, \tag{79}$$

on the set Ω_n, where $\widetilde{\boldsymbol{A}}_n$ denotes the adjugate of $\boldsymbol{A}_n$ (i.e., the matrix of cofactors) given by

$$\widetilde{\boldsymbol{A}}_n = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\left(\sum_{k=1}^n \boldsymbol{X}_{k-1}\boldsymbol{X}_{k-1}^\top - \sum_{k=1}^n \boldsymbol{X}_{k-1}\sum_{k=1}^n \boldsymbol{X}_{k-1}^\top\right)\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{80}$$

We can find the limit for the difference $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}$ by describing the asymptotic behaviour of the sequence $(\det(\boldsymbol{A}_n), \boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n)_{n\in\mathbb{N}}$.

Theorem. 4.11. *Let $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type Galton–Watson process with immigration that satisfies conditions* (CPR), (ZS) *and* (M) *with $\ell = 8$. If the process also satisfies* (ND), *then the probability of the existence of the estimators $\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)}$ and $\widehat{\boldsymbol{m}_{\varepsilon}}^{(n)}$ tends to 1 as $n \to \infty$, and further*

$$n^{1/2}(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}) \stackrel{\mathcal{D}}{\longrightarrow} \frac{(1-\lambda_-^2)^{1/2}}{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}} \frac{\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t}{\int_0^1 \mathcal{Y}_t\,\mathrm{d}t}\boldsymbol{v}_{\mathrm{left}}^\top \tag{81}$$

$$\widehat{\boldsymbol{m}_{\varepsilon}}^{(n)} - \boldsymbol{m}_{\varepsilon} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{M}}_1 \tag{82}$$

as $n \to \infty$, with $\mathcal{Y}_t := \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\varepsilon}\rangle$, $t \in \mathbb{R}_+$, where $(\boldsymbol{\mathcal{M}}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the SDE

$$\begin{aligned} \mathrm{d}\boldsymbol{\mathcal{M}}_t &= (\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\varepsilon}\rangle^+)^{1/2}\,\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \\ \boldsymbol{\mathcal{M}}_0 &= \boldsymbol{0}, \end{aligned}$$

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.

Proof. We start by finding the nonzero limit of $\det(\boldsymbol{A}_n)$. This will allow us to prove the asymptotic existence of the estimators. By the decomposition (20) and Lemma 2.2 we get

$$\begin{aligned}
\det(\boldsymbol{A}_n) &= \det\left(\sum_{k=1}^n \boldsymbol{X}_{k-1}\boldsymbol{X}_{k-1}^\top - \frac{1}{n}\sum_{k=1}^n \boldsymbol{X}_{k-1}\sum_{k=1}^n \boldsymbol{X}_{k-1}^\top\right) \\
&= \det\left(\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}\begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top - \frac{1}{n}\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top\right) \det\left(\begin{bmatrix} \boldsymbol{u}_{\mathrm{right}} & \boldsymbol{v}_{\mathrm{right}} \end{bmatrix}\right)^2 \\
&= \det\left(\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}\begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top - \frac{1}{n}\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top\right)
\end{aligned}$$

for all $n \in \mathbb{N}$. By Corollary 3.15, Lemmas 3.8 and 3.9 and the continuous mapping theorem we have

$$n^{-5}\det(\boldsymbol{A}_n) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2\right) \tag{83}$$

as $n \to \infty$. The process satisfies (ZS), therefore $\boldsymbol{m}_{\boldsymbol{\varepsilon}} \neq \boldsymbol{0}$, consequently by the SDE (61), we have

$$\mathbb{P}(\mathcal{Y}_t = 0 \ \text{ for all } \ t \in [0,1]) = 0.$$

This implies with the help of the Cauchy-Schwarz inequality, that

$$\mathbb{P}\left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2\right) > 0\right) = 1.$$

Consequently, the distribution function of

$$\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2\right)$$

is continuous at 0. Note that

$$\mathbb{P}(\Omega_n) = \mathbb{P}\left(\det(\boldsymbol{A}_n) > 0\right) = \mathbb{P}\left(n^{-5}\det(\boldsymbol{A}_n) > 0\right).$$

If the process satisfies (ND), then $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle > 0$, and by (83),

$$\begin{aligned}
\mathbb{P}(\Omega_n) &\to \mathbb{P}\left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2\right) > 0\right) \\
&= \mathbb{P}\left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2\right) > 0\right) = 1
\end{aligned}$$

as $n \to \infty$. This proves the asymptotic existence of the estimators.

Next we turn to prove convergence (81). We do this by finding stochastic expansions for the product $\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n$. We will use Corollary 3.15 again, so in fact we are proving a joint convergence of the sequence $(\det(\boldsymbol{A}_n), \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n)_{n\in\mathbb{N}}$. Using (20), we can write

$$\boldsymbol{C}_n = \left(\sum_{k=1}^n \boldsymbol{M}_k \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^\top - \frac{1}{n} \sum_{k=1}^n \boldsymbol{M}_k \sum_{k=1}^n \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^\top \right) \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}}^\top \\ \boldsymbol{v}_{\mathrm{right}}^\top \end{bmatrix},$$

$$\widetilde{\boldsymbol{A}}_n = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}}^\top \\ \boldsymbol{v}_{\mathrm{right}}^\top \end{bmatrix}^\top \left(\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top - \frac{1}{n} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top \right) \times \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}}^\top \\ \boldsymbol{v}_{\mathrm{right}}^\top \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Similarly to Lemmas 2.2 and 2.3 one can show

