Asymptotic behavior of a multi-type nearly critical Galton–Watson processes with immigration

László Győrfi[1], Márton Ispány[2], Péter Kevei[3] and Gyula Pap[4]

[1]Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Stoczek u. 2, Budapest, Hungary, H-1521;
e-mail: gyorfi@szit.bme.hu,
[2]Department of Information Technology, Faculty of Informatics, University of Debrecen, Pf.12, Debrecen, Hungary, H-4010;
e-mail: ispany.marton@inf.unideb.hu,
[3]MTA–SZTE Analysis and Stochastics Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged, Hungary, H–6720;
e-mail: kevei@math.u-szeged.hu
[4]Department of Stochastics, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged, Hungary, H–6720;
e-mail: papgy@math.u-szeged.hu

Abstract

Multi-type inhomogeneous Galton–Watson process with immigration is investigated, where the offspring mean matrix slowly converges to a critical mean matrix. Under general conditions we obtain limit distribution for the process, where the coordinates of the limit vector are not necessarily independent.

Keywords: multi-type Galton–Watson process with immigration; INAR model; nearly critical model; multi-dimensional Poisson distribution; Toeplitz matrix; inhomogeneous model.

AMS 2000 Subject Classification: Primary 60J80, Secondary 60J27; 60J85

1 Introduction

A zero start single-type inhomogeneous Galton–Watson branching process with immigration (GWI process) $(X_n)_{n\in\mathbb{Z}_+}$ is defined as

$$\begin{cases} X_n = \sum_{j=1}^{X_{n-1}} \xi_{n,j} + \varepsilon_n, & n \in \mathbb{N}, \\ X_0 = 0, \end{cases}$$

where $\{\xi_{n,j},\, \varepsilon_n : n, j \in \mathbb{N}\}$ are independent random variables with non-negative integer values such that for each $n \in \mathbb{N}$, $\{\xi_{n,j} : j \in \mathbb{N}\}$ are identically distributed. We can interpret X_n as the number of individuals in the n^{th} generation of a population, $\xi_{n,j}$ is the number of offsprings produced by the j^{th} individual belonging to the $(n-1)^{\text{th}}$ generation, and ε_n is the number of immigrants in the n^{th} generation. A zero start one-dimensional inhomogeneous integer-valued autoregressive (INAR) time series is a special single-type GWI process, such that the offspring distributions are Bernoulli.

Assume that $\varrho_n := \mathsf{E}\xi_{n,1} < \infty$ and $m_n := \mathsf{E}\varepsilon_n < \infty$. A one-dimensional inhomogeneous GWI process $(X_n)_{n\in\mathbb{Z}_+}$ is called *nearly critical* if $\varrho_n \to 1$ as $n \to \infty$. Györfi et al. [11] investigated the asymptotic behavior of nearly critical one-dimensional INAR processes with $\varrho_n < 1$, under the assumption $\sum_{n=1}^{\infty}(1-\varrho_n) = \infty$, i.e. the convergence $\varrho_n \to 1$, $n \to \infty$, is not too fast. In the followings any non-specified limit relation is meant as $n \to \infty$. It turns out in Theorem 1 [11] that in case of Bernoulli immigration the process X_n converges in distribution to a Poisson distribution with parameter λ, when $m_n/(1-\varrho_n) \to \lambda$. That is, if there is a balance between the immigration m_n and the extinction effect $1-\varrho_n$, then a non-trivial limit distribution exists. Moreover, in [11] general immigration distributions are investigated: when the factorial moments of the immigration at generation n is of order $1-\varrho_n$ then compound Poisson limit appears. These investigations were extended by Kevei [15] for general GWI processes, that is the Bernoulli assumption on the offsprings was relaxed. In the present paper we investigate the multi-type version of the previous problem.

In a multi-type homogeneous Galton–Watson process (without immigration) the main data of the process is the spectral radius $\varrho(\boldsymbol{B})$ of the mean matrix $\boldsymbol{B}$, where $\varrho(\boldsymbol{B}) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \boldsymbol{B}\}$. By classical results, a positively regular, non-singular multi-type Galton–Watson process dies out almost surely if and only if $\varrho(\boldsymbol{B}) \le 1$. The process is called subcritical, critical or supercritical if $\varrho < 1$, $= 1$ or > 1, respectively. In the multi-type setup we also consider nearly-critical processes, that is we assume that the sequence of offspring mean matrices converge to a critical limit matrix. However, contrary to the one-dimensional case, there are a lot of critical matrices, and thus a lot of nearly-critical processes. The formal definition comes below.

An inhomogeneous multi-type GWI process with d types

$$\boldsymbol{X}_n = (X_{n,1}, \dots, X_{n,d}), \quad n \in \mathbb{Z}_+,$$

defined as

$$\begin{cases} \boldsymbol{X}_n = \sum_{j=1}^{X_{n-1,1}} \boldsymbol{\xi}_{n,j,1} + \ldots + \sum_{j=1}^{X_{n-1,d}} \boldsymbol{\xi}_{n,j,d} + \boldsymbol{\varepsilon}_n, & n \in \mathbb{N}, \\ \boldsymbol{X}_0 = \boldsymbol{0}, \end{cases}$$

where $\{\boldsymbol{\xi}_{n,j,i}, \boldsymbol{\varepsilon}_n : n, j \in \mathbb{N}, i \in \{1, \dots, d\}\}$ are independent d-dimensional random vectors with non-negative integer coordinates such that for each $n \in \mathbb{N}$ and $i \in \{1, \dots, d\}$, $\{\boldsymbol{\xi}_{n,j,i} : j \in \mathbb{N}\}$ are identically distributed, and $\boldsymbol{0}$ is the zero vector. Then $X_{n,i}$ is the number of i-type individuals in the n^{th} generation of a population, $\boldsymbol{\xi}_{n,j,i}$ is the number of offsprings produced by the j^{th} individual of type i belonging to the $(n-1)^{\text{th}}$ generation, and $\boldsymbol{\varepsilon}_n$ is the number of immigrants. When the offsprings are Bernoulli distributed (see Section 2 for the definition of multidimensional Bernoulli distribution) we obtain the d-dimensional inhomogeneous INAR time series.

Suppose that the offspring and immigration means are finite. Let us denote the offspring mean matrix and the immigration mean vector in the n^{th} generation by

$$\boldsymbol{B}_n = \begin{bmatrix} \mathsf{E}\boldsymbol{\xi}_{n,1,1} \\ \vdots \\ \mathsf{E}\boldsymbol{\xi}_{n,1,d} \end{bmatrix} \in \mathbb{R}_+^{d\times d}, \qquad \mathsf{E}\boldsymbol{\varepsilon}_n = \boldsymbol{m}_n \in \mathbb{R}_+^d,$$

where the elements of $\mathbb{R}_+^d$ are d-dimensional row vectors with non-negative coordinates. Then $(\boldsymbol{B}_n)_{i,j}$ is the expected number of type-j offsprings of a single type-i particle in generation n. Then we have the recursion

$$\mathsf{E}\boldsymbol{X}_n = (\mathsf{E}\boldsymbol{X}_{n-1})\,\boldsymbol{B}_n + \boldsymbol{m}_n, \qquad n \in \mathbb{N}, \tag{1.1}$$

since

$$\begin{aligned} \mathsf{E}(\boldsymbol{X}_n \mid \boldsymbol{X}_{n-1}) &= \mathsf{E}\left(\sum_{j=1}^{X_{n-1,1}} \boldsymbol{\xi}_{n,j,1} + \ldots + \sum_{j=1}^{X_{n-1,d}} \boldsymbol{\xi}_{n,j,d} + \boldsymbol{\varepsilon}_n \,\middle|\, \boldsymbol{X}_{n-1} \right) \\ &= \sum_{j=1}^{X_{n-1,1}} \mathsf{E}\boldsymbol{\xi}_{n,j,1} + \ldots + \sum_{j=1}^{X_{n-1,d}} \mathsf{E}\boldsymbol{\xi}_{n,j,d} + \mathsf{E}\boldsymbol{\varepsilon}_n \\ &= X_{n-1,1}\,\mathsf{E}\boldsymbol{\xi}_{n,1,1} + \ldots + X_{n-1,d}\,\mathsf{E}\boldsymbol{\xi}_{n,1,d} + \mathsf{E}\boldsymbol{\varepsilon}_n = \boldsymbol{X}_{n-1}\boldsymbol{B}_n + \boldsymbol{m}_n. \end{aligned}$$

The sequence $(\boldsymbol{B}_n)_{n\in\mathbb{N}}$ of the offspring mean matrices plays a crucial role in the asymptotic behavior of the sequence $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ as $n \to \infty$. A d-dimensional inhomogeneous Galton–Watson process $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ is called *nearly critical* if $\boldsymbol{B}_n \to \boldsymbol{B}$ and $\varrho(\boldsymbol{B}) = 1$. We will investigate the asymptotic behavior of nearly critical GWI processes.

