Bernoulli **20**(4), 2014, 2247–2277
DOI: 10.3150/13-BEJ556

Asymptotic behavior of CLS estimators for 2-type doubly symmetric critical Galton–Watson processes with immigration

MÁRTON ISPÁNY[1], KRISTÓF KÖRMENDI[2,*] and GYULA PAP[2,**]

[1] *University of Debrecen, Faculty of Informatics, Department of Information Technology, Pf. 12, H-4010 Debrecen, Hungary. E-mail: ispany.marton@inf.unideb.hu*
[2] *University of Szeged, Faculty of Science, Bolyai Institute, Department of Stochastics, Aradi vértanúk tere 1, H-6720 Szeged, Hungary. E-mail: * kormendi@math.u-szeged.hu; ** papgy@math.u-szeged.hu*

In this paper, the asymptotic behavior of the conditional least squares (CLS) estimators of the offspring means (α, β) and of the criticality parameter $\varrho := \alpha + \beta$ for a 2-type critical doubly symmetric positively regular Galton–Watson branching process with immigration is described.

Keywords: conditional least squares estimator; Galton–Watson branching process with immigration

1. Introduction

Asymptotic behavior of CLS estimators for critical Galton–Watson processes is available only for single-type processes, see Wei and Winnicki [20,21] and Winnicki [22], see also the monograph of Guttorp [4]. In the present paper, the asymptotic behavior of the CLS estimators of the offspring means and criticality parameter for 2-type critical doubly symmetric positively regular Galton–Watson process with immigration is described, see Theorem 3.1. This study can be considered as the first step of examining the asymptotic behavior of the CLS estimators of parameters of multitype critical branching processes with immigration. Shete and Sriram [18] obtained convergence results for weighted CLS estimators in the supercritical case.

Let us recall the results for a single-type Galton–Watson branching process $(X_k)_{k\in\mathbb{Z}_+}$ with immigration and with initial value $X_0 = 0$. Suppose that it is critical, that is, the offspring mean equals 1. Wei and Winnicki [20] proved a functional limit theorem $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ as $n \to \infty$, where $\mathcal{X}_t^{(n)} := n^{-1} X_{\lfloor nt \rfloor}$ for $t \in \mathbb{R}_+$, $n \in \mathbb{N}$, where $\lfloor x \rfloor$ denotes the (lower) integer part of $x \in \mathbb{R}$, and $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is a (nonnegative) diffusion process with initial value $\mathcal{X}_0 = 0$ and with generator

$$Lf(x) = m_\varepsilon f'(x) + \tfrac{1}{2} V_\xi x f''(x), \qquad f \in C_{\mathrm{c}}^\infty(\mathbb{R}_+),$$

where m_ε denotes the immigration mean, V_ξ denotes the offspring variance, and $C_{\mathrm{c}}^\infty(\mathbb{R}_+)$ denotes the space of infinitely differentiable functions on $\mathbb{R}_+$ with compact support. The process $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ can also be characterized as the unique strong solution of the stochastic differential equation (SDE)

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

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process, and x^+ denotes the positive part of $x \in \mathbb{R}$. Note that this so-called square-root process is also known as Feller diffusion, or Cox–Ingersoll–Ross model in financial mathematics (see Musiela and Rutkowski [15], page 290). In fact, $(4V_\xi^{-1}\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is the square of a $4V_\xi^{-1}m_\varepsilon$-dimensional Bessel process started at 0 (see Revuz and Yor [17], XI.1.1).

Assuming that the immigration mean m_ε is known, for the conditional least squares estimator (CLSE)

$$\widehat{\alpha}_n(X_1, \dots, X_n) = \frac{\sum_{k=1}^n X_{k-1}(X_k - m_\varepsilon)}{\sum_{k=1}^n X_{k-1}^2}$$

of the offspring mean based on the observations $X_1, \dots, X_n$, one can derive

$$n\bigl(\widehat{\alpha}_n(X_1, \dots, X_n) - 1\bigr) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{X}_t \,\mathrm{d}(\mathcal{X}_t - m_\varepsilon t)}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t} \qquad \text{as } n \to \infty.$$

(Wei and Winnicki [21] contains a similar result for the CLS estimator of the offspring mean when the immigration mean is unknown.)

In Section 2, we recall some preliminaries on 2-type Galton–Watson models with immigration. Section 3 contains our main results. Sections 4, 5, 6 and 7 contain the proofs. Appendix A is devoted to the CLS estimators. In Appendix B, we present estimates for the moments of the processes involved. Appendices C and D are for a version of the continuous mapping theorem and for convergence of random step processes, respectively. For a detailed discussion of the whole paper, see Ispány *et al.* [8].

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

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

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

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

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

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

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

$$\mathbf{m}_{\boldsymbol{\xi}_i} := \mathbb{E}(\boldsymbol{\xi}_{1,1,i}) \in \mathbb{R}_+^2, \qquad \mathbf{m}_{\boldsymbol{\xi}} := [\,\mathbf{m}_{\boldsymbol{\xi}_1} \mathbf{m}_{\boldsymbol{\xi}_2}\,] \in \mathbb{R}_+^{2\times 2},$$
$$\mathbf{V}_{\boldsymbol{\xi}_i} := \operatorname{Var}(\boldsymbol{\xi}_{1,1,i}) \in \mathbb{R}^{2\times 2}, \qquad \overline{\mathbf{V}}_{\boldsymbol{\xi}} := \tfrac{1}{2}(\mathbf{V}_{\boldsymbol{\xi}_1} + \mathbf{V}_{\boldsymbol{\xi}_2}) \in \mathbb{R}^{2\times 2},$$
$$\mathbf{m}_{\boldsymbol{\varepsilon}} := \mathbb{E}(\boldsymbol{\varepsilon}_1) \in \mathbb{R}_+^2, \qquad \mathbf{V}_{\boldsymbol{\varepsilon}} := \operatorname{Var}(\boldsymbol{\varepsilon}_1) \in \mathbb{R}^{2\times 2}.$$

Note that many authors define the offspring mean matrix as $\mathbf{m}_{\boldsymbol{\xi}}^\top$. For $k \in \mathbb{Z}_+$, let $\mathcal{F}_k := \sigma(\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_k)$. By (2.1),

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

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

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

Hence, the offspring mean matrix $\mathbf{m}_{\boldsymbol{\xi}}$ plays a crucial role in the asymptotic behavior of the sequence $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$. Since $\mathbf{m}_{\boldsymbol{\xi}}$ has nonnegative entries, the Frobenius–Perron theorem (see, e.g., Horn and Johnson [7], Theorems 8.2.11 and 8.5.1) describes the behavior of the powers $\mathbf{m}_{\boldsymbol{\xi}}^k$ as $k \to \infty$. According to this behavior, a 2-type Galton–Watson process $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ with immigration is referred to respectively as *subcritical*, *critical* or *supercritical* if $\varrho < 1$, $\varrho = 1$ or $\varrho > 1$, where ϱ denotes the spectral radius of the offspring mean matrix $\mathbf{m}_{\boldsymbol{\xi}}$ (see, e.g., Athreya and Ney [1] or Quine [16]). We will consider doubly symmetric 2-type Galton–Watson processes with immigration, when the offspring mean matrix has the form

$$\mathbf{m}_{\boldsymbol{\xi}} := \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}. \tag{2.4}$$

Its spectral radius is $\varrho = \alpha + \beta$, which will be called *criticality parameter*. We will focus only on *positively regular* doubly symmetric 2-type Galton–Watson processes with immigration, that is, when there is a positive integer $k \in \mathbb{N}$ such that the entries of $\mathbf{m}_{\boldsymbol{\xi}}^k$ are positive (see Kesten and Stigum [13]), which is equivalent with $\alpha > 0$ and $\beta > 0$.

For the sake of simplicity, we consider a zero start Galton–Watson process with immigration, that is, we suppose $\mathbf{X}_0 = \mathbf{0}$. In the sequel, we always assume $\mathbf{m}_\varepsilon \neq \mathbf{0}$, otherwise $\mathbf{X}_k = \mathbf{0}$ for all $k \in \mathbb{N}$.

3. Main results

In order to find CLS estimators of the criticality parameter $\varrho = \alpha + \beta$, we introduce a further parameter $\delta := \alpha - \beta$. Then $\alpha = (\varrho + \delta)/2$ and $\beta = (\varrho - \delta)/2$, thus the recursion (4.2) can be written in the form

$$\mathbf{X}_k = \frac{1}{2}\begin{bmatrix} \varrho+\delta & \varrho-\delta \\ \varrho-\delta & \varrho+\delta \end{bmatrix}\mathbf{X}_{k-1} + \mathbf{M}_k + \mathbf{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}.$$

For each $n \in \mathbb{N}$, a CLS estimator $(\widehat{\varrho}_n, \widehat{\delta}_n)$ of (ϱ, δ) based on a sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ can be obtained by minimizing the sum of squares

$$\sum_{k=1}^{n} \left\| \mathbf{X}_k - \frac{1}{2}\begin{bmatrix} \varrho+\delta & \varrho-\delta \\ \varrho-\delta & \varrho+\delta \end{bmatrix}\mathbf{X}_{k-1} - \mathbf{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

with respect to (ϱ, δ) over $\mathbb{R}^2$, and it has the form

$$\widehat{\varrho}_n := \frac{\sum_{k=1}^n \langle \mathbf{1}, \mathbf{x}_k - \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle \langle \mathbf{1}, \mathbf{x}_{k-1}\rangle}{\sum_{k=1}^n \langle \mathbf{1}, \mathbf{x}_{k-1}\rangle^2}, \tag{3.1}$$

$$\widehat{\delta}_n := \frac{\sum_{k=1}^n \langle \widetilde{\mathbf{u}}, \mathbf{x}_k - \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle \langle \widetilde{\mathbf{u}}, \mathbf{x}_{k-1}\rangle}{\sum_{k=1}^n \langle \widetilde{\mathbf{u}}, \mathbf{x}_{k-1}\rangle^2} \tag{3.2}$$

on the set $H_n \cap \widetilde{H}_n$, where

$$\mathbf{1} := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2, \qquad \widetilde{\mathbf{u}} := \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in \mathbb{R}^2,$$

and

$$H_n := \left\{ (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{R}^2)^n : \sum_{k=1}^n \langle \mathbf{1}, \mathbf{x}_{k-1}\rangle^2 > 0 \right\}, \tag{3.3}$$

$$\widetilde{H}_n := \left\{ (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{R}^2)^n : \sum_{k=1}^n \langle \widetilde{\mathbf{u}}, \mathbf{x}_{k-1}\rangle^2 > 0 \right\}, \tag{3.4}$$

where $\mathbf{x}_0 := \mathbf{0}$ is the zero vector in $\mathbb{R}^2$. In a natural way, we extend the CLS estimators $\widehat{\varrho}_n$ and $\widehat{\delta}_n$ to the set H_n and $\widetilde{H}_n$, respectively. Moreover, for each $n \in \mathbb{N}$, any CLS estimator $(\widehat{\alpha}_n, \widehat{\beta}_n)$ of the offspring means (α, β) based on a sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ has the form

$$\begin{bmatrix} \widehat{\alpha}_n \\ \widehat{\beta}_n \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n \\ \widehat{\delta}_n \end{bmatrix}, \tag{3.5}$$

whenever the sample belongs to the set $H_n \cap \widetilde{H}_n$. For the proof see Ispány *et al.* [8], Lemma A.1.

In what follows, we always assume that $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ is a 2-type doubly symmetric Galton–Watson process with offspring means $(\alpha, \beta) \in (0, 1)^2$ such that $\alpha + \beta = 1$ (hence it is critical and positively regular), $\mathbf{X}_0 = \mathbf{0}$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8) < \infty$, and $\mathbf{m}_{\boldsymbol{\varepsilon}} \neq \mathbf{0}$. Then $\lim_{n\to\infty} \mathbb{P}((\mathbf{X}_1, \ldots, \mathbf{X}_n) \in H_n) = 1$. If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle > 0$, or if $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ and $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1\rangle^2) > 0$, then $\lim_{n\to\infty} \mathbb{P}((\mathbf{X}_1, \ldots, \mathbf{X}_n) \in \widetilde{H}_n) = 1$, see Proposition A.3.