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

therefore

$$\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n = \left(\sum_{k=1}^n \boldsymbol{M}_k \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^\top - \frac{1}{n} \sum_{k=1}^n \boldsymbol{M}_k \sum_{k=1}^n \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}^\top \right) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \times \left(\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top - \frac{1}{n} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top \right) \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}.$$

Corollary 3.15 and Lemmas 3.8 and 3.9 implies stochastic expansions

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

$$\sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top - \frac{1}{n} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix} \sum_{k=1}^n \begin{bmatrix} U_k \\ V_k \end{bmatrix}^\top = n^3 \boldsymbol{A}_{n,1} + n^{5/2} \boldsymbol{A}_{n,2} + n^2 \boldsymbol{A}_{n,3},$$

where

$$\boldsymbol{C}_{n,1} := n^{-2}\left(\sum_{k=1}^{n}\boldsymbol{M}_k U_{k-1} - \frac{1}{n}\sum_{k=1}^{n}\boldsymbol{M}_k\sum_{k=1}^{n}U_{k-1}\right)\boldsymbol{e}_1^\top$$
$$\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_1 := \left(\int_0^1 \mathcal{Y}_t\,\mathrm{d}\boldsymbol{\mathcal{M}}_t - \boldsymbol{\mathcal{M}}_1\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\right)\boldsymbol{e}_1^\top,$$
$$\boldsymbol{C}_{n,2} := n^{-3/2}\left(\sum_{k=1}^{n}\boldsymbol{M}_k V_{k-1} - \frac{1}{n}\sum_{k=1}^{n}\boldsymbol{M}_k\sum_{k=1}^{n}V_{k-1}\right)\boldsymbol{e}_2^\top$$
$$\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 := \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}}{(1-\lambda_-^2)^{1/2}}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t\,\boldsymbol{e}_2^\top,$$

and

$$\boldsymbol{A}_{n,1} := n^{-3}\left(\sum_{k=1}^{n}U_{k-1}^2 - \frac{1}{n}\left(\sum_{k=1}^{n}U_{k-1}\right)^2\right)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$
$$\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{A}}_1 := \left(\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\right)^2\right)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix},$$
$$\boldsymbol{A}_{n,2} := n^{-5/2}\left(\sum_{k=1}^{n}U_{k-1}V_{k-1} - \frac{1}{n}\sum_{k=1}^{n}U_{k-1}\sum_{k=1}^{n}V_{k-1}\right)\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \xrightarrow{\mathcal{D}} \boldsymbol{0}$$
$$\boldsymbol{A}_{n,3} := n^{-2}\left(\sum_{k=1}^{n}V_{k-1}^2 - \frac{1}{n}\left(\sum_{k=1}^{n}V_{k-1}\right)^2\right)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$$
$$\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{A}}_3 := \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}.$$

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

$$\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n = (n^5\boldsymbol{D}_{n,1} + n^{9/2}\boldsymbol{D}_{n,2} + n^4\boldsymbol{D}_{n,3} + n^{7/2}\boldsymbol{D}_{n,4})\begin{bmatrix}-\boldsymbol{v}_{\mathrm{left}}^\top\\ \boldsymbol{u}_{\mathrm{left}}^\top\end{bmatrix}, \tag{84}$$

where

$$\begin{aligned}
\boldsymbol{D}_{n,1} :&= \boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \\
&= n^{-5} \left(\sum_{k=1}^{n} \boldsymbol{M}_k U_{k-1} - \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{M}_k \sum_{k=1}^{n} U_{k-1} \right) \\
&\quad \times \left(\sum_{k=1}^{n} U_{k-1}^2 - \frac{1}{n} \left(\sum_{k=1}^{n} U_{k-1} \right)^2 \right) \boldsymbol{e}_1^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{0}
\end{aligned}$$

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

$$\begin{aligned}
\boldsymbol{D}_{n,2} &:= \boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} + \boldsymbol{C}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1, \\
\boldsymbol{D}_{n,3} &:= \boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,3} + \boldsymbol{C}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3, \\
\boldsymbol{D}_{n,4} &:= \boldsymbol{C}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,3} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3
\end{aligned}$$

as $n \to \infty$. Finally putting it all together we get

$$n^{-9/2} \boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}$$

as $n \to \infty$, where

$$\begin{aligned}
&\boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix} \\
&= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \left(\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \, \mathrm{d}t \right)^2 \right) \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \\
&\quad \times \boldsymbol{e}_2^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix} \\
&= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \left(\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \, \mathrm{d}t \right)^2 \right) \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \, \boldsymbol{v}_{\mathrm{left}}^\top.
\end{aligned}$$

We have shown the joint convergence

$$\begin{bmatrix} n^{-5}\det(\boldsymbol{A}_n) \\ n^{-9/2}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2}\left(\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\right)^2\right)\int_0^1 \mathcal{Y}_t\,\mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}}{(1-\lambda_-^2)^{1/2}}\left(\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t\,\mathrm{d}t\right)^2\right)\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t\,\boldsymbol{v}_{\mathrm{left}}^\top \end{bmatrix}$$

as $n \to \infty$. Using the continuous mapping theorem on this result gives us the desired convergence

$$n^{1/2}(\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}) = \frac{n^{-9/2}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n}{n^{-5}\det(\boldsymbol{A}_n)} \xrightarrow{\mathcal{D}} \frac{(1-\lambda_-^2)^{1/2}}{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{v}_{\mathrm{left}},\boldsymbol{v}_{\mathrm{left}}\rangle^{1/2}} \frac{\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t}{\int_0^1 \mathcal{Y}_t\,\mathrm{d}t}\boldsymbol{v}_{\mathrm{left}}^\top.$$