Homogeneous multi-type GWI processes have been introduced and studied by Quine [21, 22]. In [21] necessary and sufficient condition is given for the existence of stationary distribution in the subcritical case. A complete answer is given by Kaplan [14]. Also Mode [20] gives a sufficient condition for the existence of a stationary distribution, and in a special case he shows that the limiting distribution is a multivariate Poisson with independent components.

Branching process models are extensively used in various parts of natural sciences, among others in biology, epidemiology, physics, computer sciences. In particular, multi-type GWI processes were used to determine the asymptotic mean and covariance matrix of deleterious genes and mutant genes in a stationary population by Gladstien and Lange [9], and in non-stationary population by Lange and Fan [18]. Another rapidly developing area where multi-type GWI processes can be applied is the theory of polling systems. Resing [23] pointed out that a large variety of polling models can be described as a multi-type GWI process. Resing [23], van der Mei [19], Boon [2], Boon et al. [3] and Altman and Fiems [1] investigated several communication protocols applied in info-communication networks with differentiated services. There are different quality of services, for example, some of them are delay sensitive (telephone, on-line video, etc.), while others tolerate some delay (e-mail, internet, downloading files, etc.). Thus, the services are grouped into service classes such that each class has an own transmission protocol like priority queueing. In the papers mentioned above the d-type Galton–Watson process has been used, where the process was defined either by the sizes of the active user populations of the d service classes, or by the length of the d priority queues. For the general theory and applications of multi-type Galton–Watson processes we refer to Mode [20] and Haccou et al. [13].

The INAR time series as a particular case of GWI processes with Bernoulli offspring distribution have been investigated by several authors, see e.g. the survey of Weiß [24]. Heterogeneous INAR(1) models have been considered by Böckenholt [4] for understanding and predicting consumers' buying behavior, and Gourieroux and Jasiak [10] for modeling the premium in bonus-malus scheme of car insurance. Note that the higher order INAR(p) times series introduced by Du and Li [6] has state space representation by a multivariate INAR(1) model which is a particular case of the multi-type GWI process, see Franke and Subba Rao [7].

The paper is organized as follows. In Section 2 general sufficient conditions are given for the mean matrices $\boldsymbol{B}_n$ to get a non-trivial limit distribution for the sequence $\boldsymbol{X}_n$. In Section 3 we spell out the general theorems for some special cases of the mean matrices. We investigate here the case when the limit matrix $\boldsymbol{B} = \boldsymbol{I}$, and when $\boldsymbol{B}_n = \varrho_n \boldsymbol{B}$. The proofs are gathered in Section 4.

2 General results

First we introduce some notation. Boldface lower case letters $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{m}, \boldsymbol{\lambda}$ stand for d-dimensional (row) vectors, boldface upper case letters $\boldsymbol{A}, \boldsymbol{B}$ stand for $d \times d$ real matrices, $(\boldsymbol{x})_i$ is the i^{th} element of $\boldsymbol{x}$, $(\boldsymbol{A})_{i,j}$ is the element of $\boldsymbol{A}$ in the i^{th} row and j^{th} column. For the usual basis in $\mathbb{R}^d$ we use the notation

$$\boldsymbol{e}_1 = (1, 0, \ldots, 0), \ldots, \boldsymbol{e}_d = (0, 0, \ldots, 1),$$

and for the constant zero and constant one vector we put

$$\boldsymbol{0} = (0, 0, \ldots, 0), \ \boldsymbol{1} = (1, 1, \ldots, 1).$$

Inequalities between vectors, and between matrices are meant elementwise. For a vector $\boldsymbol{x} \in \mathbb{R}^d$ its norm is denoted by $\|\boldsymbol{x}\|$, where the norm is an arbitrary norm on the linear space $\mathbb{R}^d$. As an abuse of notation $\|\boldsymbol{A}\|$ is the operator-norm of the matrix $\boldsymbol{A}$, induced by the norm $\|\cdot\|$ on the linear space $\mathbb{R}^d$, i.e. $\|\boldsymbol{A}\| = \sup_{\boldsymbol{x}:\|\boldsymbol{x}\| \leq 1} \|\boldsymbol{x}\boldsymbol{A}\|$. Therefore, all the following statements are meant as: If there exist a norm $\|\cdot\|$, such that the conditions of the statement hold with that norm, then the conclusion holds. See the example after Proposition 1.

The distribution of a random vector $\boldsymbol{\xi}$ will be denoted by $\mathcal{L}(\boldsymbol{\xi})$. For $\boldsymbol{p} = (p_1, \ldots, p_d) \in [0, 1]^d$ with $p_1 + \ldots + p_d \leq 1$, let $\text{Be}(\boldsymbol{p})$ denote the d-dimensional Bernoulli distribution with means $p_1, \ldots, p_d$ defined by

$$\text{Be}(\boldsymbol{p})(\{\boldsymbol{e}_1\}) = p_1, \quad \ldots, \quad \text{Be}(\boldsymbol{p})(\{\boldsymbol{e}_d\}) = p_d, \qquad \text{Be}(\boldsymbol{p})(\{\boldsymbol{0}\}) = 1 - p_1 - \ldots - p_d.$$

If $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$ is a random vector with $\mathcal{L}(\boldsymbol{\xi}) = \mathrm{Be}(\boldsymbol{p})$ then $\xi_1, \dots, \xi_d$ are random variables with $\mathcal{L}(\xi_i) = \mathrm{Be}(p_i)$, $i \in \{1, \dots, d\}$ (thus $\mathsf{E}\boldsymbol{\xi} = \boldsymbol{p}$), but $\xi_1, \dots, \xi_d$ are *not* independent, hence $\mathrm{Be}(\boldsymbol{p}) \neq \mathrm{Be}(p_1) \times \ldots \times \mathrm{Be}(p_d)$.

When the offspring distributions are Bernoulli, each particle has at most one offspring. In this case $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ is an inhomogeneous INAR process, such that $\mathcal{L}(\boldsymbol{\xi}_{n,1,i}) = \mathrm{Be}(\boldsymbol{e}_i\boldsymbol{B}_n)$. Note that in this case $\boldsymbol{B}_n$ is substochastic matrix.

For $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_d) \in [0,\infty)^d$, the d-dimensional Poisson distribution with parameter $\boldsymbol{\lambda}$ is defined by $\mathrm{Po}(\boldsymbol{\lambda}) := \mathrm{Po}(\lambda_1) \times \ldots \times \mathrm{Po}(\lambda_d)$. In other words, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$ is a random vector with $\mathcal{L}(\boldsymbol{\xi}) = \mathrm{Po}(\boldsymbol{\lambda})$ whenever $\xi_1, \dots, \xi_d$ are independent random variables with $\mathcal{L}(\xi_i) = \mathrm{Po}(\lambda_i)$, $i \in \{1, \dots, d\}$. Note that $\mathsf{E}\boldsymbol{\xi} = \boldsymbol{\lambda}$.