Let $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ be the unique strong solution of the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{Y}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\,\mathrm{d}t + \sqrt{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle \mathcal{Y}_t^+}\,\mathrm{d}\mathcal{W}_t, \qquad t\in\mathbb{R}_+,\ \mathcal{Y}_0 = 0, \tag{3.6}$$

where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process.

Theorem 3.1. *We have*

$$n(\widehat{\varrho}_n - 1) \stackrel{\mathcal{D}}{\longrightarrow} \frac{\int_0^1 \mathcal{Y}_t\,\mathrm{d}(\mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle t)}{\int_0^1 \mathcal{Y}_t^2\,\mathrm{d}t}, \qquad \text{as } n\to\infty. \tag{3.7}$$

If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$, *then*

$$n^{3/2}(\widehat{\varrho}_n - 1) \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}\left(0, \frac{3\langle \mathbf{V}_{\boldsymbol{\varepsilon}}\mathbf{1}, \mathbf{1}\rangle}{\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2}\right), \qquad \text{as } n\to\infty. \tag{3.8}$$

If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle > 0$, *then*

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n - \alpha) \\ n^{1/2}(\widehat{\beta}_n - \beta) \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \sqrt{\alpha\beta}\,\frac{\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{Y}_t\,\mathrm{d}t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \text{as } n\to\infty, \tag{3.9}$$

where $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *is a standard Wiener process, independent from* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$.

If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ *and* $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1\rangle^2) > 0$, *then*

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n - \alpha) \\ n^{1/2}(\widehat{\beta}_n - \beta) \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{N}\left(0, \frac{\langle \mathbf{V}_{\boldsymbol{\varepsilon}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle}{4\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1\rangle^2)}\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \text{as } n\to\infty. \tag{3.10}$$

Remark 3.2. If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle > 0$ and $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$ then in (3.9) we have

$$\sqrt{\alpha\beta}\,\frac{\int_0^1 \mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{Y}_t\,\mathrm{d}t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \stackrel{\mathcal{D}}{=} \mathcal{N}\left(0, \frac{4}{3}\alpha\beta\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Remark 3.3. Note that the assumption $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$ is fulfilled if and only if $\xi_{1,1,1,1} + \xi_{1,1,1,2} \stackrel{\text{a.s.}}{=} 1$ and $\xi_{1,1,2,1} + \xi_{1,1,2,2} \stackrel{\text{a.s.}}{=} 1$, that is, the total number of offsprings produced by an individual of type 1 is 1, and the same holds for individuals of type 2. In a similar way, the assumption $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ is fulfilled if and only if $\alpha = \beta = \frac{1}{2}$, $\xi_{1,1,1,1} \stackrel{\text{a.s.}}{=} \xi_{1,1,1,2}$ and $\xi_{1,1,2,1} \stackrel{\text{a.s.}}{=} \xi_{1,1,2,2}$, that is, the number of offsprings of type 1 and of type 2 produced by an individual of type 1 are the same, and the same holds for individuals of type 2. Observe that the assumptions $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$ and $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ can not be fulfilled at the same time.

Condition $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1\rangle^2) > 0$ fails to hold if and only if $\varepsilon_{1,1} - \varepsilon_{1,2} \stackrel{\text{a.s.}}{=} 0$, and, under the assumption $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$, this implies $X_{k,1} \stackrel{\text{a.s.}}{=} X_{k,2}$ (see Lemma A.2), when $\mathbb{P}((\mathbf{X}_1, \ldots, \mathbf{X}_n) \in H_n \cap \widetilde{H}_n) = 0$ for all $n\in\mathbb{N}$, and hence the LSE of the offspring means (α, β) is not defined uniquely, see Appendix A.

Remark 3.4. For each $n \in \mathbb{N}$, consider the random step process

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

Theorem 5.1 implies convergence (5.3), hence

$$\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X} := \tfrac{1}{2}\mathcal{Y}\mathbf{1} \qquad \text{as } n \to \infty, \tag{3.11}$$

where the process $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the SDE (3.6) with initial value $\mathcal{Y}_0 = 0$. Note that convergence (3.11) holds even if $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$, when the unique strong solution of (3.6) is the deterministic function $\mathcal{Y}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle t$, $t \in \mathbb{R}_+$.

The SDE (3.6) has a unique strong solution $(\mathcal{Y}_t^{(y)})_{t\in\mathbb{R}_+}$ for all initial values $\mathcal{Y}_0^{(y)} = y \in \mathbb{R}$, and if $y \geq 0$, then $\mathcal{Y}_t^{(y)}$ is nonnegative for all $t \in \mathbb{R}_+$ with probability one, hence $\mathcal{Y}_t^+$ may be replaced by $\mathcal{Y}_t$ under the square root in (3.6), see, for example, Barczy *et al.* [3], Remark 3.3.

Remark 3.5. We note that in the critical positively regular case the limit distributions for the CLS estimators of the offspring means (α, β) are concentrated on the line $\{(u, v) \in \mathbb{R}^2 : u + v = 0\}$. In order to handle the difficulty caused by this degeneracy, we use an appropriate reparametrization. Surprisingly, the scaling factor of the CLS estimators of (α, β) is always $\sqrt{n}$, which is the same as in the subcritical case. The reason of this strange phenomenon can be understood from the joint asymptotic behavior of the numerator and the denominator of the CLS estimators given in Theorems 4.1, 4.2 and 4.3. The scaling factor of the estimators of the criticality parameter ϱ is usually n, except in a particular special case of $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$, when it is $n^{3/2}$. One of the decisive tools in deriving the needed asymptotic behavior is a good bound for the moments of the involved processes, see Corollary B.6.

Remark 3.6. The shape of $\int_0^1 \mathcal{Y}_t \, \mathrm{d}(\mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle t) / \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t$ in (3.7) is similar to the limit distribution of the Dickey–Fuller statistics for unit root test of AR(1) time series, see, for exmple, Hamilton [6], formulas 17.4.2 and 17.4.7, or Tanaka [19], (7.14) and Theorem 9.5.1. The shape of $\int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t / \int_0^1 \mathcal{Y}_t \, \mathrm{d}t$ in (3.9) is also similar, but it contains two independent standard Wiener processes. This phenomenon is very similar to the appearance of two independent standard Wiener processes in limit theorems for CLS estimators of the variance of the offspring and immigration distributions for critical branching processes with immigration in Winnicki [22], Theorems 3.5 and 3.8. Finally, note that the limit distribution of the CLS estimator of the criticality parameter ϱ is non-symmetric and non-normal in (3.7), and symmetric normal in (3.8), but the limit distribution of the CLS estimator of the offspring means (α, β) is always symmetric, although non-normal in (3.9).

Remark 3.7. The eighth order moment conditions on the offspring and immigration distributions in Theorem 3.1 seem to be too strong, but we note that the process $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ can be considered as a heteroscedastic time series. Indeed, $\mathbf{X}_k = \mathbf{m}_{\boldsymbol{\xi}}\mathbf{X}_{k-1} + \mathbf{m}_{\boldsymbol{\varepsilon}} + \mathbf{M}_k$, see (4.2), and by (B.1), $\mathbb{E}(\mathbf{M}_k\mathbf{M}_k^\top | \mathcal{F}_{k-1}) = X_{k-1,1}\mathbf{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\mathbf{V}_{\boldsymbol{\xi}_2} + \mathbf{V}_{\boldsymbol{\varepsilon}}$, $k \in \mathbb{N}$. That is why we think that the behavior of the process $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ is similar to GARCH models, where, even in the stable case, high moment conditions are needed for convergence of estimators such as the quasi-maximum likelihood estimator in Hall and Yao [5] or the Whittle estimator in Mikosch and Straumann [14].

4. Proof of the main results

Applying (2.2), let us introduce the sequence

$$\mathbf{M}_k := \mathbf{X}_k - \mathbb{E}(\mathbf{X}_k | \mathcal{F}_{k-1}) = \mathbf{X}_k - \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} - \mathbf{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}, \tag{4.1}$$

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

$$\mathbf{X}_k = \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} + \mathbf{m}_{\boldsymbol{\varepsilon}} + \mathbf{M}_k, \qquad k \in \mathbb{N}. \tag{4.2}$$

Next, let us introduce the sequence

$$U_k := \langle \mathbf{1}, \mathbf{X}_k \rangle = X_{k,1} + X_{k,2}, \qquad k \in \mathbb{Z}_+.$$

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

$$U_k = U_{k-1} + \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \mathbf{1}, \mathbf{M}_k \rangle, \qquad k \in \mathbb{N}, \tag{4.3}$$

since $\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} \rangle = \mathbf{1}^\top \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} = \mathbf{1}^\top \mathbf{X}_{k-1} = U_{k-1}$, because $\varrho = \alpha + \beta = 1$ implies that $\mathbf{1}$ is a left eigenvector of the mean matrix $\mathbf{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue 1. Hence, $(U_k)_{k \in \mathbb{Z}_+}$ is a nonnegative unstable AR(1) process with positive drift $\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle$ and with heteroscedastic innovation $(\langle \mathbf{1}, \mathbf{M}_k \rangle)_{k \in \mathbb{N}}$. Moreover, let

$$V_k := \langle \widetilde{\mathbf{u}}, \mathbf{X}_k \rangle = X_{k,1} - X_{k,2}, \qquad k \in \mathbb{Z}_+.$$

Note that we have

$$V_k = (\alpha - \beta) V_{k-1} + \langle \widetilde{\mathbf{u}}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle, \qquad k \in \mathbb{N}, \tag{4.4}$$

since $\langle \widetilde{\mathbf{u}}, \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} \rangle = \widetilde{\mathbf{u}}^\top \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} = (\alpha - \beta) \widetilde{\mathbf{u}}^\top \mathbf{X}_{k-1} = (\alpha - \beta) V_{k-1}$, because $\widetilde{\mathbf{u}}$ is a left eigenvector of the mean matrix $\mathbf{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue $\alpha - \beta$. Thus $(V_k)_{k \in \mathbb{N}}$ is a stable AR(1) process with drift $\langle \widetilde{\mathbf{u}}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle$ and with heteroscedastic innovation $(\langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle)_{k \in \mathbb{N}}$. Observe that

$$X_{k,1} = (U_k + V_k)/2, \qquad X_{k,2} = (U_k - V_k)/2, \qquad k \in \mathbb{Z}_+. \tag{4.5}$$

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

$$\widehat{\varrho}_n - 1 = \frac{\sum_{k=1}^n \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1}}{\sum_{k=1}^n U_{k-1}^2},$$

whenever $(\mathbf{X}_1, \ldots, \mathbf{X}_n) \in H_n$, where H_n, $n \in \mathbb{N}$, are given in (3.3). By (3.2), for each $n \in \mathbb{N}$, we have

$$\widehat{\delta}_n - \delta = \frac{\sum_{k=1}^n \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1}}{\sum_{k=1}^n V_{k-1}^2}, \tag{4.6}$$

whenever $(\mathbf{X}_1, \ldots, \mathbf{X}_n) \in \widetilde{H}_n$, where $\widetilde{H}_n$, $n \in \mathbb{N}$, are given in (3.4).

Theorem 3.1 will follow from the following statements by the continuous mapping theorem.

Theorem 4.1. *We have, as $n \to \infty$,*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \\ (4\alpha\beta)^{-1} \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\big(\mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t\big) \\ (4\alpha\beta)^{-1/2} \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}.$$

Theorem 4.2. *If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ then, as $n \to \infty$,*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-3/2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \\ (4\alpha\beta)^{-1} \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \int_0^1 \mathcal{Y}_t \, \mathrm{d}t \\ \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ (4\alpha\beta)^{-1/2} \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix},$$

where $(\widetilde{\widetilde{\mathcal{W}}}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process, independent from $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$. Note that $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ is now the deterministic function $\mathcal{Y}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t$, $t \in \mathbb{R}_+$, hence $\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 / 3$, $\int_0^1 \mathcal{Y}_t \, \mathrm{d}t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle / 2$, $\int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^1 t \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t$ and $\int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^1 t \, \mathrm{d}\widetilde{\mathcal{W}}_t$.