To prove convergence (82) we use the same method. By Corollary 3.15 we have

$$\frac{1}{n}\sum_{k=1}^n \boldsymbol{M}_k \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{M}}_1.$$

Using (78), to finish the proof we have to show

$$\left(\widehat{\boldsymbol{m}}_{\boldsymbol{\xi}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}}\right)\frac{1}{n}\sum_{k=1}^n \boldsymbol{X}_{k-1} = \frac{1}{\det(\boldsymbol{A}_n)}\boldsymbol{C}_n\widetilde{\boldsymbol{A}}_n\frac{1}{n}\sum_{k=1}^n \boldsymbol{X}_{k-1} \xrightarrow{\mathcal{D}} \boldsymbol{0}.$$

Using (20) we can write

$$\frac{1}{n}\sum_{k=1}^n \boldsymbol{X}_{k-1} = \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}}^\top \\ \boldsymbol{v}_{\mathrm{right}}^\top \end{bmatrix}^\top \frac{1}{n}\sum_{k=1}^n \begin{bmatrix} U_{k-1} \\ V_{k-1} \end{bmatrix}$$

Corollary 3.15 and (84) implies stochastic expansions

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

$$\frac{1}{n}\sum_{k=1}^n \boldsymbol{X}_{k-1} = \begin{bmatrix} \boldsymbol{u}_{\mathrm{right}}^\top \\ \boldsymbol{v}_{\mathrm{right}}^\top \end{bmatrix}^\top \left(n\boldsymbol{F}_{n,1} + n^{1/2}\boldsymbol{F}_{n,2}\right),$$

where

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

$$\boldsymbol{F}_{n,2} = n^{-3/2}\sum_{k=1}^n V_{k-1}\boldsymbol{e}_2 \xrightarrow{\mathcal{D}} \boldsymbol{0}$$

jointly as $n \to \infty$. By Lemmas 2.1 and 2.3 we have

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

Consequently we obtain a stochastic expansion

$$\boldsymbol{C}_n \widetilde{\boldsymbol{A}}_n \frac{1}{n} \sum_{k=1}^n \boldsymbol{X}_{k-1} = n^{11/2} \boldsymbol{G}_{n,1} + n^5 \boldsymbol{G}_{n,2} + n^{9/2} \boldsymbol{G}_{n,3} + n^4 \boldsymbol{G}_{n,4}$$

where

$$\begin{aligned} \boldsymbol{G}_{n,1} &= \boldsymbol{D}_{n,2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} \\ &= \left(\boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} + \boldsymbol{C}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} = \boldsymbol{0}, \end{aligned}$$

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

$$\begin{aligned} &\boldsymbol{C}_{n,1} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} \\ &\quad = n^{-13/2} \sum_{k=1}^n U_{k-1} \left(\sum_{k=1}^n U_{k-1} V_{k-1} - \frac{1}{n} \sum_{k=1}^n U_{k-1} \sum_{k=1}^n V_{k-1} \right) \\ &\qquad \times \left(\sum_{k=1}^n \boldsymbol{M}_k U_{k-1} - \frac{1}{n} \sum_{k=1}^n \boldsymbol{M}_k \sum_{k=1}^n U_{k-1} \right) \\ &\qquad \times \boldsymbol{e}_1^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{e}_1 = \boldsymbol{0}, \end{aligned}$$

and

$$\begin{aligned} &\boldsymbol{C}_{n,2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{A}_{n,1} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} \\ &\quad = n^{-13/2} \sum_{k=1}^n U_{k-1} \left(\sum_{k=1}^n U_{k-1}^2 - \frac{1}{n} \left(\sum_{k=1}^n U_{k-1} \right)^2 \right) \\ &\qquad \times \left(\sum_{k=1}^n \boldsymbol{M}_k V_{k-1} - \frac{1}{n} \sum_{k=1}^n \boldsymbol{M}_k \sum_{k=1}^n V_{k-1} \right) \\ &\qquad \times \boldsymbol{e}_2^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{e}_1 = \boldsymbol{0}. \end{aligned}$$

The other terms in the expansion are

$$\begin{aligned}
\boldsymbol{G}_{n,2} &= \boldsymbol{D}_{n,2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,2} + \boldsymbol{D}_{n,3} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} \\
&\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \, \boldsymbol{e}_1 \\
\boldsymbol{G}_{n,3} &= \boldsymbol{D}_{n,3} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,2} + \boldsymbol{D}_{n,4} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,1} \\
&\xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{C}}_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \, \boldsymbol{e}_1 \\
\boldsymbol{G}_{n,2} &= \boldsymbol{D}_{n,4} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{F}_{n,2} \xrightarrow{\mathcal{D}} \boldsymbol{0}.
\end{aligned}$$

We have

$$\begin{aligned}
&\boldsymbol{\mathcal{C}}_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_1 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \, \boldsymbol{e}_1 \\
&= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}} \rangle^{1/2}}{(1-\lambda_-^2)^{1/2}} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \left(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \right)^2 \right) \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \\
&\quad \times \boldsymbol{e}_2^\top \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \boldsymbol{e}_1 = \boldsymbol{0}
\end{aligned}$$