Introduce the generating functions

$$\begin{aligned} &F_n(\boldsymbol{x}) = \mathsf{E}\boldsymbol{x}^{\boldsymbol{X}_n}, \quad G_{n,i}(\boldsymbol{x}) = \mathsf{E}\boldsymbol{x}^{\boldsymbol{\xi}_{n,1,i}}, \quad \boldsymbol{G}_n(\boldsymbol{x}) = (G_{n,1}(\boldsymbol{x}), \dots, G_{n,d}(\boldsymbol{x})), \\ &H_n(\boldsymbol{x}) = \mathsf{E}\boldsymbol{x}^{\boldsymbol{\varepsilon}_n}, \quad \boldsymbol{x} \in [0,1]^d, \end{aligned} \tag{2.1}$$

where $\boldsymbol{x}^{\boldsymbol{k}} = x_1^{k_1} \dots x_d^{k_d}$. Conditioning argument shows the recursion $F_n(\boldsymbol{x}) = F_{n-1}(\boldsymbol{G}_n(\boldsymbol{x}))H_n(\boldsymbol{x})$, $n \geq 2$. Let denote $\overline{\boldsymbol{G}}_{n+1,n}(\boldsymbol{x}) = \boldsymbol{x}$, and if $\overline{\boldsymbol{G}}_{j+1,n}$ is defined then $\overline{\boldsymbol{G}}_{j,n}(\boldsymbol{x}) = \boldsymbol{G}_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}))$. With this notation Quine [21] proved (simple induction argument shows) that we have

$$F_n(\boldsymbol{x}) = \prod_{j=1}^{n} H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})). \tag{2.2}$$

It turns out that due to the near-criticality under general conditions $H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) \approx 1$, for each j, thus

$$H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) \approx \exp\left\{H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1\right\},$$

therefore it is reasonable to define the accompanying compound Poisson probability generating function

$$\widetilde{F}_n(\boldsymbol{x}) = \exp\left\{\sum_{j=1}^{n}\left[H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1\right]\right\}. \tag{2.3}$$

We prove in Lemma 3 that under some conditions

$$\lim_{n\to\infty}(F_n(\boldsymbol{x}) - \widetilde{F}_n(\boldsymbol{x})) = 0.$$

Therefore to determine the asymptotic properties of $\boldsymbol{X}_n$ we have to investigate the sum

$$\sum_{j=1}^{n}\left[H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1\right].$$

We can compute explicitly the generating function when both the immigration and the offsprings have Bernoulli distribution. Indeed, when $\mathcal{L}(\boldsymbol{\varepsilon}_n) = \mathrm{Be}(\boldsymbol{m}_n)$, the immigration generating function is

$$H_n(\boldsymbol{x}) = \mathsf{E}\boldsymbol{x}^{\boldsymbol{\varepsilon}_n} = m_{n,1}x_1 + \ldots + m_{n,d}x_d + 1 - (m_{n,1} + \ldots + m_{n,d}) = 1 + (\boldsymbol{x} - \boldsymbol{1})\boldsymbol{m}_n^\top;$$

when the offspring distributions are also Bernoulli then $\boldsymbol{G}_n(\boldsymbol{x}) = \boldsymbol{1} + (\boldsymbol{x} - \boldsymbol{1})\boldsymbol{B}_n^\top$, and so $\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) = \boldsymbol{1} + (\boldsymbol{x} - \boldsymbol{1})\boldsymbol{B}_{[j,n]}^\top$, where

$$\boldsymbol{B}_{[j,n]} := \begin{cases} \boldsymbol{B}_{j+1} \ldots \boldsymbol{B}_n, & \text{for } 0 \le j \le n-1, \\ \boldsymbol{I}, & \text{for } j = n. \end{cases}$$

Note that in this paper the multivariate Bernoulli distribution is defined in such a way that its generating function is a first order polynomial which is a particular case of a more general definition of the multivariate Bernoulli distribution, see Krummenauer [17, Definition 1]. Thus (2.3) reads as

$$\widetilde{F}_n(\boldsymbol{x}) = \exp\left\{ \sum_{j=1}^n \left[(\boldsymbol{x} - \boldsymbol{1})\boldsymbol{B}_{[j,n]}^\top \boldsymbol{m}_j^\top \right] \right\}.$$

Observe that the recursion (1.1) implies

$$\mathsf{E}\boldsymbol{X}_n = \sum_{j=1}^n \boldsymbol{m}_j \, \boldsymbol{B}_{[j,n]}. \tag{2.4}$$

This can be obtained also by differentiating F_n in (2.2). Putting

$$\boldsymbol{A}_{j,n} = (\boldsymbol{B} - \boldsymbol{B}_j)\boldsymbol{B}_{[j,n]}, \qquad n \in \mathbb{N}, \quad j \in \{1, \ldots, n\}, \tag{2.5}$$

we may rewrite (2.4) as

$$\mathsf{E}\boldsymbol{X}_n = \sum_{j=1}^n \boldsymbol{m}_j (\boldsymbol{B} - \boldsymbol{B}_j)^{-1} \boldsymbol{A}_{j,n},$$

whenever the inverse $(\boldsymbol{B} - \boldsymbol{B}_j)^{-1}$ exists for each $j = 1, \ldots, n$.

These computations shows the necessity for a summability method defined by the offspring mean matrices. We will make of use of the following matrix version of Toeplitz theorem (see, e.g., in Fritz [8]).

Lemma 1. *Let $\boldsymbol{A}_{j,n} \in \mathbb{R}^{d\times d}$, $n \in \mathbb{N}$, $j = 1, 2, \ldots, n$ be matrices such that*

$$\lim_{n\to\infty} \max_{1\le j\le n} \|\boldsymbol{A}_{j,n}\| = 0, \tag{2.6}$$

$$\lim_{n\to\infty} \sum_{j=1}^n \boldsymbol{A}_{j,n} = \boldsymbol{A}, \tag{2.7}$$

$$\sup_{n\in\mathbb{N}} \sum_{j=1}^n \|\boldsymbol{A}_{j,n}\| < \infty. \tag{2.8}$$

Then for any convergent sequence of vectors $\boldsymbol{x}_n \to \boldsymbol{x}$

$$\sum_{j=1}^{n} \boldsymbol{x}_j \boldsymbol{A}_{j,n} \to \boldsymbol{x}\boldsymbol{A}.$$

In fact, the Lemma holds with the weaker assumption $\|\boldsymbol{A}_{j,n}\| \to 0$, for all $j \in \mathbb{N}$, instead of (2.6). However, for the proof of Lemma 3 the stronger version is needed.

Lemma 2. *Assume that the sequence of mean matrices* $(\boldsymbol{B}_n)_{n\in\mathbb{N}}$ *satisfies the following conditions:*

(B1) $\lim_{n\to\infty} \boldsymbol{B}_n = \boldsymbol{B}$*, for some matrix* $\boldsymbol{B}$*;*

(B2) $\|\boldsymbol{B}_n\| \leq 1$ *and* $\boldsymbol{B} - \boldsymbol{B}_n$ *is invertible whenever* $n \geq n_0$ *for some* n_0*;*

(B3) $\lim_{n\to\infty} \|\boldsymbol{B}_{[j,n]}\| = 0$ *for any fixed* j*;*

(B4) $\lim_{n\to\infty} \sum_{j=1}^{n} (\boldsymbol{B} - \boldsymbol{B}_j)\boldsymbol{B}_{[j,n]} = \boldsymbol{A}$ *for some limit matrix* $\boldsymbol{A}$*;*

(B5) $\sup_n \sum_{j=1}^{n} \|(\boldsymbol{B} - \boldsymbol{B}_j)\boldsymbol{B}_{[j,n]}\| < \infty$.

Then the triangular matrix array $(\boldsymbol{A}_{j,n} = (\boldsymbol{B} - \boldsymbol{B}_j)\boldsymbol{B}_{[j,n]})_{j,n}$ *satisfies the conditions of Lemma 1.*

The following two general theorems give sufficient condition for the convergence of $\boldsymbol{X}_n$. It turns out that in case of Bernoulli offspring and immigration only conditions (2.6)–(2.8) have to be assured. Note that when the offspring distribution is Bernoulli, then the limit matrix $\boldsymbol{B}$ in (B1) is necessarily substochastic.