Theorem 4.3. *If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ then, as $n \to \infty$,*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-1} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-1/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \\ \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\big(\mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t\big) \\ \big[\langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big)\big]^{1/2} \widetilde{\mathcal{W}}_1 \end{bmatrix}.$$

5. Proof of Theorem 4.1

Consider the sequence of stochastic processes

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

with

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

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$, where $\otimes$ denotes Kronecker product of matrices. Theorem 4.1 follows from Lemma A.1 and the following theorem (this will be explained after Theorem 5.1).

Theorem 5.1. *We have*

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

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

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

with initial value $\mathcal{Z}_0 = \mathbf{0}$, *where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *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*

$$\gamma(t, \mathbf{x}) := \begin{bmatrix} \langle \mathbf{1}, (\mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} & 0 \\ \langle \mathbf{1}, (\mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{3/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} & 0 \\ 0 & \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \langle \mathbf{1}, \mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \end{bmatrix}$$

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

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

The SDE (5.2) has the form

$$\mathrm{d}\mathcal{Z}_t = \begin{bmatrix} \mathrm{d}\mathcal{M}_t \\ \mathrm{d}\mathcal{N}_t \\ \mathrm{d}\mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \langle \mathbf{1}, (\mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \,\mathrm{d}\mathcal{W}_t \\ \langle \mathbf{1}, (\mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{3/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \,\mathrm{d}\mathcal{W}_t \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

Ispány and Pap [9] proved that the first 2-dimensional equation of this SDE has a unique strong solution $(\mathcal{M}_t)_{t\in\mathbb{R}_+}$ with initial value $\mathcal{M}_0 = \mathbf{0}$, and $(\mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+$ may be replaced by $\mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}}$ (see the proof of [9, Theorem 3.1]). Thus, the SDE (5.2) has a unique strong solution with initial value $\mathcal{Z}_0 = \mathbf{0}$, and we have

$$\mathcal{Z}_t = \begin{bmatrix} \mathcal{M}_t \\ \mathcal{N}_t \\ \mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \,\mathrm{d}\mathcal{W}_s \\ \int_0^t \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \,\mathrm{d}\mathcal{M}_s \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^t \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \,\mathrm{d}\widetilde{\mathcal{W}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

By the method of the proof of $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ in Theorem 3.1 in Barczy *et al.* [3], applying Lemma C.2, one can easily derive

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathcal{Z}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix}, \qquad \text{as } n \to \infty, \tag{5.3}$$

where

$$\mathcal{X}_t^{(n)} := n^{-1} \mathbf{X}_{\lfloor nt \rfloor}, \qquad \mathcal{X}_t := \tfrac{1}{2} \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}, \qquad t \in \mathbb{R}_+,\ n \in \mathbb{N},$$

see Ispány *et al.* [8], page 10. Now, with the process

$$\mathcal{Y}_t := \langle \mathbf{1}, \mathcal{X}_t \rangle = \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle, \qquad t \in \mathbb{R}_+,$$

we have

$$\mathcal{X}_t = \tfrac{1}{2} \mathcal{Y}_t \mathbf{1}, \qquad t \in \mathbb{R}_+.$$

By Itô's formula, we obtain that the process $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ satisfies the SDE (3.6). Next, similarly to the proof of (A.2), by Lemma C.3, convergence (5.3) and Lemma A.1 with $U_{k-1} = \langle \mathbf{1}, \mathbf{X}_{k-1} \rangle$ implies

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \mathbf{1}, \mathcal{X}_t \rangle^2 \, \mathrm{d}t \\ \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \int_0^1 \langle \mathbf{1}, \mathcal{X}_t \rangle \, \mathrm{d}t \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\langle \mathbf{1}, \mathcal{M}_t \rangle \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\big\langle \widetilde{\mathbf{u}}, \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\mathcal{W}}_t \big\rangle \end{bmatrix},$$

as $n \to \infty$. This limiting random vector can be written in the form as given in Theorem 4.1, since $\langle \mathbf{1}, \mathcal{X}_t \rangle = \mathcal{Y}_t$, $\langle \mathbf{1}, \mathcal{M}_t \rangle = \langle \mathbf{1}, \mathcal{X}_t \rangle - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t = \mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t$ and $\langle \widetilde{\mathbf{u}}, \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\mathcal{W}}_t \rangle = \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle^{1/2} \widetilde{\mathcal{W}}_t$ for all $t \in \mathbb{R}_+$ with a (one-dimensional) standard Wiener process $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$.

Proof of Theorem 5.1. In order to show convergence $\mathcal{Z}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{Z}$, we apply Theorem D.1 with the special choices $\mathcal{U} := \mathcal{Z}$, $\mathbf{U}_k^{(n)} := \mathbf{Z}_k^{(n)}$, $n, k \in \mathbb{N}$, $(\mathcal{F}_k^{(n)})_{k \in \mathbb{Z}_+} := (\mathcal{F}_k)_{k \in \mathbb{Z}_+}$ and the function γ which is defined in Theorem 5.1. Note that the discussion after Theorem 5.1 shows that the SDE (5.2) admits a unique strong solution $(\mathcal{Z}_t^{\mathbf{z}})_{t \in \mathbb{R}_+}$ for all initial values $\mathcal{Z}_0^{\mathbf{z}} = \mathbf{z} \in (\mathbb{R}^2)^3$.

Now we show that conditions (i) and (ii) of Theorem D.1 hold. The conditional variance $\mathbb{E}(\mathbf{Z}_k^{(n)} (\mathbf{Z}_k^{(n)})^\top | \mathcal{F}_{k-1})$ has the form

$$\begin{bmatrix} n^{-2} & n^{-3} U_{k-1} & n^{-5/2} V_{k-1} \\ n^{-3} U_{k-1} & n^{-4} U_{k-1}^2 & n^{-7/2} U_{k-1} V_{k-1} \\ n^{-5/2} V_{k-1} & n^{-7/2} U_{k-1} V_{k-1} & n^{-3} V_{k-1}^2 \end{bmatrix} \otimes \mathbf{V}_{\mathbf{M}_k}$$

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

$$\begin{bmatrix} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle & \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 & \mathbf{0} \\ \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 & \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \end{bmatrix} \otimes \overline{\mathbf{V}}_\xi$$

for $s \in \mathbb{R}_+$, where we used that $\langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^+ = \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$. Indeed, by (4.1), we get

$$\begin{aligned} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \mathbf{1}, \mathbf{X}_k - \mathbf{m}_{\boldsymbol{\xi}} \mathbf{X}_{k-1} - \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \mathbf{1}, s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} \langle \mathbf{1}, \mathbf{X}_{\lfloor ns \rfloor} \rangle + \frac{ns - \lfloor ns \rfloor}{n} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \in \mathbb{R}_+ \end{aligned} \tag{5.4}$$

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

In order to check condition (i) of Theorem D.1, we need to prove that for each $T > 0$, as $n \to \infty$,

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{V}_{\mathbf{M}_k} - \int_0^t \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.5}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbf{V}_{\mathbf{M}_k} - \int_0^t \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \overline{\mathbf{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.6}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \mathbf{V}_{\mathbf{M}_k} - \int_0^t \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^3 \overline{\mathbf{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.7}$$

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

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.9}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{5.10}$$

First, we show (5.5). By (5.4), $\int_0^t \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \, \mathrm{d}s$ has the form

$$\frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k + \frac{nt - \lfloor nt \rfloor}{n^2} U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2}{2n^2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle.$$

Using Lemma B.1, we obtain

$$\mathbf{V}_{\mathbf{M}_k} = U_{k-1} \overline{\mathbf{V}}_{\boldsymbol{\xi}} + \tfrac{1}{2} V_{k-1} (\mathbf{V}_{\boldsymbol{\xi}_1} - \mathbf{V}_{\boldsymbol{\xi}_2}) + \mathbf{V}_{\boldsymbol{\varepsilon}}. \tag{5.11}$$

Thus, in order to show (5.5), it suffices to prove

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

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

as $n \to \infty$. Using (B.4) with $(\ell, i, j) = (2, 1, 1)$ and (B.5) with $(\ell, i, j) = (2, 1, 0)$, we have (5.12). Clearly, (5.13) follows from $|nt - \lfloor nt \rfloor| \leq 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, thus we conclude (5.5). The convergences (5.6) and (5.7) can be checked in a similar way.

Next, we turn to prove (5.8). By (5.11) and (B.4), we get

$$n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbf{V}_{\mathbf{M}_k} - \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\mathbf{V}}_{\boldsymbol{\xi}} \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.14}$$

as $n \to \infty$ for all $T > 0$. Using (5.6), in order to prove (5.8), it is sufficient to show that

$$n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbf{V}_{\mathbf{M}_k} - \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\mathbf{V}}_{\boldsymbol{\xi}} \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.15}$$

as $n \to \infty$ for all $T > 0$. By (5.11), $\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbf{V}_{\mathbf{M}_k}$ has the form

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 \overline{\mathbf{V}}_{\boldsymbol{\xi}} + \frac{1}{2} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^3 (\mathbf{V}_{\boldsymbol{\xi}_1} - \mathbf{V}_{\boldsymbol{\xi}_2}) + \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbf{V}_{\boldsymbol{\varepsilon}}.$$

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

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

hence (5.15) will follow from

$$n^{-3} \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right\| \xrightarrow{\mathbb{P}} 0, \tag{5.16}$$

as $n \to \infty$ for all $T > 0$. By the method of the proof of Lemma A.1, we obtain a decomposition of $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2$ as a sum of a martingale and some negligible terms, namely,

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 = \frac{1}{4\alpha\beta} \sum_{k=2}^{\lfloor nt \rfloor} \bigl[U_{k-1} V_{k-1}^2 - \mathbb{E}\bigl(U_{k-1} V_{k-1}^2 | \mathcal{F}_{k-2}\bigr)\bigr]$$

$$+ \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2 - \frac{(\alpha - \beta)^2}{4\alpha\beta} U_{\lfloor nt \rfloor - 1} V_{\lfloor nt \rfloor - 1}^2 + \mathrm{O}(n)$$

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

Using (B.6) with $(\ell, i, j) = (8, 1, 2)$ we have

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

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

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

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

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

as $n \to \infty$. Using (B.4) with $(\ell, i, j) = (2, 1, 1)$, $(\ell, i, j) = (4, 0, 2)$, $(\ell, i, j) = (2, 1, 0)$ and $(\ell, i, j) = (2, 0, 1)$, we have (5.17) and (5.18). By (B.5) with $(\ell, i, j) = (4, 1, 2)$ and by (B.5), we have (5.19). Thus, we conclude (5.8). Convergences (5.9) and (5.10) can be proved similarly.

Finally, we check condition (ii) of Theorem D.1, that is, the conditional Lindeberg condition

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\bigl(\|\mathbf{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\mathbf{Z}_k^{(n)}\| > \theta\}} | \mathcal{F}_{k-1}\bigr) \xrightarrow{\mathbb{P}} 0, \qquad \text{as } n \to \infty \tag{5.20}$$

for all $\theta > 0$ and $T > 0$. We have $\mathbb{E}(\|\mathbf{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\mathbf{Z}_k^{(n)}\| > \theta\}} | \mathcal{F}_{k-1}) \leq \theta^{-2} \mathbb{E}(\|\mathbf{Z}_k^{(n)}\|^4 | \mathcal{F}_{k-1})$ and

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

Hence, for all $\theta > 0$ and $T > 0$, we have

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

since $\mathbb{E}(\|\mathbf{M}_k\|^4) = \mathrm{O}(k^2)$, $\mathbb{E}(\|\mathbf{M}_k\|^4 U_{k-1}^4) \leq \sqrt{\mathbb{E}(\|\mathbf{M}_k\|^8)\mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)$ and $\mathbb{E}(\|\mathbf{M}_k\|^4 V_{k-1}^4) \leq \sqrt{\mathbb{E}(\|\mathbf{M}_k\|^8)\mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)$ by Corollary B.6. Here we call the attention 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 B.6. This yields (5.20). □

6. Proof of Theorem 4.2

This is similar to the proof of Theorem 4.1. Consider the sequence of stochastic processes

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

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$. Theorem 4.2 follows from Lemma A.1 and the following theorem (this will be explained after Theorem 6.1).