Finally

$$\left(\widehat{\boldsymbol{m}_{\boldsymbol{\xi}}}^{(n)} - \boldsymbol{m}_{\boldsymbol{\xi}} \right) \frac{1}{n} \sum_{k=1}^n \boldsymbol{X}_{k-1} = \frac{\boldsymbol{G}_{n,2} + n^{-1/2} \boldsymbol{G}_{n,3} + n^{-1} \boldsymbol{G}_{n,4}}{n^{-5} \det(\boldsymbol{A}_n)} \xrightarrow{\mathcal{D}} \boldsymbol{0}$$

proves the statement. □

Theorem. 4.12. *Let* $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ *be a 2-type Galton–Watson process with immigration satisfying conditions* (CPR), (ZS) *and* (M) *with* $\ell = 8$. *If the process satisfies* (ND) *as well, then the probability of the existence of the estimator* $\widehat{\varrho}_n$ *tends to 1 as* $n \to \infty$, *and further*

$$n(\widehat{\varrho}_n - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{Y}_t \,\mathrm{d}(\mathcal{Y}_t - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_\varepsilon \rangle t) - (\mathcal{Y}_1 - \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_\varepsilon \rangle) \int_0^1 \mathcal{Y}_t \,\mathrm{d}t}{\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \right)^2} \tag{85}$$

as $n \to \infty$.

Proof. We can follow the proof of Theorem 4.8 up to the point of expressing d_n in (77). Then by stochastic expansion (84) we have $\boldsymbol{D}_{n,2}\boldsymbol{u}_{\mathrm{right}} = 0$ for all $n \in \mathbb{N}$.

Therefore

$$nd_n = \frac{4\,(1-\lambda_-)\,\boldsymbol{u}_{\mathrm{left}}^\top \left(\boldsymbol{D}_{n,3} + n^{-1/2}\boldsymbol{D}_{n,4}\right)\boldsymbol{u}_{\mathrm{right}}}{n^{-5}\det(A_n)},$$

where

$$\boldsymbol{u}_{\mathrm{left}}^\top \left(\boldsymbol{D}_{n,3} + n^{-1/2}\boldsymbol{D}_{n,4}\right)\boldsymbol{u}_{\mathrm{right}} \xrightarrow{\mathcal{D}} \boldsymbol{u}_{\mathrm{left}}^\top \left(\boldsymbol{\mathcal{C}}_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\mathcal{A}}_3 \begin{bmatrix} -\boldsymbol{v}_{\mathrm{left}}^\top \\ \boldsymbol{u}_{\mathrm{left}}^\top \end{bmatrix}\right)\boldsymbol{u}_{\mathrm{right}}$$
$$= \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t\, \boldsymbol{u}_{\mathrm{left}}^\top \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}\boldsymbol{\mathcal{M}}_t - \boldsymbol{\mathcal{M}}_1 \int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right) \boldsymbol{u}_{\mathrm{left}}^\top \boldsymbol{u}_{\mathrm{right}}$$

Putting it all together we have

$$\frac{4\,(1-\lambda_-)}{2\left(\sqrt{\left(\widehat{\alpha}_n - \widehat{\delta}_n\right)^2 + 4\widehat{\beta}_n\widehat{\gamma}_n} + (1-\lambda_-)\right)} \xrightarrow{\mathbb{P}} 1,$$

and

$$\begin{bmatrix} \boldsymbol{u}_{\mathrm{left}}^\top \left(\boldsymbol{D}_{n,3} + n^{-1/2}\boldsymbol{D}_{n,4}\right)\boldsymbol{u}_{\mathrm{right}} \\ n^{-5}\det(A_n) \end{bmatrix}$$
$$\xrightarrow{\mathcal{D}} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{v}_{\mathrm{left}}, \boldsymbol{v}_{\mathrm{left}}\rangle}{1-\lambda_-^2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \begin{bmatrix} \boldsymbol{u}_{\mathrm{left}}^\top \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}\boldsymbol{\mathcal{M}}_t - \boldsymbol{\mathcal{M}}_1 \int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right) \\ \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)^2 \end{bmatrix}.$$

Using $\langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{\mathcal{M}}_t\rangle = \mathcal{Y}_t + \langle \boldsymbol{u}_{\mathrm{left}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle t$, we can conclude (85). □

5 A discussion of the results

In this section we discuss the results in Sections 3 and 4. We point out open questions, possible avenues of improvements and investigations.

5.1 A discussion of the toolkit

In Section 3 we developed a toolkit for studying the asymptotic properties of estimators of the offspring mean matrix in critical 2-type Galton–Watson processes with immigration. We introduced a decomposition of the process based on the eigenvalues and eigenvectors of the offspring mean matrix, then we investigated the asymptotic behaviour of these building blocks. It was a two step process, first we introduced zero limit theorems to understand which terms will be negligible in an expression, then we proved a joint limit theorem finding the non-zero limits of the building block. This limit was described in Corollary 3.15.

One way these results could be improved is by relaxing the moment conditions. While the question of estimating the offspring mean only requires the existence of the first moment we require the existence of the 8th moments. It is reasonable to expect that the moment condition (M) could be relaxed to $\ell = 4$. If someone sets out to achieve this, then there are two points in the proof that needs improving. The first is checking condition (ii) of Theorem 3.12, as expressed in formula (60), and then the multiple application of convergence (25) of Corollary 3.6. Out of these two, the latter seems more difficult to improve.

A possible direction of generalization could be to examine d-type Galton–Watson processes, where $d > 2$. This dissertation deals with the 2-type case inspired by the results available for the single-type case. The advantage of working in 2-dimensions is that we can solve quadratic equations therefore we can explicitly describe the eigenvalues and eigenvectors in terms of the elements of the offspring mean matrix. This advantage is lost in higher dimensions. One could possibly gain insight into the problem by first trying to solve cases with heavy restrictions on the structure of the offspring main matrix, this is what we did for the 2-type case[10]. To give an impression as to what would be needed for the general case we point to the proofs of Lemmas 2.1 and 2.3. The first proof only uses general properties of eigenvectors, while the next one requires us to write out the coordinates of the vectors in question, for the general case all proofs would need to be carried out in the same manner as the first one.