Theorem 1. *Let* $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ *be an inhomogeneous GWI process such that both the offspring and the immigration have Bernoulli distribution and* (B1)–(B5) *hold. If*

(M) $\lim_{n\to\infty} \boldsymbol{m}_n(\boldsymbol{B} - \boldsymbol{B}_n)^{-1} = \boldsymbol{\lambda}$,

then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \mathrm{Po}(\boldsymbol{\lambda}\boldsymbol{A}),$$

where $\boldsymbol{A}$ *is given in* (B4).

The Bernoulli distribution of the offsprings and the immigration is a very restrictive condition. In the following theorems we weaken these assumptions.

The interesting new feature in the following theorem is that the components of the limit are dependent in general. We need some further notation.

For a multi-index $\boldsymbol{j} = (j_1, \ldots, j_d) \in \mathbb{Z}_+^d$ let denote $m_{n,\boldsymbol{j}}$ the $\boldsymbol{j}^{\text{th}}$ factorial moment of the immigration $\boldsymbol{\varepsilon}_n = (\varepsilon_{n,1}, \ldots, \varepsilon_{n,d})$, that is

$$m_{n,\boldsymbol{j}} = \mathsf{E}\left(\prod_{i=1}^{d} \varepsilon_{n,i}(\varepsilon_{n,i} - 1)\ldots(\varepsilon_{n,i} - j_i + 1)\right) = D^{\boldsymbol{j}} H_n(\mathbf{1}),$$

where for a multi-index $\boldsymbol{j}$

$$D^{\boldsymbol{j}} H_n(\mathbf{1}) = \frac{\partial^{|\boldsymbol{j}|}}{\partial^{j_1} x_1 \dots \partial^{j_d} x_d} H_n(\mathbf{1}),$$

$|\boldsymbol{j}| = j_1 + \ldots + j_d$, and the derivatives are meant as the left derivatives.

We cannot circumvent the fairly inconvenient notation below and in Lemma 4, because formulas (2.10) and (2.11) are not easily translated to the multi-index notation.

Theorem 2. *Let $(\boldsymbol{X}_n)_{n \in \mathbb{Z}_+}$ be an inhomogeneous GWI process with Bernoulli offspring distributions, such that* (B1)–(B5) *hold. Moreover, assume that for some $k \geq 2$*

$$\lim_{n \to \infty} \|(\boldsymbol{B} - \boldsymbol{B}_n)^{-1}\| \max_{|\boldsymbol{j}|=k} D^{\boldsymbol{j}} H_n(\mathbf{1}) = 0, \tag{2.9}$$

and for each $i = 1, 2, \ldots, k-1$, for each $1 \leq \ell_{i+1}, \ldots, \ell_{2i} \leq d$, the limit

$$\lim_{n \to \infty} \sum_{j=1}^{n} \frac{1}{i!} \sum_{\ell_1, \ldots, \ell_i = 1}^{d} \frac{\partial^i H_j(\mathbf{1})}{\partial x_{\ell_1} \ldots \partial x_{\ell_i}} \left(\boldsymbol{B}_{[j,n]}\right)_{\ell_1, \ell_{i+1}} \ldots \left(\boldsymbol{B}_{[j,n]}\right)_{\ell_i, \ell_{2i}} =: \Lambda_{i;\ell_{i+1},\ldots,\ell_{2i}} \tag{2.10}$$

exists. Then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \boldsymbol{Y},$$

where

$$\mathsf{E}\boldsymbol{x}^{\boldsymbol{Y}} = \exp\left\{ \sum_{i=1}^{k-1} \sum_{\ell_{i+1},\ldots,\ell_{2i}=1}^{d} \Lambda_{i;\ell_{i+1},\ldots,\ell_{2i}} (x_{\ell_{i+1}} - 1) \ldots (x_{\ell_{2i}} - 1) \right\}. \tag{2.11}$$

Note that if (2.10) holds, then necessarily $\Lambda_{i;\cdot}$ is symmetric in the sense that for any permutation π we have $\Lambda_{i;\ell_1,\ell_2,\ldots,\ell_i} = \Lambda_{i;\ell_{\pi 1},\ell_{\pi 2},\ldots,\ell_{\pi i}}$. In particular, $\Lambda_{2;j,k} = \Lambda_{2;k,j}$, we use this in Example 1.

A simple sufficient condition which guarantees (2.9) is that there are at most $k-1$ immigrants in any generation. The other condition is more difficult to check, however for $i = 1, 2$ we can write it in a simpler form.

For $i = 1$ condition (2.10) is just the convergence

$$\sum_{j=1}^{n} \boldsymbol{m}_j \boldsymbol{B}_{[j,n]} \to \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d),$$

with $\Lambda_{1;\ell} = \lambda_\ell$. Since the matrix array $(\boldsymbol{A}_{j,n})$ satisfies the conditions of Lemma 1 we see that the convergence above follows from condition (M). As a consequence we obtain

Corollary 1. *Let $(\boldsymbol{X}_n)_{n \in \mathbb{Z}_+}$ be an inhomogeneous GWI process with Bernoulli offspring distributions, such that* (B1)–(B5) *hold. Moreover, assume* (M) *and*

$$\lim_{n \to \infty} \mathsf{E}\|\boldsymbol{\varepsilon}_n\|^2 \, \|(\boldsymbol{B} - \boldsymbol{B}_n)^{-1}\| = 0. \tag{2.12}$$

Then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \mathrm{Po}(\boldsymbol{\lambda}\boldsymbol{A}).$$

For $i = 2$ condition (2.10) takes the form

$$\frac{1}{2}\sum_{j=1}^{n} \boldsymbol{B}_{[j,n]}^{\top} \Delta_j \boldsymbol{B}_{[j,n]} \to \boldsymbol{\Lambda}_2,$$

with $\Lambda_{2;k,\ell} = (\boldsymbol{\Lambda}_2)_{k,\ell}$, where $\Delta_j = \Delta H_j(\mathbf{1})$ is the Hesse-matrix of the immigration generating function at $\mathbf{1}$; i.e. $(\Delta H_j(\mathbf{1}))_{k,\ell} = \frac{\partial^2}{\partial x_k \partial x_\ell} H_j(\mathbf{1})$.

Example 1. The following simple example shows that the limit may have dependent components even in a simple case. Let $d = 2$, $\mathcal{L}(\boldsymbol{\xi}_{n,1,i}) = \mathrm{Be}((1-n^{-1})\boldsymbol{e}_i)$, $i = 1, 2$, and $\mathsf{P}(\boldsymbol{\varepsilon}_n = \mathbf{0}) = 1 - n^{-1}$, $\mathsf{P}(\boldsymbol{\varepsilon}_n = \mathbf{1}) = n^{-1}$ for all $n \in \mathbb{N}$, that is in the n^{th} generation each particle survives with probability $1 - n^{-1}$, and with probability n^{-1} a type-1 and a type-2 particle immigrate together. Then we have

$$\boldsymbol{B}_n = \left(1 - \frac{1}{n}\right)\boldsymbol{I}, \quad \text{and} \quad H_n(x_1, x_2) = 1 - \frac{1}{n} + \frac{x_1 x_2}{n}.$$

Clearly condition (2.9) holds with $k = 3$. The relevant quantities are $\boldsymbol{B} = \boldsymbol{I}$, $\boldsymbol{m}_n = \frac{1}{n}(1, 1)$,

$$\Delta_n = \frac{1}{n}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{m}_n(\boldsymbol{B} - \boldsymbol{B}_n)^{-1} = \frac{1}{n}(1, 1)n\boldsymbol{I} = (1, 1),$$

and

$$\sum_{j=1}^{n} \boldsymbol{B}_{[j,n]}^{\top} \Delta_j \boldsymbol{B}_{[j,n]} = \sum_{j=1}^{n} \frac{j}{n^2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \to \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We see that $\Lambda_{1;1} = \Lambda_{1;2} = 1$ and $\Lambda_{2;1,2} = \Lambda_{2;2,1} = 1/4$, $\Lambda_{2;1,1} = \Lambda_{2;2,2} = 0$. Thus