Theorem 6.1. *If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ then*

$$\mathcal{Z}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{Z}, \qquad \text{as } n \to \infty, \tag{6.1}$$

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

$$\mathrm{d}\mathcal{Z}_t = \gamma(t, \mathcal{Z}_t) \begin{bmatrix} \mathrm{d}\mathcal{W}_t \\ \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{6.2}$$

with initial value $\mathcal{Z}_0 = \mathbf{0}$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$, $(\widetilde{\widetilde{\mathcal{W}}}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent standard Wiener processes of dimension 2, 1 and 2, respectively, and $\gamma(t, \mathbf{x})$ is a block diagonal matrix with the matrices $\langle \mathbf{1}, (\mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2}$, $\langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t$ and $(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta})^{1/2} \langle \mathbf{1}, \mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2}$ in its diagonal for each $t \in \mathbb{R}_+$ and $\mathbf{x} = (\mathbf{x}_1, x_2, \mathbf{x}_3) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2$.

As in the case of Theorem 4.1, the SDE (6.2) has a unique strong solution with initial value $\mathcal{Z}_0 = \mathbf{0}$, for which we have

$$\mathcal{Z}_t = \begin{bmatrix} \mathcal{M}_t \\ \mathcal{N}_t \\ \mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \mathcal{Y}_t^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\mathcal{W}_s \\ \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^t s \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_s \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^t \mathcal{Y}_t \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\mathcal{W}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+,$$

where now $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ yields $\mathcal{Y}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t$, $t \in \mathbb{R}_+$. One can again easily derive

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathcal{Z}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix}, \qquad \text{as } n \to \infty, \tag{6.3}$$

where

$$\mathcal{X}_t^{(n)} := n^{-1} \mathbf{X}_{\lfloor nt \rfloor}, \qquad \mathcal{X}_t := \frac{1}{2} \langle \mathbf{1}, \mathcal{M}_t + t \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1} = \frac{t}{2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1},$$

for $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$, since $\mathcal{X}_t = \frac{1}{2} \mathcal{Y}_t \mathbf{1} = \frac{t}{2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}$, $t \in \mathbb{R}_+$. Next, similarly to the proof of (A.2), by Lemma C.3, convergence (6.3) and Lemma A.1 with $U_{k-1} = \langle \mathbf{1}, \mathbf{X}_{k-1} \rangle$ imply

$$\sum_{k=1}^n \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-3/2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \mathbf{1}, \mathcal{X}_t \rangle^2 \, \mathrm{d}t \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right) \int_0^1 \langle \mathbf{1}, \mathcal{X}_t \rangle \, \mathrm{d}t \\ \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^1 t \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ \left(\frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d} \langle \widetilde{\mathbf{u}}, \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\mathcal{W}}_t \rangle \end{bmatrix},$$

as $n \to \infty$. This limiting random vector can be written in the form as given in Theorem 4.2 since $\langle \mathbf{1}, \mathcal{X}_t \rangle = \mathcal{Y}_t = \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle t$, and $\langle \widetilde{\mathbf{u}}, \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\mathcal{W}}_t \rangle = \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle^{1/2} \widetilde{\mathcal{W}}_t$ for all $t \in \mathbb{R}_+$ with a (one-dimensional) standard Wiener process $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$.

Proof of Theorem 6.1. Similar to the proof of Theorem 5.1. The conditional variance $\mathbb{E}(\mathbf{Z}_k^{(n)} (\mathbf{Z}_k^{(n)})^\top | \mathcal{F}_{k-1})$ has the form

$$\begin{bmatrix} n^{-2} \mathbf{V}_{\mathbf{M}_k} & n^{-5/2} U_{k-1} \mathbf{V}_{\mathbf{M}_k} \mathbf{1} & n^{-5/2} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \\ n^{-5/2} U_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} & n^{-3} U_{k-1}^2 \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \mathbf{1} & n^{-3} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \\ n^{-5/2} V_{k-1} \mathbf{V}_{\mathbf{M}_k} & n^{-3} U_{k-1} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \mathbf{1} & n^{-3} V_{k-1}^2 \mathbf{V}_{\mathbf{M}_k} \end{bmatrix}$$

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

$$\begin{bmatrix} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\xi} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 s^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \overline{\mathbf{V}}_{\xi} \end{bmatrix}$$

for $s \in \mathbb{R}_+$.

In order to check condition (i) of Theorem D.1, we need to prove only that for each $T > 0$,

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \right\| \xrightarrow{\mathbb{P}} 0, \tag{6.4}$$

$$\sup_{t \in [0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \mathbf{1} - \int_0^t \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 s^2 \, \mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0, \tag{6.5}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \right\| \xrightarrow{\mathbb{P}} 0, \tag{6.6}$$

as $n \to \infty$, since the rest, namely, (5.5), (5.8) and (5.9) have already been proved.

Clearly, $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ implies $\langle \mathbf{V}_{\boldsymbol{\xi}_1} \mathbf{1}, \mathbf{1} \rangle = 0$ and $\langle \mathbf{V}_{\boldsymbol{\xi}_2} \mathbf{1}, \mathbf{1} \rangle = 0$. For each $i \in \{1, 2\}$, we have $\langle \mathbf{V}_{\boldsymbol{\xi}_i} \mathbf{1}, \mathbf{1} \rangle = \mathbf{1}^\top \mathbf{V}_{\boldsymbol{\xi}_i} \mathbf{1} = (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1})^\top (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}) = \|\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}\|^2$, hence we obtain $\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1} = \mathbf{0}$, thus $\mathbf{V}_{\boldsymbol{\xi}_i} \mathbf{1} = \mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}) = \mathbf{0}$, and hence $\mathbf{1}^\top \mathbf{V}_{\boldsymbol{\xi}_i} = \mathbf{0}$, implying also $\mathbf{1}^\top \overline{\mathbf{V}}_{\boldsymbol{\xi}} = \mathbf{0}$.

First we show (6.4). By (5.11), $\mathbf{1}^\top \overline{\mathbf{V}}_{\boldsymbol{\xi}} = \mathbf{0}$ and $\mathbf{1}^\top \mathbf{V}_{\boldsymbol{\xi}_i} = \mathbf{0}$ for $i \in \{1, 2\}$, we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbf{1}^\top \mathbf{V}_{\boldsymbol{\varepsilon}}, \tag{6.7}$$

hence using (B.4) with $(\ell, i, j) = (2, 1, 0)$, we conclude (6.4).

Now we turn to check (6.5). By (5.11),

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} \mathbf{1} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \mathbf{1}^\top \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle,$$

hence, in order to show (6.5), it suffices to prove

$$\sup_{t \in [0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 - \frac{t^3}{3} \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \right| \xrightarrow{\mathbb{P}} 0, \qquad \text{as } n \to \infty. \tag{6.8}$$

We have

$$\left|\frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 - \frac{t^3}{3}\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2\right| \leq \frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor}\left|U_{k-1}^2 - (k-1)^2\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2\right| \\ + \left|\frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor}(k-1)^2 - \frac{t^3}{3}\right|\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2,$$

where

$$\sup_{t\in[0,T]}\left|\frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor}(k-1)^2 - \frac{t^3}{3}\right| \to 0, \qquad \text{as } n\to\infty,$$

hence, in order to show (6.5), it suffices to prove

$$\frac{1}{n^3}\sum_{k=1}^{\lfloor nT\rfloor}\left|U_k^2 - k^2\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2\right| \xrightarrow{\mathbb{P}} 0, \qquad \text{as } n\to\infty. \tag{6.9}$$

For all $k\in\mathbb{N}$, by Remark 3.3, $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}}\mathbf{1}, \mathbf{1}\rangle = 0$ implies

$$\begin{aligned} U_k &= \sum_{j=1}^{X_{k-1,1}}(\xi_{k,j,1,1}+\xi_{k,j,1,2}) + \sum_{j=1}^{X_{k-1,2}}(\xi_{k,j,2,1}+\xi_{k,j,2,2}) + (\varepsilon_{k,1}+\varepsilon_{k,2}) \\ &\overset{\text{a.s.}}{=} X_{k-1,1}+X_{k-1,2}+\varepsilon_{k,1}+\varepsilon_{k,2} = U_{k-1} + \langle \mathbf{1}, \boldsymbol{\varepsilon}_k\rangle, \end{aligned}$$

hence $U_k = \sum_{i=1}^k \langle \mathbf{1}, \boldsymbol{\varepsilon}_i\rangle$. By Kolmogorov's maximal inequality,

$$\mathbb{P}\Big(n^{-1}\max_{k\in\{1,\ldots,\lfloor nT\rfloor\}}\left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right| \geq \varepsilon\Big) \leq n^{-2}\varepsilon^{-2}\operatorname{Var}(U_{\lfloor nT\rfloor}) \\ = \frac{\lfloor nT\rfloor}{n^2\varepsilon^2}\operatorname{Var}\big(\langle \mathbf{1}, \boldsymbol{\varepsilon}_1\rangle^2\big) \to 0$$

as $n\to\infty$ for all $\varepsilon>0$, thus

$$n^{-1}\max_{k\in\{1,\ldots,\lfloor nT\rfloor\}}\left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right| \xrightarrow{\mathbb{P}} 0, \qquad \text{as } n\to\infty.$$

We have

$$\left|U_k^2 - k^2\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2\right| \leq \left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right|^2 + 2k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right|,$$

hence

$$n^{-2}\max_{k\in\{1,\ldots,\lfloor nT\rfloor\}}\left|U_k^2 - k^2\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle^2\right| \leq \Big(n^{-1}\max_{k\in\{1,\ldots,\lfloor nT\rfloor\}}\left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right|\Big)^2 \\ + \frac{2\lfloor nT\rfloor}{n^2}\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\max_{k\in\{1,\ldots,\lfloor nT\rfloor\}}\left|U_k - k\langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}}\rangle\right| \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$. Consequently,

$$\frac{1}{n^3} \sum_{k=1}^{\lfloor nT \rfloor} \big| U_{k-1}^2 - (k-1)^2 \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| \\ \leq \frac{\lfloor nT \rfloor}{n^3} \max_{k \in \{1, \ldots, \lfloor nT \rfloor\}} \big| U_{k-1}^2 - (k-1)^2 \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$, thus we conclude (6.9), and hence (6.5).

Finally, we check (6.6). By (5.11),

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbf{V}_{\mathbf{M}_k} = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbf{V}_{\boldsymbol{\varepsilon}},$$

hence using (B.4) with $(\ell, i, j) = (2, 1, 1)$, we conclude (6.6). Condition (ii) of Theorem D.1 can be checked as in case of Theorem 5.1. □

7. Proof of Theorem 4.3

This proof is also similar to the proof of Theorem 4.1. Consider the sequence of stochastic processes

$$\mathcal{Z}_t^{(n)} := \begin{bmatrix} \mathcal{M}_t^{(n)} \\ \mathcal{N}_t^{(n)} \\ \mathcal{P}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{Z}_k^{(n)} \qquad \text{with } \mathbf{Z}_k^{(n)} := \begin{bmatrix} n^{-1} \mathbf{M}_k \\ n^{-2} \mathbf{M}_k U_{k-1} \\ n^{-1/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix}$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$. Theorem 4.3 follows from Lemma A.2 and the following theorem (this will be explained after Theorem 7.1).

Theorem 7.1. *If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ then*

$$\mathcal{Z}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{Z}, \qquad \text{as } n \to \infty, \tag{7.1}$$

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

$$\mathrm{d}\mathcal{Z}_t = \gamma(t, \mathcal{Z}_t) \begin{bmatrix} \mathrm{d}\mathcal{W}_t \\ \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{7.2}$$

with initial value $\mathcal{Z}_0 = \mathbf{0}$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent standard Wiener processes of dimension 2 and 1, respectively, and $\gamma : \mathbb{R}_+ \times (\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}) \to \mathbb{R}^{5 \times 3}$ is defined by

$$\gamma(t, \mathbf{x}) := \begin{bmatrix} \langle \mathbf{1}, (\mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} & \mathbf{0} \\ \langle \mathbf{1}, (\mathbf{x}_1 + t\mathbf{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{3/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} & \mathbf{0} \\ \mathbf{0} & \big[\langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2 \big) \big]^{1/2} \end{bmatrix}$$

for $t \in \mathbb{R}_+$ *and* $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, x_3) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$.