5.2 A discussion of the estimates

In Section 4 we used our toolkit to examine the asymptotic behaviour of the estimators for the offspring mean matrix under three set of assumptions. All of these estimators were obtained by the conditional least squares method, however we'd like to note, that the toolkit is not restricted to these types of estimators. Any estimator obtained by any heuristics can be examined this way as long as it can be written as a continuous function of our building blocks.

Subsection 4.1 reproduces our first published results [10] on this subject. It was proved using our toolkit, however in a more obscure form, not as streamlined and clearly structured as it is presented here. Later we treated the general case [16], those results are reproduced in Subsection 4.2. Finally Subsection 4.3 offers a small extension in treating the immigration mean as unknown parameter.

In all three cases the estimates for the offspring mean matrix and the criticality parameter found to be weakly consistent and their limit behaviour is described with $n^{1/2}$ and n scaling respectively. A next logical step would be to use these limit distributions as a basis for constructing test, however that is not feasible. The problem is that in all cases the limiting distributions depend on the very parameters we are estimating in an intricate way, namely the process $\mathcal{Y}_t, t \in \mathbb{R}_+$ introduced in Remark 3.14 that appears in our limits depend on the eigenvectors of the offspring mean matrix, see SDE (61). At this time we found no way to work around this problem. Nevertheless these are the only consistent estimators available in the literature of critical 2-type Galton–Watson processes.

We also note that Subsection 4.3 contains a joint estimator of both the offspring and the immigration means. Unfortunately the estimator for the immigration mean requires no scaling for a limit, therefore even weak consistency cannot be established. The problem of estimating the immigration mean in branching processes is a great deal harder than estimating the offspring mean, we didn't expect strong results in this area.

A Summary

In this section we summarize the results. For the sake of conserving space we keep formulas and references to other parts of the thesis to a minimum.

As the title suggests the goal of this thesis is to estimate the offspring mean matrix in critical 2-type Galton–Watson processes with immigration. After some preliminaries in Section 2 this is achieved in two large steps. First, in Section 3 we establish a toolkit for asymptotic study of estimators, then in Section 4 we introduce estimators and use those tools we developed to examine their asymptotic properties. The core of the thesis ends with Section 5 a discussion of the results as well as some open questions. Below is a short summary of the ideas and key insights that went into these results and some description of what was achieved.

We begin with introducing the process that we are working with, it is a simple generalization of single-type Galton–Watson processes with immigration. Going from one dimension to two means that now we are dealing with vectors and matrices and we have the tools of linear algebra at our disposal, for example we make good use of the Frobenius–Perron theorem. Then we define a classification of these processes, based on the spectral radius of the offspring mean matrix, we distinguish 3 categories: subcritical, critical, and supercritical, for this thesis we focus on the critical case. We spend some time exploring the eigenvalues and eigenvectors of the offspring mean matrix, as they are used to describe the limiting behaviour of the process and will be instrumental in a decomposition introduced later. Theorem 2.4 by Ispány and Pap [12] describes the aforementioned limiting behaviour of the process. This limit is curious, because it is degenerate in the sense that it is concentrated on a single line whose direction is determined by the right Perron vector of the offspring mean matrix. Finally we introduce a set of conditions that we will reference throughout the thesis.

Generally one would find an estimator and then apply Theorem 2.4 alongside some form of the continuous mapping theorem to describe its limiting behaviour. We do exactly that using the conditional least squares method, and we find that Theorem 2.4 is insufficient for our purposes. The inverse of some matrix $\boldsymbol{A}_n$ appears in the formulation whose limit can be described using the continuous mapping theorem, but whose limit is non-invertible and our attempt fails here. We theorize that the problem is Theorem 2.4 being incomplete and point to the curious phenomenon about the limit described in it as starting point for further investigation. This concludes the preliminaries.

Section 3 opens with a decomposition of the process based on the (left) eigenvectors of the offspring mean matrix. We introduce the random variables U_k in (12) as the *well-behaved* part of our decomposition, we know their limiting behaviour, it coincides with the underlying 1-dimensional stochastic process in Theorem 2.4. Then we introduce the random variables V_k in (16) as the *problematic* part of the decomposition.This part doesn't contribute to the limit of the process, but as we will see it later it does in case of the estimator.

Any estimator based on observing the process can be rewritten in terms of the variables U_k and V_k. This is purely a theoretical tool, if we only observe the process then we don't have information on the eigenvectors of the offspring mean matrix, thus we cannot tell the values of U_k or V_k, however if we understand their limiting behaviour then we can use them to prove theorems about the estimators. The rest of this section is devoted to the asymptotic study of these variables and various expressions of them.

We prove some bounds on the growth of the moments in Lemma 3.5, then we use those to establish a set of zero-limit theorems in Corollary 3.6. We don't have to prove this corollary as it can be done in exactly the same way as in Barczy et al. [3, Corollary 9.2]. These lower bounds on the scaling necessary to get a limit of 0 are not sharp, and more importantly not good enough for our future proofs. Where we need it later we improve these bounds, these results are expressed in Lemmas 3.8, 3.9, and 3.10. We use Theorem 3.12 by Ispány and Pap [11] to prove our main result, it gives a set of sufficient conditions under which random step processes formed from martingale differences converge to a diffusion process. Our main results are contained in Theorem 3.13, although they are easier to grasp in Corollary 3.15, where we formulated them in way that best suits our purpose.