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \boldsymbol{Y}, \quad \text{where} \quad \mathsf{E}\boldsymbol{x}^{\boldsymbol{Y}} = \exp\left\{x_1 - 1 + x_2 - 1 + \frac{(x_1 - 1)(x_2 - 1)}{2}\right\}.$$

Let U, V, W be independent Poisson random variables with parameters $\lambda_1, \lambda_2, \mu$, respectively. The generating function of $(U + W, V + W)$ is given by

$$\begin{aligned} \mathsf{E}x_1^{U+W} x_2^{V+W} &= \mathsf{E}x_1^U \mathsf{E}x_2^V \mathsf{E}(x_1 x_2)^W \\ &= \exp\left\{\lambda_1(x_1 - 1) + \lambda_2(x_2 - 1) + \mu(x_1 x_2 - 1)\right\}, \end{aligned}$$

therefore the distribution of the limit Y is the distribution of the vector $(U + W, V + W)$ where U, V, W are iid Poisson(1/2). The distribution is called bivariate Poisson distribution, with parameters λ_1, λ_2, and μ, see Johnson et al. [12] p.124, or Kocherlakota et al. [16].

In general, when in the exponent in (2.11) none of the terms are divisible with x_i^2 for any i (e.g. at most 1 particle immigrates for any given type), then the components of the limit Y in Theorem 2 can be represented as the sum of independent Poisson random variables. Assume that the conditions of Theorem 2 hold, with $k = 3$ in (2.9), and for the limits in (2.10) $\Lambda_{2;i,i} = 0$ for all i. Then the limit random vector $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ can be represented as

$$Y_i = U_i + \sum_{j \neq i} U_{i,j}, \quad i = 1, \dots, d,$$

where $(U_i)_{i=1}^d$ and $(U_{i,j})_{1 \le i < j \le d}$ are independent Poisson random variables, with parameters a_i and $a_{i,j}$ respectively, with

$$a_i = \Lambda_{1;i} - 2 \sum_{j=1}^{d} \Lambda_{2;i,j}, \quad i = 1, \dots, d, \quad a_{i,j} = 2\Lambda_{2;i,j}, \quad 1 \le i < j \le d,$$

and $U_{i,j} := U_{j,i}$ for $i > j$. It is not difficult to show that (2.9) with $k = 3$ and (2.10) imply that the coefficients above are non-negative. Simple computation shows that the generating function of $\boldsymbol{Y}$ agrees with the one given in Theorem 2. Clearly, this construction extends for $k \ge 3$. The appearing limiting distributions are the so-called multivariate Poisson distributions; for further properties see [12], p.139. Note that this multivariate Poisson distribution appears as a limit in the multivariate version of the law of small numbers, see Krummenauer [17, Theorem 1]. Hence, Theorem 2 can be interpreted as a general law of small numbers for inhomogeneous GWI processes. Also note the difference between the multivariate Poisson distribution introduced here and the d-dimensional Poisson distribution defined before (2.1).

In the next theorem the condition on the offspring distribution is relaxed, though (2.14) means that the offspring distribution has to be very close to a Bernoulli distribution. Note that in this case we assume that the limit matrix is the unit matrix $\boldsymbol{I}$, in which case condition (B4) automatically holds, with limit matrix $\boldsymbol{A} = \boldsymbol{I}$. We return to this question in Subsection 3.1. To state the theorem we introduce the notation

$$m_2(n) = \max_{1 \le i,j,l \le d} \frac{\partial^2}{\partial x_j \partial x_l} G_{n,i}(\mathbf{1}). \tag{2.13}$$

Theorem 3. *Let $(\boldsymbol{X}_n)_{n \in \mathbb{Z}_+}$ be an inhomogeneous GWI process with Bernoulli immigration, such that $\boldsymbol{B} = \boldsymbol{I}$ and* (B1)–(B5) *hold. Moreover, assume* (M) *and*

$$\lim_{n \to \infty} m_2(n) \|(\boldsymbol{I} - \boldsymbol{B}_n)^{-1}\| = 0. \tag{2.14}$$

Then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \mathrm{Po}(\boldsymbol{\lambda}).$$

In the single-type case general immigration distribution is investigated and convergence to a compound Poisson distribution is proved in Theorem 4 in [11], and in Theorem 3 in [15]. In case of more general offspring distribution existence of negative binomial limit is showed in Theorem 5 in [15]. However, in our multi-type scenario the computations with general immigration or (and) with general offspring distribution become intractably complicated.

Finally, we note that if $\prod_{n=1}^{\infty} \boldsymbol{B}_n$ exists and is not the 0 matrix, then the process $\boldsymbol{X}_n$ converges when $\sum_{n=1}^{\infty} \boldsymbol{m}_n$ is finite. This case can be handled similarly as in the single-type scenario in [15].

3 Special cases and examples

In what follows we investigate some special cases for the sequence of mean matrices, and we give sufficient conditions for the existence of the distributional limit, which are easier to handle than the ones given in Theorem 1.

3.1 The case $\boldsymbol{B} = \boldsymbol{I}$

When the critical limit matrix is the identity matrix then one of the most complicated assumption, (B4) in Theorem 1, holds automatically. In this case $\boldsymbol{A}_{j,n} = \boldsymbol{B}_{[j,n]} - \boldsymbol{B}_{[j-1,n]}$, and so $\sum_{j=1}^{n} \boldsymbol{A}_{j,n}$ is a telescopic sum.

Proposition 1. *Suppose that*

(I1) $\lim_{n\to\infty} \boldsymbol{B}_n = \boldsymbol{I}$;

(I2) *there is an* n_0 *such that* $\|\boldsymbol{B}_n\| < 1$ *for all* $n \geq n_0$;

(I3) $\lim_{n\to\infty} \|\boldsymbol{B}_{[j,n]}\| = 0$ *for all* $j \in \mathbb{N}$;

(I4) *there is an* n_0 *such that* $\sup\limits_{n\geq n_0} \frac{\|\boldsymbol{I}-\boldsymbol{B}_n\|}{1-\|\boldsymbol{B}_n\|} < \infty$ *or* $\boldsymbol{A}_{j,n} \in \mathbb{R}_+^{d\times d}$ *for all* $n \geq n_0$ *and all* $j \in \{1, \ldots, n\}$.

Then the triangular matrix array $(\boldsymbol{A}_{j,n})_{j,n}$ *satisfies the conditions of Lemma 1.*

Note that condition (I2) guarantees the existence of the inverse $(\boldsymbol{I} - \boldsymbol{B}_n)^{-1}$ in (M) for $n \geq n_0$.

As we mentioned, the norm can be arbitrary operator norm. It is easy to construct examples, such that some conditions hold in one operator norm, and fail in another. For instance, let

$$\boldsymbol{B}_n = \begin{bmatrix} 1 - \frac{1}{n} & \frac{1}{n} \\ 0 & 1 - \frac{2}{n} \end{bmatrix}.$$

Then in column sum norm (induced by the ℓ_∞ norm on $\mathbb{R}^d$) condition (I2) and (I4) hold, while in the row sum norm (induced by the ℓ_1 norm on $\mathbb{R}^d$) even condition (I2) fails.

For jointly diagonalizable offspring mean matrices a better result is available, namely, condition (I4) above can be omitted.

Proposition 2. *Suppose that conditions* (I1)–(I3) *of Proposition 1 hold, and the offspring mean matrices are of the form*

$$\boldsymbol{B}_n = \boldsymbol{U} \operatorname{diag}(\varrho_{n,1}, \dots, \varrho_{n,d}) \boldsymbol{U}^\top, \qquad n \in \mathbb{N},$$

where $\boldsymbol{U} \in \mathbb{R}^{d\times d}$ *is an orthogonal matrix. Then the triangular matrix array* $(\boldsymbol{A}_{j,n})_{j,n}$ *satisfies the conditions of Lemma 1.*

As a consequence we obtain that the corresponding versions of Theorem 1 and 2 can be stated. For example the following holds.