As in the case of Theorem 4.1, the SDE (7.2) has a unique strong solution with initial value $\mathcal{Z}_0 = \mathbf{0}$, for which we have

$$\mathcal{Z}_t = \begin{bmatrix} \mathcal{M}_t \\ \mathcal{N}_t \\ \mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \mathcal{Y}_s^{1/2} \overline{\mathbf{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\mathcal{W}_s \\ \int_0^t \mathcal{Y}_s \, \mathrm{d}\mathcal{M}_s \\ \big[\langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big)\big]^{1/2} \widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

One can again easily derive

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathcal{Z}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix}, \qquad \text{as } n \to \infty, \tag{7.3}$$

where

$$\mathcal{X}_t^{(n)} := n^{-1} \mathbf{X}_{\lfloor nt \rfloor}, \qquad \mathcal{X}_t := \tfrac{1}{2} \langle \mathbf{1}, \mathcal{M}_t + t\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}, \qquad t \in \mathbb{R}_+, \ n \in \mathbb{N}.$$

Next, similarly to the proof of (A.2), by Lemma C.3, convergence (7.3) and Lemma A.2 imply

$$\sum_{k=1}^n \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-1} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \mathbf{M}_k \rangle U_{k-1} \\ n^{-1/2} \langle \widetilde{\mathbf{u}}, \mathbf{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \mathbf{1}, \mathcal{X}_t \rangle^2 \, \mathrm{d}t \\ \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\langle \mathbf{1}, \mathcal{M}_t \rangle \\ \big[\langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big)\big]^{1/2} \widetilde{\mathcal{W}}_1 \end{bmatrix},$$

as $n \to \infty$. Note that this convergence holds even in case $\mathbb{E}[\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2] = 0$. The limiting random vector can be written in the form as given in Theorem 4.3, since $\langle \mathbf{1}, \mathcal{X}_t \rangle = \mathcal{Y}_t$ and $\langle \mathbf{1}, \mathcal{M}_t \rangle = \mathcal{Y}_t - \langle \mathbf{1}, \mathbf{m}_{\boldsymbol{\varepsilon}} t \rangle$ for all $t \in \mathbb{R}_+$.

Proof of Theorem 7.1. Similar to the proof of Theorem 5.1. The conditional variance $\mathbb{E}(\mathbf{Z}_k^{(n)} (\mathbf{Z}_k^{(n)})^\top | \mathcal{F}_{k-1})$ has the form

$$\begin{bmatrix} n^{-2} \mathbf{V}_{\mathbf{M}_k} & n^{-3} U_{k-1} \mathbf{V}_{\mathbf{M}_k} & n^{-3/2} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \widetilde{\mathbf{u}} \\ n^{-3} U_{k-1} \mathbf{V}_{\mathbf{M}_k} & n^{-4} U_{k-1}^2 \mathbf{V}_{\mathbf{M}_k} & n^{-5/2} U_{k-1} V_{k-1} \mathbf{V}_{\mathbf{M}_k} \widetilde{\mathbf{u}} \\ n^{-3/2} V_{k-1} \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} & n^{-5/2} U_{k-1} V_{k-1} \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} & n^{-1} V_{k-1}^2 \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} \widetilde{\mathbf{u}} \end{bmatrix}$$

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

$$\begin{bmatrix} \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\mathbf{V}}_{\xi} & \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \overline{\mathbf{V}}_{\xi} & \mathbf{0} \\ \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \overline{\mathbf{V}}_{\xi} & \langle \mathbf{1}, \mathcal{M}_s^{(n)} + s\mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^3 \overline{\mathbf{V}}_{\xi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \end{bmatrix}$$

for $s \in \mathbb{R}_+$.

In order to check condition (i) of Theorem D.1, we need to prove only that for each $T > 0$,

$$\sup_{t\in[0,T]} \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} \widetilde{\mathbf{u}} - t \langle \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \right| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{7.4}$$

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^{3/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{7.5}$$

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{7.6}$$

as $n \to \infty$, since the rest, namely, (5.5), (5.6) and (5.7), have already been proved.

Clearly, $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ implies $\langle \mathbf{V}_{\boldsymbol{\xi}_1} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ and $\langle \mathbf{V}_{\boldsymbol{\xi}_2} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$. For each $i \in \{1, 2\}$, we have $\langle \mathbf{V}_{\boldsymbol{\xi}_i} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = \widetilde{\mathbf{u}}^\top \mathbf{V}_{\boldsymbol{\xi}_i} \widetilde{\mathbf{u}} = (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \widetilde{\mathbf{u}})^\top (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \widetilde{\mathbf{u}}) = \|\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \widetilde{\mathbf{u}}\|^2$, hence we obtain $\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \widetilde{\mathbf{u}} = \mathbf{0}$, thus $\mathbf{V}_{\boldsymbol{\xi}_i} \widetilde{\mathbf{u}} = \mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} (\mathbf{V}_{\boldsymbol{\xi}_i}^{1/2} \widetilde{\mathbf{u}}) = \mathbf{0}$, and hence $\widetilde{\mathbf{u}}^\top \mathbf{V}_{\boldsymbol{\xi}_i} = \mathbf{0}$.

First, we show (7.4). By (5.11),

$$\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \widetilde{\mathbf{u}}^\top \mathbf{V}_{\mathbf{M}_k} \widetilde{\mathbf{u}} = \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \widetilde{\mathbf{u}}^\top \mathbf{V}_{\boldsymbol{\varepsilon}} \widetilde{\mathbf{u}},$$

hence, in order to show (7.4), it suffices to prove

$$\sup_{t\in[0,T]} \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 - t \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \right| \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

For all $k \in \mathbb{N}$, by Remark 3.3, $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ implies

$$\begin{aligned} V_k &= \sum_{j=1}^{X_{k-1,1}} (\xi_{k,j,1,1} - \xi_{k,j,1,2}) + \sum_{j=1}^{X_{k-1,2}} (\xi_{k,j,2,1} - \xi_{k,j,2,2}) + (\varepsilon_{k,1} - \varepsilon_{k,2}) \\ &\stackrel{\text{a.s.}}{=} \varepsilon_{k,1} - \varepsilon_{k,2} = \langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_k \rangle. \end{aligned}$$

We have

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 - t \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big) \right| &\leq \frac{1}{n} \left| \sum_{k=1}^{\lfloor nt \rfloor} \big[\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_{k-1} \rangle^2 - \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_{k-1} \rangle^2\big) \big] \right| \\ &\quad + \frac{|nt - \lfloor nt \rfloor|}{n} \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_k \rangle^2\big), \end{aligned}$$

where $|nt - \lfloor nt \rfloor| \leq 1$, hence, in order to show (7.4), it suffices to prove

$$\begin{aligned}&\frac{1}{n}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt\rfloor}\big[\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2-\mathbb{E}\big(\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2\big)\big]\right|\\&\quad=\frac{1}{n}\max_{N\in\{1,\dots,\lfloor nT\rfloor\}}\left|\sum_{k=1}^{N}\big[\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2-\mathbb{E}\big(\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2\big)\big]\right|\stackrel{\mathbb{P}}{\longrightarrow}0.\end{aligned}\tag{7.7}$$

Applying Kolmogorov's maximal inequality, we obtain

$$\begin{aligned}&\mathbb{P}\left(n^{-1}\max_{N\in\{1,\dots,\lfloor nT\rfloor\}}\left|\sum_{k=1}^{N}\big[\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2-\mathbb{E}\big(\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2\big)\big]\right|\geq\varepsilon\right)\\&\leq\frac{1}{n^2\varepsilon^2}\operatorname{Var}\left(\sum_{k=1}^{\lfloor nT\rfloor}\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2\right)=\frac{\lfloor nT\rfloor}{n^2\varepsilon^2}\operatorname{Var}\big(\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle^2\big)\to0,\qquad\text{as } n\to\infty\end{aligned}$$

for all $\varepsilon > 0$, thus we conclude (7.7), and hence (7.4).

Now we turn to check (7.5). By (5.11),

$$\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}\widetilde{\mathbf{u}}^\top\mathbf{V}_{\mathbf{M}_k}=\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}\widetilde{\mathbf{u}}^\top\mathbf{V}_{\boldsymbol{\varepsilon}}.$$

Again by the strong law of large numbers, $n^{-1}\sum_{k=1}^{\lfloor nT\rfloor}|V_{k-1}| \xrightarrow{\text{a.s.}} t\mathbb{E}(|\langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_1\rangle|)$ as $n \to \infty$ for all $T > 0$, hence we conclude (7.5).

Finally, we check (7.6). By (5.11),

$$\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}V_{k-1}\widetilde{\mathbf{u}}^\top\mathbb{E}\big(\mathbf{M}_k\mathbf{M}_k^\top|\mathcal{F}_{k-1}\big)=\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}V_{k-1}\widetilde{\mathbf{u}}^\top\mathbf{V}_{\boldsymbol{\varepsilon}}.$$

Applying $V_k = \langle\widetilde{\mathbf{u}},\boldsymbol{\varepsilon}_k\rangle$, $k \in \mathbb{N}$, and Corollary B.6, we have $\mathbb{E}(|U_{k-1}V_{k-1}|) \leq \sqrt{\mathbb{E}(U_{k-1}^2)\mathbb{E}(V_{k-1}^2)} = \mathrm{O}(k)$, which clearly implies (7.6). Condition (ii) of Theorem D.1 can be checked again as in case of Theorem 5.1. □

Appendix A: CLS estimators

In order to analyse existence and uniqueness of the estimators given in (3.1), (3.2) and (3.5) in case of a critical doubly symmetric 2-type Galton–Watson process, that is, when $\varrho = 1$, we need the following approximations.

Lemma A.1. *We have*

$$n^{-2}\left(\sum_{k=1}^{n} V_k^2 - \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \sum_{k=1}^{n} U_{k-1}\right) \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad \text{as } n \to \infty.$$

Proof. In order to prove the statement, we derive a decomposition of $\sum_{k=1}^{n} V_k^2$ as a sum of a martingale and some negligible terms. Using recursion (4.4), Lemma B.1 and (4.5), we obtain

$$\begin{aligned}
\mathbb{E}\big(V_k^2 | \mathcal{F}_{k-1}\big) &= (\alpha-\beta)^2 V_{k-1}^2 + 2(\alpha-\beta)\langle \widetilde{\mathbf{u}}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle V_{k-1} + \langle \widetilde{\mathbf{u}}, \mathbf{m}_{\boldsymbol{\varepsilon}} \rangle^2 \\
&\quad + \widetilde{\mathbf{u}}^\top \mathbb{E}\big(\mathbf{M}_k \mathbf{M}_k^\top | \mathcal{F}_{k-1}\big) \widetilde{\mathbf{u}} \\
&= (\alpha-\beta)^2 V_{k-1}^2 + \tfrac{1}{2} \widetilde{\mathbf{u}}^\top (\mathbf{V}_{\boldsymbol{\xi}_1} + \mathbf{V}_{\boldsymbol{\xi}_2}) \widetilde{\mathbf{u}} U_{k-1} + \text{constant} + \text{constant} \times V_{k-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{k=1}^{n} V_k^2 &= \sum_{k=1}^{n} \big[V_k^2 - \mathbb{E}\big(V_k^2 | \mathcal{F}_{k-1}\big)\big] + (\alpha-\beta)^2 \sum_{k=1}^{n} V_{k-1}^2 + \widetilde{\mathbf{u}}^\top \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}} \sum_{k=1}^{n} U_{k-1} \\
&\quad + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{n} V_{k-1}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\sum_{k=1}^{n} V_k^2 &= \frac{1}{1-(\alpha-\beta)^2} \sum_{k=1}^{n} \big[V_k^2 - \mathbb{E}\big(V_k^2 | \mathcal{F}_{k-1}\big)\big] \\
&\quad + \frac{1}{1-(\alpha-\beta)^2} \langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle \sum_{k=1}^{n} U_{k-1} \\
&\quad - \frac{(\alpha-\beta)^2}{1-(\alpha-\beta)^2} V_n^2 + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{n} V_{k-1}.
\end{aligned} \tag{A.1}$$

Using (B.6) with $(\ell, i, j) = (8, 0, 2)$, we obtain

$$\frac{1}{n^2} \sum_{k=1}^{n} \big[V_k^2 - \mathbb{E}\big(V_k^2 | \mathcal{F}_{k-1}\big)\big] \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad \text{as } n \to \infty.$$

By Corollary B.6, we obtain $\mathbb{E}(V_n^2) = \mathrm{O}(n)$, and hence $n^{-2} V_n^2 \stackrel{\mathbb{P}}{\longrightarrow} 0$. Moreover, $n^{-2} \times \sum_{k=1}^{n} V_{k-1} \stackrel{\mathbb{P}}{\longrightarrow} 0$ as $n \to \infty$ follows by (B.4) with the choices $(\ell, i, j) = (4, 0, 1)$. Consequently, by (A.1), we obtain the statement, since $1 - (\alpha-\beta)^2 = 4\alpha\beta$. $\square$

Lemma A.2. *If* $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$, *then*

$$n^{-1} \sum_{k=1}^{n} V_k^2 \xrightarrow{\text{a.s.}} \mathbb{E}\big(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2\big), \qquad \text{as } n \to \infty,$$

and $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) = 0$ *if and only if* $X_{k,1} \stackrel{\text{a.s.}}{=} X_{k,2}$ *for all* $k \in \mathbb{N}$.