The proof of Theorem 3.13 is just a careful checking of the conditions of Theorem 3.12 using our set of supporting Lemmas built up in the beginning of Section 3. We note that the difficulty of this part is in figuring out the correct limit, that can be proven. This was done via an iterative process, we had an educated guess on the limit, we tried to prove it, but failed repeatedly, each failure giving us some insight and getting us closer to the correct formulation.

Briefly in Section 3 we identified a set of building blocks with known limiting behaviour. Any estimator we can build (or reformulate) using these building blocks can be examined using our toolkit.

Section 4 contains 3 subsections, each dedicated to finding and analysing the estimator under different assumptions. Finding the formula for the estimators is a simple minimization problem. Since we know from our earlier observations that the estimator contains the inverse of a matrix whose limit is not invertible we use the adjugate matrix to express the inverse before trying to find the limit distribution.

We gain insight into why Theorem 2.4 was insufficient when we use stochastic expansions. It is a method by which we write an expression as a sum of parts where we now the right scaling and non-zero limit for each part, see for example (73). When multiplying stochastic expansions together sometimes the leading term vanishes, it is because in 2-dimension you can multiply together 2 non-zero matrices with the result being a zero matrix, to see an example of that look at formula (75). This enables the lower order terms to dictate the limit behaviour, this the reason why we needed to work on the problematic part of our decomposition.

We call the spectral radius of the offspring mean matrix the criticality parameter and as it can be expressed as a function of the matrix elements we can naturally define an estimator for it. In the doubly symmetric case in subsection 4.1 with the restrictions on the matrix structure is is a linear function of the matrix elements, therefore it is easy to handle. In the general cases however the spectral radius is a non-linear function of the matrix elements and require quite a bit of work to establish asymptotic results.

The last section, Section 5 contains a discussion on the results and some open questions. It discusses how one could try to relax the rather high moment conditions of our theorems and also sheds some light on the difficulty of generalizing the results to an arbitrary number of types. We also discuss there that while our theorems prove that the estimators for the offspring mean matrix are weakly consistent and describe their limiting behaviour there is rather big obstacle in their application. We cannot construct statistical tests using these results as the limit distributions depend on the very things we aim to estimate in an intricate way, namely they appear in the drift and volatility term of the stochastic differential equation describing the underlying one-dimensional process of Corollary 3.15, see SDE (61) for more details.

Apart from the results themselves the biggest takeaway is, that whenever a standard method fails, it is worthwhile to understand why it did. Understanding the reason of failure often reveals other potential angles of attack on the problem.

B Összefoglaló

Ebben a fejezetben összefoglaljuk a dolgozat eredményeit. A helytakarékosság jegyében a formulákat és a dolgozat többi részére való hivatkozást igyekszünk minimális szinten tartani.

Mint az a címből is kiderül a disszertáció célja az utódeloszlás várható érték mátrixának becslése kritikus, kéttípusos, bevándorlásos Galton–Watson folyamatokban. Miután bevezetjük a szükséges előismereteket a 2. Fejezetben, a tényleges becslés és annak aszimptotikus vizsgálata két nagy részből tevődik össze. Először a 3. Fejezetben felépítünk egy eszköztárat ezen modellbeli becslések aszimptotikus vizsgálatára, majd a 4. Fejezetben bevezetjük a vizsgálni kívánt becsléseket és az eszköztárunk felhasználásával határeloszlás tételeket adunk rájuk. A disszertáció lényegi tartalma az 5. Fejezettel zárul, mely az eredmények diszkusszióját tartalmazza. Az alábbiakban a kutatás során felhasznált kulcsfontosságú ötletek és észrevételek egy rövid összefoglalása olvasható.

Értelemszerűen a modell bemutatásával kezdünk, ami egyszerű általánosítása az egytípusos, bevándorlásos Galton–Watson folyamatoknak. Mivel áttérünk egy dimenzióról kettőre, itt már vektorokkal és mátrixokkal dolgozunk, ez lehetővé teszi, hogy lineáris algebrai tételeket alkalmazzunk, például a Frobenius–Perron tétel kimondottan hasznunkra válik. Ezután a bevezetjük a folyamatok klasszifikálását, ez az utódeloszlás várható érték mátrixának spektrálsugara alapján történik, 3 csoportot különböztetünk meg, ezek a szubkritikus, a kritikus, és a szuperkritikus. Jelen disszertáció a kritikus esetre fókuszál. Foglalkozunk még a várható érték mátrix sajátvektoraival, erre a folyamat aszimptotikus viselkedésének leírásához van szükség, valamint ezen vektorok képezik az alapját a később bevezetett felbontásnak. A 2.4 Tétel leírja az előbb említett aszimptotikus viselkedést. A határeloszlás különös, ugyanis degenerált abban az értelemben, hogy a sík egyetlen egyenesére van korlátozva, melynek irányát a várható érték mátrix jobb oldali Perron vektora határozza meg. Végezetül az itt felhasznált kritériumoknak saját nevet és jelölést adunk, mivel ezekre a későbbiekben többször hivatkozunk.

Általában ilyen problémákban kézenfekvő módszer a becslések vizsgálatára a folytonos leképezések tételének alkalmazása a 2.4 Tétellel karöltve. Mi is ezt tesszük, becslési módszernek pedig a feltételes legkisebb négyzetek módszerét alkalmazzuk, azonban azt találjuk, hogy a 2.4 Tétel nem elégséges a céljainkhoz. A becslések vizsgálatakor megjelenik egy $\boldsymbol{A}_n$ mátrix inverze és ugyan a folytonos leképezések tételével adhatunk határeloszlás tételt a mátrixra, azonban a határértékben megjelenő mátrix nem invertálható, így a módszerünk itt megbukik. Azt sejtjük, hogy a probléma abban áll, hogy a 2.4 Tétel nem ad elég átfogó képet a becslés építőelemeinek aszimptotikus viselkedéséről, ezért az előbb említett különös jelenség nyomán kezdünk vizsgálódni.