Theorem 4. *Let* $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ *be an inhomogeneous GWI process with Bernoulli offspring and immigration distributions. Assume that either conditions of Proposition 1 or Proposition 2 are satisfied, and for the immigration* (M) *holds. Then*

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \operatorname{Po}(\boldsymbol{\lambda}).$$

Remark 1. The statement of Proposition 2 for the special case $\boldsymbol{B}_n = \varrho_n \boldsymbol{I}$, $n \in \mathbb{N}$, with $\varrho_n \in [0,1]$, $n \in \mathbb{N}$, also follows from Proposition 1; in this case Theorem 4 imply the appropriate results for one-dimensional inhomogeneous INAR processes due to Györfi et al. [11].

Note that under the assumption of Proposition 2 conditions (I1)–(I3) of Proposition 1 are equivalent to

(I1') $\lim_{n\to\infty} \varrho_{n,i} = 1$ for all $i \in \{1,\dots,d\}$;

(I2') $\max_{1\le i\le d} \varrho_{n,i} < 1$ for all $n \ge n_0$;

(I3') $\prod_{n=j}^{\infty} \varrho_{n,i} = 0$ for all $j \in \mathbb{N}$ and all $i \in \{1,\dots,d\}$,

respectively. Remark that conditions (I3) and (I3') are also equivalent to

(I3") $\sum_{n=1}^{\infty} (1-\varrho_{n,i}) = +\infty$ for all $i \in \{1,\dots,d\}$.

The following example shows that Proposition 2 can really perform better for jointly diagonalizable offspring mean matrices than Proposition 1.

Example 2. Let $d = 2$, $\varrho_{n,1} = 1 - 1/n$, $\varrho_{n,2} = 1 - 1/\sqrt{n}$,

$$\boldsymbol{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \text{ hence } \boldsymbol{B}_n = \begin{bmatrix} 1 - \frac{\sqrt{n}+1}{2n} & \frac{\sqrt{n}-1}{2n} \\ \frac{\sqrt{n}-1}{2n} & 1 - \frac{\sqrt{n}+1}{2n} \end{bmatrix}.$$

Then conditions (I1)–(I3) of Proposition 1 are trivially satisfied, but condition (I4) of Proposition 1 fails to hold. Indeed,

$$\boldsymbol{A}_{n,n} = \boldsymbol{I} - \boldsymbol{B}_n = \begin{bmatrix} \frac{\sqrt{n}+1}{2n} & -\frac{\sqrt{n}-1}{2n} \\ -\frac{\sqrt{n}-1}{2n} & \frac{\sqrt{n}+1}{2n} \end{bmatrix} \notin \mathbb{R}_+^{2\times 2}, \quad \|\boldsymbol{B}_n\| = \left\| \operatorname{diag}\left(1-\frac{1}{n}, 1-\frac{1}{\sqrt{n}}\right)\right\| = 1-\frac{1}{n}$$

and

$$\begin{aligned} \|\boldsymbol{I} - \boldsymbol{B}_n\| &= \left\| \boldsymbol{U}\left(\boldsymbol{I} - \operatorname{diag}\left(1-\frac{1}{n}, 1-\frac{1}{\sqrt{n}}\right)\right)\boldsymbol{U}^\top \right\| = \left\| \boldsymbol{I} - \operatorname{diag}\left(1-\frac{1}{n}, 1-\frac{1}{\sqrt{n}}\right)\right\| \\ &= \left\| \operatorname{diag}\left(\frac{1}{n}, \frac{1}{\sqrt{n}}\right)\right\| = \frac{1}{\sqrt{n}} \end{aligned}$$

imply $\sup\limits_{n\geq n_0} \frac{\|\boldsymbol{I}-\boldsymbol{B}_n\|}{1-\|\boldsymbol{B}_n\|} = \infty$. Here we used the simple fact that the norm of a normal element in a C^*-algebra is equal to its spectral radius.

3.2 The case $\boldsymbol{B}_n = \varrho_n \boldsymbol{B}$

In this subsection we assume that $\boldsymbol{B}_n = \varrho_n \boldsymbol{B}$, for all $n \in \mathbb{N}$, where $\boldsymbol{B}$ is a substochastic matrix, and $\varrho_n < 1, \varrho_n \to 1$ and $\sum_{n=1}^\infty (1-\varrho_n) = \infty$. In this special case $\boldsymbol{B}_{[j,n]} = \varrho_{[j,n]} \boldsymbol{B}^{n-j}$, with $\varrho_{[j,n]} = \varrho_{j+1} \dots \varrho_n$. Put $a_{j,n} = \varrho_{[j,n]}(1-\varrho_j)$, then $\boldsymbol{A}_{j,n} = \boldsymbol{B}_{[j,n]}(\boldsymbol{B} - \boldsymbol{B}_j) = a_{j,n} \boldsymbol{B}^{n-j+1}$.

To apply Theorem 1 or 2 in this case, the missing condition is again (B4). In the following statement we give a rather general condition for the existence of the limit matrix. The key point is a slight modification of the proof of Theorem 5.2.1 in Doob [5].

Proposition 3. *Let $(\varrho_n)_{n\in\mathbb{N}}$ be a sequence, such that $\varrho_n < 1$, $\varrho_n \to 1$, $\sum_{n=1}^\infty (1-\varrho_n) = \infty$ and $(1-\varrho_n)/(1-\varrho_{n+1}) \to 1$. Then for any matrix $\boldsymbol{B}$ such that $\|\boldsymbol{B}\| \leq 1$ the limit*

$$\lim_{n\to\infty} \sum_{k=1}^n \varrho_{[k,n]}(1-\varrho_k)\boldsymbol{B}^{n-k} = \boldsymbol{A}$$

exists, and $\boldsymbol{BA} = \boldsymbol{AB} = \boldsymbol{A} = \boldsymbol{A}^2$.

It will be clear from the proof that whenever $\boldsymbol{B}$ is stochastic the limit $\boldsymbol{A}$ is stochastic too.

Note that in the single-type case no additional assumption is needed on the sequence $(\varrho_n)_{n\in\mathbb{N}}$, see [11] or [15]. Indeed, the condition $\sum_{n=1}^\infty (1-\varrho_n) = \infty$ implies that the numerical triangular array $(a_{j,n} = \varrho_{[j,n]}(1-\varrho_j))$ satisfies the following conditions, which are the 1-dimensional analog of

(2.6), (2.7) and (2.8):

$$\lim_{n\to\infty} \max_{1\le j\le n} a_{j,n} = 0,$$
$$\lim_{n\to\infty} \sum_{j=1}^{n} a_{j,n} = 1, \tag{3.1}$$
$$\sup_{n\ge 1} \sum_{j=1}^{n} |a_{j,n}| < \infty.$$

The following example shows that when dealing with matrices the additional assumption is in fact necessary.

Example 3. Let

$$\varrho_n = \begin{cases} 1, & \text{if } n \text{ is odd,} \\ 1 - \frac{2}{n}, & \text{if } n \text{ is even,} \end{cases}$$

and

$$\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then $\varrho_{[k,n]} = [k/2]/[n/2]$, where $[\cdot]$ stands for the (lower) integer part, and so

$$\varrho_{[k,n]}(1-\varrho_k) = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ \frac{1}{[n/2]}, & \text{if } k \text{ is even.} \end{cases}$$

Thus we obtain

$$\sum_{k=1}^{2n} \boldsymbol{A}_{k,2n} = \sum_{k=1}^{2n} \boldsymbol{B}^{2n-k+1} \varrho_{[k,2n]}(1-\varrho_k) \to \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

while

$$\sum_{k=1}^{2n+1} \boldsymbol{A}_{k,2n+1} = \sum_{k=1}^{2n+1} \boldsymbol{B}^{2n-k+2} \varrho_{[k,2n+1]}(1-\varrho_k) \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

thus the limit does not exist.