Proof. By Remark 3.3, $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ implies $V_k \stackrel{\text{a.s.}}{=} \varepsilon_{k,1} - \varepsilon_{k,2} = \langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle$ for all $k \in \mathbb{N}$, hence the convergence follows from the strong law of large numbers. Clearly $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) = 0$ is equivalent to $\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1 \rangle = \varepsilon_{1,1} - \varepsilon_{1,2} \stackrel{\text{a.s.}}{=} 0$, and hence it is equivalent to $X_{k,1} - X_{k,2} \stackrel{\text{a.s.}}{=} 0$ for all $k \in \mathbb{N}$. □

Now we can prove existence and uniqueness of CLS estimators of the offspring means and of the criticality parameter.

Proposition A.3. *We have* $\lim_{n\to\infty} \mathbb{P}((\mathbf{X}_1, \ldots, \mathbf{X}_n) \in H_n) = 1$, *where* H_n *is defined in* (3.3), *and hence the probability of the existence of a unique CLS estimator* $\widehat{\varrho}_n$ *converges to* 1 *as* $n \to \infty$, *and this CLS estimator has the form given in* (3.1) *whenever the sample* $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ *belongs to the set* H_n.

If $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle > 0$, *or if* $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle = 0$ *and* $\mathbb{E}(\langle \widetilde{\mathbf{u}}, \boldsymbol{\varepsilon} \rangle^2) > 0$, *then* $\lim_{n\to\infty} \mathbb{P}((\mathbf{X}_1, \ldots, \mathbf{X}_n) \in \widetilde{H}_n) = 1$, *where* $\widetilde{H}_n$ *is defined in* (3.4), *and hence the probability of the existence of unique CLS estimators* $\widehat{\delta}_n$ *and* $(\widehat{\alpha}_n, \widehat{\beta}_n)$ *converges to* 1 *as* $n \to \infty$. *The CLS estimator* $\widehat{\delta}_n$ *has the form given in* (3.2) *whenever the sample* $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ *belongs to the set* $\widetilde{H}_n$. *The CLS estimator* $(\widehat{\alpha}_n, \widehat{\beta}_n)$ *has the form given in* (3.5) *whenever the sample* $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ *belongs to the set* $H_n \cap \widetilde{H}_n$.

Proof. Recall convergence $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X} = \frac{1}{2}\mathcal{Y}\mathbf{1}$ from (3.11). By Lemmas C.2 and C.3 one can show

$$\frac{1}{n^3} \sum_{k=1}^{n} \big(X_{k-1,1}^2 + X_{k-1,2}^2\big) \xrightarrow{\mathcal{D}} \frac{1}{2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t, \qquad \text{as } n \to \infty, \tag{A.2}$$

see Ispány *et al.* [8], Proposition A.4. Since $\mathbf{m}_{\boldsymbol{\varepsilon}} \neq \mathbf{0}$, by the SDE (3.6), we have $\mathbb{P}(\mathcal{Y}_t = 0, t \in [0, 1]) = 0$, which implies that $\mathbb{P}(\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t > 0) = 1$. Consequently, the distribution function of $\int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t$ is continuous at 0, and hence, by (A.2),

$$\mathbb{P}\left(\sum_{k=1}^{n} \langle \mathbf{1}, \mathbf{X}_{k-1} \rangle^2 > 0\right) \to \mathbb{P}\left(\frac{1}{2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t > 0\right) = 1, \qquad \text{as } n \to \infty.$$

Now suppose that $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle > 0$ holds. In a similar way, using Lemma A.1, convergence (3.11), and Lemmas C.2 and C.3, one can show

$$\frac{1}{n^2} \sum_{k=1}^{n} \langle \widetilde{\mathbf{u}}, \mathbf{X}_{k-1} \rangle^2 \xrightarrow{\mathcal{D}} \frac{\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}} \rangle}{4\alpha\beta} \int_0^1 \mathcal{Y}_t \, \mathrm{d}t, \qquad \text{as } n \to \infty,$$

implying

$$\mathbb{P}\left(\sum_{k=1}^{n}\langle\widetilde{\mathbf{u}}, \mathbf{X}_{k-1}\rangle^2 > 0\right) \to \mathbb{P}\left(\int_0^1 \mathcal{Y}_t \, \mathrm{d}t > 0\right) = 1, \qquad \text{as } n \to \infty,$$

hence we obtain the statement under the assumption $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle > 0$.

Next, we suppose that $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ and $\mathbb{E}(\langle\widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}\rangle^2) > 0$ hold. Then

$$\mathbb{P}\left(\sum_{k=1}^{n}\langle\widetilde{\mathbf{u}}, \mathbf{X}_{k-1}\rangle^2 > 0\right) = \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} V_{k-1}^2 > 0\right) \to 1, \qquad \text{as } n \to \infty,$$

since Lemma A.2 yields $n^{-1}\sum_{k=1}^n V_{k-1}^2 \xrightarrow{\mathbb{P}} \mathbb{E}(\langle\widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}_1\rangle^2) > 0$, and hence we conclude the statement under the assumptions $\langle \overline{\mathbf{V}}_{\boldsymbol{\xi}} \widetilde{\mathbf{u}}, \widetilde{\mathbf{u}}\rangle = 0$ and $\mathbb{E}(\langle\widetilde{\mathbf{u}}, \boldsymbol{\varepsilon}\rangle^2) > 0$. □

Appendix B: Estimations of moments

In the proof of Theorem 3.1, good bounds for moments of the random vectors and variables $(\mathbf{M}_k)_{k\in\mathbb{Z}_+}$, $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$, $(U_k)_{k\in\mathbb{Z}_+}$ and $(V_k)_{k\in\mathbb{Z}_+}$ are extensively used. First note that, for all $k \in \mathbb{N}$, $\mathbb{E}(\mathbf{M}_k | \mathcal{F}_{k-1}) = \mathbf{0}$ and $\mathbb{E}(\mathbf{M}_k) = \mathbf{0}$, since $\mathbf{M}_k = \mathbf{X}_k - \mathbb{E}(\mathbf{X}_k | \mathcal{F}_{k-1})$.

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

$$\mathbb{E}\big(\mathbf{M}_k\mathbf{M}_k^\top | \mathcal{F}_{k-1}\big) = X_{k-1,1}\mathbf{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\mathbf{V}_{\boldsymbol{\xi}_2} + \mathbf{V}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}. \tag{B.1}$$

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

$$\begin{aligned}\mathbb{E}\big(\mathbf{M}_k^{\otimes 3} | \mathcal{F}_{k-1}\big) = {} & X_{k-1,1}\mathbb{E}\big[(\boldsymbol{\xi}_{1,1,1} - \mathbb{E}(\boldsymbol{\xi}_{1,1,1})^{\otimes 3}\big] \\ & + X_{k-1,2}\mathbb{E}\big[(\boldsymbol{\xi}_{1,1,2} - \mathbb{E}(\boldsymbol{\xi}_{1,1,2})^{\otimes 3}\big] + \mathbb{E}\big[(\boldsymbol{\varepsilon}_1 - \mathbb{E}(\boldsymbol{\varepsilon}_1)^{\otimes 3}\big], \qquad k \in \mathbb{N}.\end{aligned} \tag{B.2}$$

Proof. By (2.1) and (4.1), $\mathbf{M}_k$ has the form

$$\sum_{j=1}^{X_{k-1,1}} \big(\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1})\big) + \sum_{j=1}^{X_{k-1,2}} \big(\boldsymbol{\xi}_{k,j,2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2})\big) + \big(\boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k)\big) \tag{B.3}$$

for all $k \in \mathbb{N}$. The random vectors $\{\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1}), \boldsymbol{\xi}_{k,j.2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2}), \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k) : j \in \mathbb{N}\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean vector, thus we conclude (B.1) and (B.2). □

Lemma B.2. *Let $(\boldsymbol{\zeta}_k)_{k\in\mathbb{N}}$ be independent and identically distributed random vectors with values in $\mathbb{R}^d$ such that $\mathbb{E}(\|\boldsymbol{\zeta}_1\|^\ell) < \infty$ with some $\ell \in \mathbb{N}$.*

(i) *Then there exists* $\mathbf{Q} = (Q_1, \dots, Q_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, *where* $Q_1, \dots, Q_{d^\ell}$ *are polynomials having degree at most* $\ell - 1$ *such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_1 + \cdots + \boldsymbol{\zeta}_N)^{\otimes \ell}\big) = N^\ell \big[\mathbb{E}(\boldsymbol{\zeta}_1)\big]^{\otimes \ell} + \mathbf{Q}(N), \qquad N \in \mathbb{N}, N \geq \ell.$$

(ii) *If* $\mathbb{E}(\boldsymbol{\zeta}_1) = \mathbf{0}$, *then there exists* $\mathbf{R} = (R_1, \dots, R_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, *where* $R_1, \dots, R_{d^\ell}$ *are polynomials having degree at most* $\lfloor \ell/2 \rfloor$ *such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_1 + \cdots + \boldsymbol{\zeta}_N)^{\otimes \ell}\big) = \mathbf{R}(N), \qquad N \in \mathbb{N}, N \geq \ell.$$

The coefficients of the polynomials $\mathbf{Q}$ *and* $\mathbf{R}$ *depend on the moments* $\mathbb{E}(\boldsymbol{\zeta}_{i_1} \otimes \cdots \otimes \boldsymbol{\zeta}_{i_\ell})$, $i_1, \dots, i_\ell \in \{1, \dots, N\}$.

Proof. (i) We have

$$\begin{aligned} &\mathbb{E}\big((\boldsymbol{\zeta}_1 + \cdots + \boldsymbol{\zeta}_N)^{\otimes \ell}\big) \\ &\quad = \sum_{\substack{s \in \{1,\dots,\ell\}, k_1,\dots,k_s \in \mathbb{Z}_+, \\ k_1 + 2k_2 + \cdots + s k_s = \ell, k_s \neq 0}} \binom{N}{k_1} \binom{N-k_1}{k_2} \cdots \binom{N-k_1-\cdots-k_{s-1}}{k_s} \\ &\qquad \times \sum_{(i_1,\dots,i_\ell) \in P^{(N,\ell)}_{k_1,\dots,k_s}} \mathbb{E}(\boldsymbol{\zeta}_{i_1} \otimes \cdots \otimes \boldsymbol{\zeta}_{i_\ell}), \end{aligned}$$

where the set $P^{(N,\ell)}_{k_1,\dots,k_s}$ consists of permutations of all the multisets containing pairwise different elements $j_{k_1}, \dots, j_{k_s}$ of the set $\{1, \dots, N\}$ with multiplicities $k_1, \dots, k_s$, respectively. Since

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

is a polynomial of the variable N having degree $k_1 + \cdots + k_s \leq \ell$, there exists $\mathbf{P} = (P_1, \dots, P_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, where $P_1, \dots, P_{d^\ell}$ are polynomials having degree at most ℓ such that $\mathbb{E}((\boldsymbol{\zeta}_1 + \cdots + \boldsymbol{\zeta}_N)^{\otimes \ell}) = \mathbf{P}(N)$. A term of degree ℓ can occur only in case $k_1 + \cdots + k_s = \ell$, when $k_1 + 2k_2 + \cdots + s k_s = \ell$ implies $s = 1$ and $k_1 = \ell$, thus the corresponding term of degree ℓ is $N(N-1)\cdots(N-\ell+1)[\mathbb{E}(\boldsymbol{\zeta}_1)]^{\otimes \ell}$, hence we obtain the statement. Part (ii) can be proved in a similar way. □

Lemma B.2 can be generalized in the following way.