A 3. Fejezet elején bevezetjük a folyamat felbontását a várható érték mátrix baloldali sajátvektorai alapján. Elsőként az U_k változókat vezetjük be a (12) képletben, ezeket nevezzük a felbontás *jól viselkedő* részének, mivel határeloszlása egybeesik a 2.4 Tételben megbúvó egydimenziós folyamattal. A felbontás másik tagját, a *problémás* részt, a (16) formulában bevezetett V_k változók képezik. Ezen változók szerepe nem tükröződik a folyamat aszimptotikus viselkedésében, azonban mint később látni fogjuk a becslések aszimptotikájához már hozzájárulnak.

Bármely becslés amit a folyamat megfigyelésével felírhatunk átírható az előbb bevezetett felbontás szerint. Megjegyezzük, hogy ez pusztán elméleti eszköz, ha csak a folyamatot figyeljük meg, akkor nem rendelkezünk információval a várható érték mátrix sajátvektorairól és így nem ismerjük az U_k és V_k értékeket. Azonban amennyiben megértjük ezen változók aszimptotikus viselkedését, akkor ezt felhasználhatjuk a becslésekre vonatkozó határeloszlás tételek bizonyításához. A fejezet további részei ezen változók és különböző kifejezéseik vizsgálatával foglalkoznak.

A 3.5 Lemmában a várható értékek növekedési rátájára adunk felső korlátokat majd ezek segítségével a 3.6 Következményben nullához tartó határérték tételeket igazolunk. Ezt a következményt nem vezetjük le, csupán megadjuk a megfelelő hivatkozást egy hasonló következményhez, melynek bizonyítása lépésről lépésre átültethető a mi modellünkre. Az így kapott alsó becslések a 0 határértékhez szükségez skálázás nagyságrendjére nem élesek, sőt nem elegendőek a később bizonyítandó tételeinkhez. Éppen ezért speciális esetekben javítunk a becsléseinken, ezen eredményeket a 3.8, 3.9, és 3.10 Lemmák írják le. A felbontásra vonatkozó nemnulla határeloszlás tételünket a 3.12 Tétel segítségével bizonyítjuk. Ispány Márton és Pap Gyula ezen eredménye elégéseges feltételt ad martingálkülönbségekből képzett lépcsős függvények diffúziós folyamathoz való konvergenciájára. A fejezet fő eredménye a 3.13 Tétel, melynek eredményeit a 3.15 Következményben oly módon fogalmaztunk meg, mely a legcélszerűbb a későbbi alkalmazásuk szempontjából.

A fő tétel bizonyítása nem más mint a 3.12 Tétel feltételeinek ellenőrzése a korábban bevezetett segédlemmáink felhasználásával. A tétel kitalálásában a határeloszlás leírása jelentette a legnagyobb nehézséget, hogy pontosan mihez is konvergál az általunk vizsgált folyamat. A helyes választ egy iteratív folyamat végeredményeként kaptuk, ahol a határeloszlásra adott intuitív tippünkből indulva minden egyes sikertelen bizonyítási kísérlettel egyre közelebb kerültünk a helyes eredményhez.

Röviden összefoglalva a 3. Fejezetben azonosítottunk néhány építőelemet melynek ismerjük az aszimptotikus viselkedését. Az így kapott eszköztár felhasználható ezen építőelemek segítségével kifejezett becslések vizsgálatára.

A 4. Fejezet 3 alfejezetből áll, mindegyik az utódeloszlás várható érték mátrixának becslésével foglalkozik különböző feltételezések mellett. A becslés meghatározása egy egyszerű minimalizálási feladat. Mivel az előző sikertelen próbálkozásunkból tudjuk, hogy ahol megjelenik a mátrix inverze ott az a határérték nem invertálható ezért eleve az adjungált mátrix segítségével írjuk fel az inverzet.

Mikor sztochasztikus kifejtést használunk, akkor értjük meg, hogy miért nem elégséges a 2.4 Tétel. Ez egy olyan módszer, ahol a vizsgálni kívánt kifejezést felírjuk olyan tagok összegeként melynek ismerjük a megfelelő skálázását és nemnulla határértékét, ilyenre az olvasó a (73) képletben talál példát. Mikor sztochasztikus kifejtéseket szorzunk össze, akkor a legmagasabb rangú tag eltűnhet, ez azért van mert 2 dimenzióban össze tudunk szorozni nemnulla mátrixokat úgy, hogy az eredmény nullmátrix legyen. Erre szolgáltat példát a (75) összefüggés. Így lehetséges, hogy kisebb rendű tagok is szerepet játsszanak a határeloszlásban, emiatt volt szükség a felbontásunk *problémás* felének vizsgálatára.

Az utódeloszlás várható érték mátrixának spektrálsugarára kritikussági paraméterként is hivatkozhatunk. Mivel ez a mennyiség kifejezhető a mátrix elemeinek függvényeként, ezért természetes módon kapjuk a becslését a mátrix becsléséből. A duplán szimmetrikus esetben, a 4.1 Fejezetben a mátrix struktúrájára tett megszorításaink miatt az összefüggés a spektrálsugár és a mátrix elemei között lineáris, következésképpen a kritikussági paraméter becslése könnyen kezelhető. Az általános esetben sajnos ennél bonyolultabb a helyzet, jóval több munkára van szükség a határeloszlás vizsgálatához.