Using Proposition 3 and Theorem 2 we obtain the following

Theorem 5. *Assume that the mean matrix of the Bernoulli offspring distribution has the form*

$$\boldsymbol{B}_n = \varrho_n \boldsymbol{B},$$

where $\boldsymbol{B}$ is an invertible substochastic matrix, and $\varrho_n < 1$, $\varrho_n \to 1$, $\sum_{k=1}^{\infty}(1-\varrho_n) = \infty$ and $(1-\varrho_n)/(1-\varrho_{n+1}) \to 1$. Moreover, assume that for some $k \geq 2$

$$\lim_{n\to\infty} \frac{\max_{|\boldsymbol{j}|=k} D^{\boldsymbol{j}} H_n(\boldsymbol{1})}{1-\varrho_n} = 0, \quad \lim_{n\to\infty} \frac{\boldsymbol{m}_n}{1-\varrho_n} = \boldsymbol{\lambda},$$

and for each $i = 2, \ldots, k-1$, for each $1 \leq \ell_{i+1}, \ldots, \ell_{2i} \leq d$ the limit

$$\lim_{n\to\infty} \sum_{j=1}^{n} \frac{(\varrho_{[j,n]})^i}{i!} \sum_{\ell_1,\ldots,\ell_i=1}^{d} \frac{\partial^i H_j(\boldsymbol{1})}{\partial x_{\ell_1}\ldots\partial x_{\ell_i}} \left(\boldsymbol{B}^{n-j}\right)_{\ell_1,\ell_{i+1}} \ldots \left(\boldsymbol{B}^{n-j}\right)_{\ell_i,\ell_{2i}} =: \Lambda_{i;\ell_{i+1},\ldots,\ell_{2i}}$$

exists. Then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \boldsymbol{Y},$$

where

$$\mathsf{E}\boldsymbol{x}^{\boldsymbol{Y}} = \exp\left\{(\boldsymbol{x}-\boldsymbol{1})(\boldsymbol{\lambda}\boldsymbol{A})^\top + \sum_{i=2}^{k-1} \sum_{\ell_{i+1},\ldots,\ell_{2i}=1}^{d} \Lambda_{i;\ell_{i+1},\ldots,\ell_{2i}}(x_{\ell_{i+1}}-1)\ldots(x_{\ell_{2i}}-1)\right\},$$

where the matrix $\boldsymbol{A}$ is given by Proposition 3.

The next theorem gives more freedom on the mean matrix $\boldsymbol{B}_n$, however stronger assumption on the limit matrix $\boldsymbol{B}$ is needed. The inequality for the matrices are meant elementwise.

Theorem 6. *Assume that for the mean matrix of the Bernoulli offspring distribution*

$$\vartheta_n \boldsymbol{B} \leq \boldsymbol{B}_n \leq \varrho_n \boldsymbol{B}$$

holds, where $\vartheta_n \leq \varrho_n < 1$, $\vartheta_n \to 1$, $\varrho_n \to 1$, $\sum_{n=1}^{\infty}(1-\varrho_n) = \infty$, and $(\varrho_n - \vartheta_n)/(1-\varrho_n) \to 0$, and for the immigration

$$\boldsymbol{m}_n(\boldsymbol{B}-\boldsymbol{B}_n)^{-1} \to \boldsymbol{\lambda}.$$

If either (a) $\boldsymbol{B}^n \to \boldsymbol{A}$ for some matrix $\boldsymbol{A}$ or (b) $\|\boldsymbol{B}\| \leq 1$ and $(1-\varrho_{n+1})/(1-\varrho_n) \to 1$ holds then

$$\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \mathrm{Po}(\boldsymbol{\lambda}\boldsymbol{A}),$$

where in case (b) the matrix $\boldsymbol{A}$ is given in Proposition 3.

4 Proofs

Before the proofs, we gather some simple inequalities, which we use frequently without further reference. If $a_k, b_k \in [-1, 1]$, $k = 1, \ldots, n$, then

$$\left| \prod_{k=1}^{n} a_k - \prod_{k=1}^{n} b_k \right| \le \sum_{k=1}^{n} |a_k - b_k|.$$

For $x \in (-1, 1)$ we have $|\mathrm{e}^x - 1 - x| \le x^2$.

For a vector-vector function $\boldsymbol{H} : \mathbb{R}^d \to \mathbb{R}^d$ the symbol $\nabla \boldsymbol{H}$ denotes

$$\begin{bmatrix} \frac{\partial}{\partial x_1} H_1 & \ldots & \frac{\partial}{\partial x_d} H_1 \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} H_d & \ldots & \frac{\partial}{\partial x_d} H_d \end{bmatrix}.$$

By the multivariate mean-value theorem, and the monotonicity of the derivatives, for a vector of generating functions $\boldsymbol{G} = (\boldsymbol{G}_1, \ldots, \boldsymbol{G}_d)$, for $\boldsymbol{x} \in [0, 1]^d$

$$\mathbf{1} - \boldsymbol{G}(\boldsymbol{x}) \le (\mathbf{1} - \boldsymbol{x}) \nabla \boldsymbol{G}(\mathbf{1})^\top. \tag{4.1}$$

4.1 Proofs for Section 2

Since $\|\boldsymbol{B}_{[j,n]}\| \to 0$ for any j, we may and do assume that $n_0 = 1$ in conditions (B2), (I2) and (I4).

Proof of Lemma 2. Condition (2.6) simply follows from (B1), (B2) and (B3). Conditions (2.7) and (2.8) are the same as (B4), and (B5), respectively. □

Remark 2. It is worth to note that after rearranging the sum in (B4) and using (B3) we obtain that (B4) is equivalent to the convergence

$$\lim_{n \to \infty} \sum_{j=1}^{n} \boldsymbol{B}_{[j,n]} = \widetilde{\boldsymbol{B}},$$

where the relation between $\boldsymbol{A}$ and $\widetilde{\boldsymbol{B}}$ is given by $(\boldsymbol{B} - \boldsymbol{I})\widetilde{\boldsymbol{B}} + \boldsymbol{B} = \boldsymbol{A}$.

Recall the definitions (2.2) and (2.3).

Lemma 3. *Assume conditions* (B1)–(B5), (M). *Then for any* $\boldsymbol{x} \in [0, 1]^d$ *we have*

$$\lim_{n \to \infty} |F_n(\boldsymbol{x}) - \widetilde{F}_n(\boldsymbol{x})| = 0.$$

Proof. Since $\boldsymbol{B}_n = \nabla \boldsymbol{G}_n(\mathbf{1})$ we have

$$\nabla \overline{\boldsymbol{G}}_{j+1,n}(\mathbf{1}) = \nabla \boldsymbol{G}_{j+1}(\mathbf{1}) \nabla \boldsymbol{G}_{j+2}(\mathbf{1}) \dots \nabla \boldsymbol{G}_n(\mathbf{1}) = \boldsymbol{B}_{[j,n]}.$$

By the scalar version of (4.1) we have

$$\begin{aligned}
|F_n(\boldsymbol{x}) - \widetilde{F}_n(\boldsymbol{x})| &\leq \sum_{j=1}^{n} \left| \mathrm{e}^{H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}))-1} - 1 - (H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1) \right| \\
&\leq \sum_{j=1}^{n} \left(H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1 \right)^2 \\
&\leq \sum_{j=1}^{n} \left((\mathbf{1} - \overline{\boldsymbol{G}}_{[j+1,n]}(\boldsymbol{x})) \boldsymbol{m}_j^\top \right)^2 \\
&\leq \sum_{j=1}^{n} \left((\mathbf{1} - \boldsymbol{x}) \boldsymbol{B}_{[j,n]}^\top \boldsymbol{m}_j^\top \right)^2.
\end{aligned}$$