Lemma B.3. *For each* $i \in \mathbb{N}$, *let* $(\boldsymbol{\zeta}_{i,k})_{k \in \mathbb{N}}$ *be independent and identically distributed random vectors with values in* $\mathbb{R}^d$ *such that* $\mathbb{E}(\|\boldsymbol{\zeta}_{i,1}\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Let* $j_1, \dots, j_\ell \in \mathbb{N}$.

(i) *Then there exists* $\mathbf{Q} = (Q_1, \dots, Q_{d^\ell}) : \mathbb{R}^\ell \to \mathbb{R}^{d^\ell}$, *where* $Q_1, \dots, Q_{d^\ell}$ *are polynomials of* ℓ *variables having degree at most* $\ell - 1$ *such that*

$$\begin{aligned}&\mathbb{E}\big((\boldsymbol{\zeta}_{j_1,1} + \dots + \boldsymbol{\zeta}_{j_1,N_1}) \otimes \dots \otimes (\boldsymbol{\zeta}_{j_\ell,1} + \dots + \boldsymbol{\zeta}_{j_\ell,N_\ell})\big)\\ &\quad = N_1 \cdots N_\ell \mathbb{E}(\boldsymbol{\zeta}_{j_1,1}) \otimes \dots \otimes \mathbb{E}(\boldsymbol{\zeta}_{j_\ell,1}) + \mathbf{Q}(N_1, \dots, N_\ell)\end{aligned}$$

for $N_1, \dots, N_\ell \in \mathbb{N}$ *with* $N_1 \geq \ell, \dots, N_\ell \geq \ell$.

(ii) *If* $\mathbb{E}(\boldsymbol{\zeta}_{j_1,1}) = \dots = \mathbb{E}(\boldsymbol{\zeta}_{j_\ell,1}) = \mathbf{0}$, *then there exists* $\mathbf{R} = (R_1, \dots, R_{d^\ell}) : \mathbb{R}^\ell \to \mathbb{R}^{d^\ell}$, *where* $R_1, \dots, R_{d^\ell}$ *are polynomials of* ℓ *variables having degree at most* $\lfloor \ell/2 \rfloor$ *such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_{j_1,1} + \dots + \boldsymbol{\zeta}_{j_1,N_1}) \otimes \dots \otimes (\boldsymbol{\zeta}_{j_\ell,1} + \dots + \boldsymbol{\zeta}_{j_\ell,N_\ell})\big) = \mathbf{R}(N_1, \dots, N_\ell)$$

for $N_1, \dots, N_\ell \in \mathbb{N}$ *with* $N_1 \geq \ell, \dots, N_\ell \geq \ell$.

The coefficients of the polynomials $\mathbf{Q}$ *and* $\mathbf{R}$ *depend on the moments* $\mathbb{E}(\boldsymbol{\zeta}_{j_1,i_1} \otimes \dots \otimes \boldsymbol{\zeta}_{j_\ell,i_\ell})$, $i_1 \in \{1, \dots, N_1\}, \dots, i_\ell \in \{1, \dots, N_\ell\}$.

Lemma B.4. *If* $(\alpha, \beta) \in [0, 1]$ *with* $\alpha + \beta = 1$, *then the matrix* $\mathbf{m}_{\boldsymbol{\xi}}$ *defined in* (2.4) *has eigenvalues* 1 *and* $\alpha - \beta$, *and the powers of* $\mathbf{m}_{\boldsymbol{\xi}}$ *take the form*

$$\mathbf{m}_{\boldsymbol{\xi}}^j = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2}(\alpha - \beta)^j \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad j \in \mathbb{Z}_+.$$

Consequently, $\|\mathbf{m}_{\boldsymbol{\xi}}^j\| = \mathrm{O}(1)$, *that is*, $\sup_{j \in \mathbb{N}} \|\mathbf{m}_{\boldsymbol{\xi}}^j\| < \infty$.

Lemma B.5. *Let* $(\mathbf{X}_k)_{k \in \mathbb{Z}_+}$ *be a* 2-*type doubly symmetric Galton–Watson process with immigration with offspring means* $(\alpha, \beta) \in [0, 1]$ *such that* $\alpha + \beta = 1$ (*hence it is critical*). *Suppose* $\mathbf{X}_0 = \mathbf{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then* $\mathbb{E}(\|\mathbf{X}_k\|^\ell) = \mathrm{O}(k^\ell)$, *that is*, $\sup_{k \in \mathbb{N}} k^{-\ell} \mathbb{E}(\|\mathbf{X}_k\|^\ell) < \infty$.

Proof. The statement is clearly equivalent with $\mathbb{E}(|P(X_{k,1}, X_{k,2})|) \leq c_P k^\ell$, $k \in \mathbb{N}$, for all polynomials P of two variables having degree at most ℓ, where c_P depends only on P.

If $\ell = 1$, then (2.3) and Lemma B.4 imply

$$\mathbb{E}(\mathbf{X}_k) = \sum_{j=0}^{k-1} \mathbf{m}_{\boldsymbol{\xi}}^j \mathbf{m}_{\boldsymbol{\varepsilon}} = \left(\frac{k}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1 - (\alpha - \beta)^k}{4\beta}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right)\mathbf{m}_{\boldsymbol{\varepsilon}},$$

for all $k \in \mathbb{N}$, which yields the statement.

Using part (i) of Lemma B.3 and separating the terms having degree 2 and less than 2, we obtain

$$\begin{aligned}\mathbb{E}\big(\mathbf{X}_k^{\otimes 2} | \mathcal{F}_{k-1}\big) &= X_{k-1,1}^2 \mathbf{m}_{\boldsymbol{\xi}_1}^{\otimes 2} + X_{k-1,2}^2 \mathbf{m}_{\boldsymbol{\xi}_2}^{\otimes 2} + X_{k-1,1} X_{k-1,2} (\mathbf{m}_{\boldsymbol{\xi}_1} \otimes \mathbf{m}_{\boldsymbol{\xi}_2} + \mathbf{m}_{\boldsymbol{\xi}_2} \otimes \mathbf{m}_{\boldsymbol{\xi}_1})\\ &\quad + \mathbf{Q}_2(X_{k-1,1}, X_{k-1,2})\end{aligned}$$

$$= (X_{k-1,1}\mathbf{m}_{\xi_1} + X_{k-1,2}\mathbf{m}_{\xi_2})^{\otimes 2} + \mathbf{Q}_2(X_{k-1,1}, X_{k-1,2})$$
$$= (\mathbf{m}_{\boldsymbol{\xi}}\mathbf{X}_{k-1})^{\otimes 2} + \mathbf{Q}_2(X_{k-1,1}, X_{k-1,2}) = \mathbf{m}_{\boldsymbol{\xi}}^{\otimes 2}\mathbf{X}_{k-1}^{\otimes 2} + \mathbf{Q}_2(X_{k-1,1}, X_{k-1,2}),$$

where $\mathbf{Q}_2 = (Q_{2,1}, Q_{2,2}, Q_{2,3}, Q_{2,4}) : \mathbb{R}^2 \to \mathbb{R}^4$, and $Q_{2,1}, Q_{2,2}, Q_{2,3}$ and $Q_{2,4}$ are polynomials of two variables having degree at most 1. Hence

$$\mathbb{E}(\mathbf{X}_k^{\otimes 2}) = \mathbf{m}_{\boldsymbol{\xi}}^{\otimes 2}\mathbb{E}(\mathbf{X}_{k-1}^{\otimes 2}) + \mathbb{E}[\mathbf{Q}_2(X_{k-1,1}, X_{k-1,2})].$$

In a similar way,

$$\mathbb{E}(\mathbf{X}_k^{\otimes \ell}) = \mathbf{m}_{\boldsymbol{\xi}}^{\otimes \ell}\mathbb{E}(\mathbf{X}_{k-1}^{\otimes \ell}) + \mathbb{E}[\mathbf{Q}_\ell(X_{k-1,1}, X_{k-1,2})],$$

where $\mathbf{Q}_\ell = (Q_{\ell,1}, \dots, Q_{\ell,2^\ell}) : \mathbb{R}^2 \to \mathbb{R}^{2^\ell}$, and $Q_{\ell,1}, \dots, Q_{\ell,2^\ell}$ are polynomials of two variables having degree at most $\ell - 1$, implying

$$\begin{aligned}\mathbb{E}(\mathbf{X}_k^{\otimes \ell}) &= \sum_{j=1}^{k} (\mathbf{m}_{\boldsymbol{\xi}}^{\otimes \ell})^{k-j}\mathbb{E}[\mathbf{Q}_\ell(X_{j-1,1}, X_{j-1,2})] \\ &= \sum_{j=0}^{k-1} (\mathbf{m}_{\boldsymbol{\xi}}^{\otimes \ell})^{j}\mathbb{E}[\mathbf{Q}_\ell(X_{k-j-1,1}, X_{k-j-1,2})] \\ &= \sum_{j=0}^{k-1} (\mathbf{m}_{\boldsymbol{\xi}}^{j})^{\otimes \ell}\mathbb{E}[\mathbf{Q}_\ell(X_{k-j-1,1}, X_{k-j-1,2})].\end{aligned}$$

Let us suppose now that the statement holds for $1, \dots, \ell - 1$. Then

$$\mathbb{E}[|Q_{\ell,i}(X_{k-j-1,1}, X_{k-j-1,2})|] \le c_{Q_{\ell,i}} k^{\ell-1}, \qquad k \in \mathbb{N},\ i \in \{1, \dots, 2^\ell\}.$$

By Lemma B.4 $\|(\mathbf{m}_{\boldsymbol{\xi}}^{j})^{\otimes \ell}\| = \mathrm{O}(1)$, hence we obtain the assertion for ℓ. □

Corollary B.6. *Let $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type doubly symmetric Galton–Watson process with immigration having offspring means $(\alpha, \beta) \in (0, 1)^2$ such that $\alpha + \beta = 1$ (hence it is critical and positively regular). Suppose $\mathbf{X}_0 = \mathbf{0}$, and $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ with some $\ell \in \mathbb{N}$. Then $\mathbb{E}(\|\mathbf{X}_k\|^\ell) = \mathrm{O}(k^\ell)$, $\mathbb{E}(\mathbf{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor})$, $\mathbb{E}(U_k^\ell) = \mathrm{O}(k^\ell)$ and $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$ for $j \in \mathbb{Z}_+$ with $2j \le \ell$.*

Proof. The first statement is just Lemma B.5. Next, we turn to prove $\mathbb{E}(\mathbf{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor})$. Using (B.3), part (ii) of Lemma B.3, and that the random vectors $\{\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1}), \boldsymbol{\xi}_{k,j.2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2}), \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k) : j \in \mathbb{N}\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean vector, we obtain $\mathbb{E}(\mathbf{M}_k^{\otimes \ell} | \mathcal{F}_{k-1}) = \mathbf{R}(X_{k-1,1}, X_{k-1,2})$ with $\mathbf{R} = (R_1, \dots, R_{2^\ell}) : \mathbb{R}^2 \to \mathbb{R}^{2\ell}$, where $R_1, \dots, R_{2^\ell}$ are polynomials of two variables having degree at most $\ell/2$. Hence $\mathbb{E}(\mathbf{M}_k^{\otimes \ell}) = \mathbb{E}(\mathbf{R}(X_{k-1,1}, X_{k-1,2}))$. By Lemma B.5, we conclude $\mathbb{E}(\mathbf{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor})$. The rest of the proof can be carried out as in Corollary 9.1 of Barczy *et al.* [2]. □

The next corollary can be derived as Corollary 9.2 of Barczy *et al.* [2].