Az utolsó, 5. Fejezet az eredmények és néhány nyitott kérdés diszkusszziója. Szó esik benne arról, hogy milyen módon lehetne a meglehetősen magas momentumfeltételek gyengíteni, valamint rávilágít, hogy milyen nehézségekkel nézne szembe az, aki magasabb dimenziókra próbálná általánosítani az eredményeket. Arról is itt esik említés, hogy ugyan a határeloszlás tételeink leírják a becslések aszimptotikus viselkedését a statisztikai próbák konstruálásának van egy nagy akadálya, a határeloszlás elég összetett módon függ a becsülni kívánt mennyiségtől. Az utódeloszlás várható érték mátrixának sajátvektorai megjelennek a a határeloszlás leírásához használt sztochasztikus folyamatot meghatározó sztochasztikus differenciálegyenlet együtthatóiban, mely a (61) képletben található.

Az eredmények mellett ezen kutatás fő tanúságaként azt emelnénk ki, hogy ahol a bevett módszerek csődöt mondanak, ott érdemes alaposan megvizsgálni, hogy mely ponton bukik meg a folyamat, ez ugyanis gyakran nyomként szolgálhat arra nézve, hogy milyen irányból érdemes megközelíteni a problémát.

C List of the author's publications

Below is a list of the author's publications, and a short explanation on how do these connect to the results of this thesis.

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.

This publication contains the first results we had about the core problem of this dissertation. We use heavy restrictions on the structure of the offspring mean matrix to prove limit theorems. The main results are reproduced in Section 4.1 of this thesis.

Körmendi, K. and Pap, G. (2018). Statistical inference of 2-type critical Galton–Watson processes with immigration. *Statistical Inference for Stochastic Processes* **21(1)** 169–190.

In this article we treat the general case, this work is the basis of the dissertation. The main results are reproduced in Section 4.2. The publication contains two limit theorems which are not included in this thesis: one in the critical case if (ND) doesn't hold, and another that describes the asymptotic behaviour of the estimator in the subcritical case.

Barczy, M., Körmendi, K. and Pap, G. (2015). Statistical inference for 2-type doubly symmetric critical irreducible continuous state and continuous time branching processes with immigration. *Journal of Multivariate Analysis* **139(2015)** 92–123.

The method of this thesis has been adapted to continuous state and continuous time branching processes, albeit with heavy restrictions on the structure of these models. The results are in the above publication.

Barczy, M., Körmendi, K. and Pap, G. (2016). Statistical inference for critical continuous state and continuous time branching processes with immigration. *Metrika* **79(7)** 789–816.

During the study of 2-type models the authors realized that results are not available even in for the single-type version of the problem, so they solved and published it.

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] Danka, T. and Pap, G. (2016). Asymptotic behavior of critical indecomposable multi-type branching processes with immigration. *European Series in Applied and Industrial Mathematics (ESAIM). Probability and Statistics.* **20** 238–260.

[5] Heyde, C. C. and Seneta, E. (1972). Estimation Theory for Growth and Immigration Rates in a Multiplicative Process *Journal of Applied Probability* **9(2)** 235–256.

[6] Heyde, C. C. and Seneta, E. (1974). Notes on "Estimation Theory for Growth and Immigration Rates in a Multiplicative Process" *Journal of Applied Probability* **11(3)** 572–577.

[7] Haccou, P., Jagers, P. and Vatutin, V. 2005. *Branching Processes*, Cambridge University Press, Cambridge.

[8] Horn, R. A. and Johnson, Ch. R. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

[9] Ikeda, N. and Watanabe, S. (1992). *Stochastic Differential Equations and Diffusion Processes, Volume 24 (2nd edition)* North Holland

[10] 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.

[11] Ispány, M. and Pap, G. (2010). A note on weak convergence of step processes. *Acta Mathematica Hungarica* **126(4)** 381–395.

[12] 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.

[13] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[14] KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *The Annals of Mathematical Statistics* **37(5)** 1211–1223.

[15] KLIMKO, L. A. and NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* **6(3)** 629–642.

[16] KÖRMENDI, K. and PAP, G. (2018). Statistical inference of 2-type critical Galton–Watson processes with immigration. *Statistical Inference for Stochastic Processes* **21(1)** 169–190.

[17] MODE, C. J. (1971). *Multitype Branching Processes: Theory and Applications*, American-Elsevier, New York.

[18] NELSON, P. I. (1980). A Note on Strong Consistency of Least Squares Estimators in Regression Models with Martingale Difference Errors. *The Annals of Statistics* **8(5)** 1057–1064.

[19] 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.

[20] QUINE, M. P. (1970). The multi-type Galton–Watson process with immigration. *Journal of Applied Probability* **7(2)** 411–422.

[21] QUINE, M. P. and DURHAM, P. (1977). Estimation for multitype branching processes. *Journal of Applied Probability* **14(4)** 829–835.

[22] SHETE, S. and SRIRAM, T. N. (2003). A note on estimation in multitype supercritical branching processes with immigration. *Sankhyā: The Indian Journal of Statistics* **65(1)** 107–121.

[23] 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.

[24] 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.

[25] Wei, C. Z. and Winnicki, J. (1990). Estimation of the means in the branching process with immigration. *The Annals of Statistics* **18** 1757–1773.

[26] 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.

[27] Winnicki, J. (1991). Estimation of the variances in the branching process with immigration. *Probability Theory and Related Fields* **88(1)** 77–106.