Corollary B.7. *Let* $(\mathbf{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type doubly symmetric Galton–Watson process with immigration having offspring means* $(\alpha,\beta)\in(0,1)^2$ *such that* $\alpha+\beta=1$ *(hence, it is critical and positively regular). Suppose* $\mathbf{X}_0=\mathbf{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell)<\infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell)<\infty$ *with some* $\ell\in\mathbb{N}$. *Then*

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

$$n^{-\kappa}\sum_{k=1}^{n}\left|U_k^iV_k^j\right|\xrightarrow{\mathbb{P}}0,\qquad \text{as } n\to\infty,\tag{B.4}$$

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

$$n^{-\kappa}\sup_{t\in[0,T]}\left|U_{\lfloor nt\rfloor}^iV_{\lfloor nt\rfloor}^j\right|\xrightarrow{\mathbb{P}}0,\qquad \text{as } n\to\infty,\tag{B.5}$$

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

$$n^{-\kappa}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt\rfloor}\left[U_k^iV_k^j-\mathbb{E}\left(U_k^iV_k^j\,|\,\mathcal{F}_{k-1}\right)\right]\right|\xrightarrow{\mathbb{P}}0,\qquad \text{as } n\to\infty.\tag{B.6}$$

Remark B.8. In the special case $(\ell,i,j)=(2,1,0)$, one can improve (B.5), namely, one can show

$$n^{-\kappa}\sup_{t\in[0,T]}U_{\lfloor nt\rfloor}\xrightarrow{\mathbb{P}}0,\qquad \text{as } n\to\infty \text{ for } \kappa>1,\tag{B.7}$$

see Barczy *et al.* [2].

Appendix C: A version of the continuous mapping theorem

A function $f:\mathbb{R}_+\to\mathbb{R}^d$ is called *càdlàg* if it is right continuous with left limits. Let $\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d)$ and $\mathsf{C}(\mathbb{R}_+,\mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg and continuous functions on $\mathbb{R}_+$, respectively. Let $\mathcal{B}(\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d))$ denote the Borel σ-algebra on $\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d)$ for the metric defined in Jacod and Shiryaev [11], Chapter VI, (1.26) (with this metric $\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d)$ is a complete and separable metric space and the topology induced by this metric is the so-called Skorokhod topology). For $\mathbb{R}^d$-valued stochastic processes $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ and $(\mathcal{Y}_t^{(n)})_{t\in\mathbb{R}_+}$, $n\in\mathbb{N}$, with càdlàg paths, we write $\mathcal{Y}^{(n)}\xrightarrow{\mathcal{D}}\mathcal{Y}$ if the distribution of $\mathcal{Y}^{(n)}$ on the space $(\mathsf{D}(\mathbb{R}_+,\mathbb{R}),\mathcal{B}(\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d)))$ converges weakly to the distribution of $\mathcal{Y}$ on the space $(\mathsf{D}(\mathbb{R}_+,\mathbb{R}),\mathcal{B}(\mathsf{D}(\mathbb{R}_+,\mathbb{R}^d)))$ as $n\to\infty$. Concerning the notation $\xrightarrow{\mathcal{D}}$ we note that if ξ and ξ_n, $n\in\mathbb{N}$, are random elements with values in a metric space (E,d), then we also denote by $\xi_n\xrightarrow{\mathcal{D}}\xi$ 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 d.

The following version of continuous mapping theorem can be found, for example, in Kallenberg [12, Theorem 3.27].

Lemma C.1. *Let (S, d_S) and (T, d_T) be metric spaces and $(\xi_n)_{n\in\mathbb{N}}$, ξ be random elements with values in S such that $\xi_n \stackrel{\mathcal{D}}{\longrightarrow} \xi$ as $n \to \infty$. Let $f : S \to T$ and $f_n : S \to T$, $n \in \mathbb{N}$, be measurable mappings and $C \in \mathcal{B}(S)$ such that $\mathbb{P}(\xi \in C) = 1$ and $\lim_{n\to\infty} d_T(f_n(s_n), f(s)) = 0$ if $\lim_{n\to\infty} d_S(s_n, s) = 0$ and $s \in C$. Then $f_n(\xi_n) \stackrel{\mathcal{D}}{\longrightarrow} f(\xi)$, as $n \to \infty$.*

For the case $S = \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $T = \mathbb{R}^q$ (or $T = \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), where $d, q \in \mathbb{N}$, we formulate a consequence of Lemma C.1.

For functions f and f_n, $n \in \mathbb{N}$, in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, we write $f_n \stackrel{\text{lu}}{\longrightarrow} f$ if $(f_n)_{n\in\mathbb{N}}$ converges to f locally uniformly, that is, if $\sup_{t\in[0,T]} \|f_n(t) - f(t)\| \to 0$ as $n \to \infty$ for all $T > 0$. For measurable mappings $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, we will denote by $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ the set of all functions $f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that $\Phi_n(f_n) \to \Phi(f)$ (or $\Phi_n(f_n) \to\stackrel{\text{lu}}{\longrightarrow} \Phi(f)$) whenever $f_n \stackrel{\text{lu}}{\longrightarrow} f$ with $f_n \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$.

We will use the following version of the continuous mapping theorem several times, see, for example, Ispány and Pap [10], Lemma 3.1.

Lemma C.2. *Let $d, q \in \mathbb{N}$, and $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ and $(\mathcal{U}_t^{(n)})_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, be $\mathbb{R}^d$-valued stochastic processes with càdlàg paths such that $\mathcal{U}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \mathcal{U}$. Let $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, be measurable mappings such that there exists $C \subset C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ with $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ and $\mathbb{P}(\mathcal{U} \in C) = 1$. Then $\Phi_n(\mathcal{U}^{(n)}) \stackrel{\mathcal{D}}{\longrightarrow} \Phi(\mathcal{U})$.*

In order to apply Lemma C.2, we will use the following statement several times, see Barczy *et al.* [2], Lemma B.3.

Lemma C.3. *Let $d, p, q \in \mathbb{N}$, $h : \mathbb{R}^d \to \mathbb{R}^q$ be a continuous function and $K : [0, 1] \times \mathbb{R}^{2d} \to \mathbb{R}^p$ be a function such that for all $R > 0$ there exists $C_R > 0$ such that*

$$\big\| K(s, x) - K(t, y)\big\| \leq C_R\big(|t - s| + \|x - y\|\big) \tag{C.1}$$

for all $s, t \in [0, 1]$ and $x, y \in \mathbb{R}^{2d}$ with $\|x\| \leq R$ and $\|y\| \leq R$. Moreover, let us define the mappings $\Phi, \Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^{q+p}$, $n \in \mathbb{N}$, by

$$\Phi_n(f) := \left(h\big(f(1)\big), \frac{1}{n}\sum_{k=1}^{n} K\left(\frac{k}{n}, f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right)\right)\right),$$

$$\Phi(f) := \left(h\big(f(1)\big), \int_0^1 K\big(u, f(u), f(u)\big)\,\mathrm{d}u\right)$$

for all $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$. *Then the mappings* Φ *and* Φ_n, $n \in \mathbb{N}$, *are measurable*, *and* $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$.

Appendix D: Convergence of random step processes

We recall a result about convergence of random step processes towards a diffusion process, see Ispány and Pap [10]. This result is used for the proof of convergence (5.1).

Theorem D.1. *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*

$$\mathrm{d}\mathcal{U}_t = \boldsymbol{\gamma}(t, \mathcal{U}_t)\,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+, \tag{D.1}$$

with initial value $\mathcal{U}_0 = \mathbf{u}_0$ *for all* $\mathbf{u}_0 \in \mathbb{R}^d$, *where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *is an* r*-dimensional standard Wiener process. Let* $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ *be a solution of* (D.1) *with initial value* $\mathcal{U}_0 = \mathbf{0} \in \mathbb{R}^d$.

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

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

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

(i) $\sup_{t\in[0,T]} \| \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{E}(\mathbf{U}_k^{(n)} (\mathbf{U}_k^{(n)})^\top \mid \mathcal{F}_{k-1}^{(n)}) - \int_0^t \boldsymbol{\gamma}(s, \mathcal{U}_s^{(n)}) \boldsymbol{\gamma}(s, \mathcal{U}_s^{(n)})^\top \,\mathrm{d}s \| \xrightarrow{\mathbb{P}} 0$,

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

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

Note that in (i) of Theorem D.1, $\|\cdot\|$ denotes a matrix norm, while in (ii) it denotes a vector norm.

Acknowledgements

The authors have been partially supported by the Hungarian Chinese Intergovernmental S&T Cooperation Programme for 2011–2013 under Grant No. 10-1-2011-0079. M. Ispány has been partially supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund. K. Körmendi was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 "National Excellence Program". G. Pap has been partially supported by the Hungarian Scientific Research Fund under Grant No. OTKA T-079128.

References

[1] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. New York: Springer. MR0373040

[2] Barczy, M., Ispány, M. and Pap, G. (2012). Asymptotic behavior of CLS estimators for unstable INAR(2) models. Available at http://arxiv.org/abs/1202.1617.

[3] Barczy, M., Ispány, M. and Pap, G. (2011). Asymptotic behavior of unstable INAR(p) processes. *Stochastic Process*. *Appl*. **121** 583–608. MR2763097

[4] Guttorp, P. (1991). *Statistical Inference for Branching Processes*. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1254434

[5] Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71** 285–317. MR1956860

[6] Hamilton, J.D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton Univ. Press. MR1278033

[7] Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183

[8] Ispány, M., Körmendi, K. and Pap, G. (2012). Asymptotic behavior of CLS estimators for 2-type critical Galton–Watson processes with immigration. Available at http://arxiv.org/abs/1210.8315.

[9] Ispány, M. and Pap, G. (2012). Asymptotic behavior of critical primitive multi-type branching processes with immigration. Available at http://arxiv.org/abs/1205.0388.

[10] Ispány, M. and Pap, G. (2010). A note on weak convergence of random step processes. *Acta Math*. *Hungar*. **126** 381–395. MR2629664

[11] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Berlin: Springer. MR1943877

[12] Kallenberg, O. (1997). *Foundations of Modern Probability*. *Probability and Its Applications* (*New York*). New York: Springer. MR1464694

[13] Kesten, H. and Stigum, B.P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann*. *Math*. *Statist*. **37** 1211–1223. MR0198552

[14] Mikosch, T. and Straumann, D. (2002). Whittle estimation in a heavy-tailed GARCH(1, 1) model. *Stochastic Process*. *Appl*. **100** 187–222. MR1919613

[15] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*. *Applications of Mathematics* (*New York*) **36**. Berlin: Springer. MR1474500

[16] Quine, M.P. (1970). The multi-type Galton–Watson process with immigration. *J*. *Appl*. *Probability* **7** 411–422. MR0263168

[17] Revuz, D. and Yor, M. (2001). *Continuous Martingales and Brownian Motion*, 3rd ed., corrected 2nd printing. Berlin: Springer. MR1725357.

[18] Shete, S. and Sriram, T.N. (2003). A note on estimation in multitype supercritical branching processes with immigration. *Sankhyā* **65** 107–121. MR2016780

[19] Tanaka, K. (1996). *Time Series Analysis*: *Nonstationary and Noninvertible Distribution Theory*. *Wiley Series in Probability and Statistics*. New York: Wiley. MR1397269

[20] Wei, C.Z. and Winnicki, J. (1989). Some asymptotic results for the branching process with immigration. *Stochastic Process*. *Appl*. **31** 261–282. MR0998117

[21] Wei, C.Z. and Winnicki, J. (1990). Estimation of the means in the branching process with immigration. *Ann*. *Statist*. **18** 1757–1773. MR1074433

[22] Winnicki, J. (1991). Estimation of the variances in the branching process with immigration. *Probab*. *Theory Related Fields* **88** 77–106. MR1094078

Received December 2012 and revised August 2013