Since

$$\left| (\mathbf{1} - \boldsymbol{x}) \boldsymbol{B}_{[j,n]}^\top \boldsymbol{m}_j^\top \right| \leq \|\mathbf{1}\| \cdot \|\boldsymbol{A}_{j,n}\| \max_{k \geq 1} \|\boldsymbol{m}_k (\boldsymbol{B} - \boldsymbol{B}_k)^{-1}\|,$$

by (2.6), (2.8) and (M) the sum above converges to 0, as stated. □

Proof of Theorem 1. Since the generating function of $\mathrm{Po}(\boldsymbol{\lambda}) = \mathrm{Po}(\lambda_1) \times \dots \times \mathrm{Po}(\lambda_d)$ has the form

$$\mathrm{e}^{\lambda_1(x_1-1)} \dots \mathrm{e}^{\lambda_d(x_d-1)} = \mathrm{e}^{(\boldsymbol{x}-\mathbf{1})\boldsymbol{\lambda}^\top}, \qquad \boldsymbol{x} \in [0,1]^d,$$

by Lemma 3 we only have to show that

$$\sum_{j=1}^{n} \boldsymbol{m}_j \boldsymbol{B}_{[j,n]} \to \boldsymbol{\lambda} \boldsymbol{A},$$

for all $\boldsymbol{x} \in [0,1]^d$. This holds according to Lemma 1 and our assumption. □

Since $D^{\boldsymbol{j}} H_n(\mathbf{1}) = m_{n,\boldsymbol{j}}$, the multivariate Taylor expansion gives the following.

Lemma 4. *If $\mathsf{E}(\|\boldsymbol{\varepsilon}_n\|^k) < \infty$ for some $k \in \mathbb{N}$ then for all $\boldsymbol{x} \in [0,1]^d$*

$$\begin{aligned}
H_n(\boldsymbol{x}) &= \sum_{\boldsymbol{\ell} \in \mathbb{Z}_+^d,\, |\boldsymbol{\ell}| < k} \frac{m_{n,\boldsymbol{\ell}}}{\boldsymbol{\ell}!} (\boldsymbol{x} - \mathbf{1})^{\boldsymbol{\ell}} + R_{n,k}(\boldsymbol{x}) \\
&= 1 + \sum_{i=1}^{k-1} \frac{1}{i!} \sum_{\ell_1,\dots,\ell_i=1}^{d} \frac{\partial^i H_n(\mathbf{1})}{\partial x_{\ell_1} \dots \partial x_{\ell_i}} (x_{\ell_1} - 1) \dots (x_{\ell_i} - 1) + R_{n,k}(\boldsymbol{x}),
\end{aligned}$$

where $\boldsymbol{\ell}! := \ell_1! \dots \ell_d!$ *for* $\boldsymbol{\ell} = (\ell_1, \dots, \ell_d) \in \mathbb{Z}_+^d$, *and*

$$|R_{n,k}(\boldsymbol{x})| \le \sum_{\boldsymbol{\ell} \in \mathbb{Z}_+^d,\, |\boldsymbol{\ell}|=k} \frac{m_{n,\boldsymbol{\ell}}}{\boldsymbol{\ell}!} (\mathbf{1} - \boldsymbol{x})^{\boldsymbol{\ell}} \le d^k \|\mathbf{1} - \boldsymbol{x}\|^k \max_{|\boldsymbol{\ell}|=k} D^{\boldsymbol{\ell}} H_n(\mathbf{1}).$$

Proof of Theorem 2. Since the offsprings are Bernoulli distributed

$$\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) = \mathbf{1} + (\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top,$$

therefore by Lemma 3 it is enough to show that the convergence

$$\sum_{j=1}^n \left[H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1 \right] \to \sum_{i=1}^{k-1} \sum_{\ell_{i+1},\dots,\ell_{2i}=1}^d \Lambda_{i;\ell_{i+1},\dots,\ell_{2i}} (x_{\ell_{i+1}} - 1) \dots (x_{\ell_{2i}} - 1)$$

holds. Using Lemma 4 we may write

$$\begin{aligned}\sum_{j=1}^n \left[H_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) - 1 \right] &= \sum_{j=1}^n \sum_{i=1}^{k-1} \frac{1}{i!} \sum_{\ell_1,\dots,\ell_i=1}^d \frac{\partial^i H_j(\mathbf{1})}{\partial x_{\ell_1} \dots \partial x_{\ell_i}} \big((\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top \big)_{\ell_1} \dots \big((\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top \big)_{\ell_i} \\ &\quad + \sum_{j=1}^n R_{j,k} \big(\mathbf{1} + (\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top \big).\end{aligned}$$

Since, for $m \in \{1, \dots, i\}$

$$\Big((\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top \Big)_{\ell_m} = \sum_{\ell_{i+m}=1}^d \big(\boldsymbol{B}_{[j,n]} \big)_{\ell_m, \ell_{i+m}} (x_{\ell_{i+m}} - 1),$$

by (2.10) the first term converges for any $i \in \{1, \dots, k-1\}$,

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{i!} \sum_{j=1}^n \sum_{\ell_1,\dots,\ell_i=1}^d \frac{\partial^i H_j(\mathbf{1})}{\partial x_{\ell_1} \dots \partial x_{\ell_i}} ((\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top)_{\ell_1} \dots ((\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top)_{\ell_i} \\ &= \sum_{\ell_{i+1},\dots,\ell_{2i}=1}^d \Lambda_{i;\ell_{i+1},\dots,\ell_{2i}} (x_{\ell_{i+1}} - 1) \dots (x_{\ell_{2i}} - 1).\end{aligned}$$

Using Lemma 4 for the second term we have

$$\begin{aligned}\sum_{j=1}^n |R_{j,k}(\mathbf{1} + (\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top)| &\le \sum_{j=1}^n d^k \max_{|\boldsymbol{\ell}|=k} D^{\boldsymbol{\ell}} H_j(\mathbf{1}) \|\boldsymbol{B}_{[j,n]}\|^k \\ &\le d^k \sum_{j=1}^n \max_{|\boldsymbol{\ell}|=k} D^{\boldsymbol{\ell}} H_j(\mathbf{1}) \|(\boldsymbol{B} - \boldsymbol{B}_j)^{-1}\| \cdot \|\boldsymbol{A}_{j,n}\|,\end{aligned}$$

which goes to 0, due to (2.9). □

4.2 Proofs for Section 3

We start with the case when the limit matrix is the identity.

Lemma 5. *For any $d \geq 1$ there exists positive constant C_d such that*

$$\sum_{j=1}^{n} \|\boldsymbol{A}_j\| \leq C_d \left\| \sum_{j=1}^{n} \boldsymbol{A}_j \right\|$$

for all $n \in \mathbb{N}$ and for all matrices $\boldsymbol{A}_j \in \mathbb{R}_+^{d \times d}$, $j \in \{1, \ldots, n\}$.

Proof. The norms of a finite dimensional vector space are equivalent, hence there are positive constants $c_d, \widetilde{c}_d$ such that

$$c_d \sum_{i=1}^{d} \sum_{k=1}^{d} |a_{i,k}| \leq \|\boldsymbol{A}\| \leq \widetilde{c}_d \sum_{i=1}^{d} \sum_{k=1}^{d} |a_{i,k}|$$

for all matrices $\boldsymbol{A} = (a_{i,j})_{i,j \in \{1,\ldots,d\}} \in \mathbb{R}^{d \times d}$. Put $(\boldsymbol{A}_j)_{i,k} = a_{j;i,k}$. Consequently,

$$\begin{aligned} \sum_{j=1}^{n} \|\boldsymbol{A}_j\| &\leq \widetilde{c}_d \sum_{j=1}^{n} \sum_{i=1}^{d} \sum_{k=1}^{d} |a_{j;i,k}| = \widetilde{c}_d \sum_{j=1}^{n} \sum_{i=1}^{d} \sum_{k=1}^{d} a_{j;i,k} \\ &= \widetilde{c}_d \sum_{i=1}^{d} \sum_{k=1}^{d} \sum_{j=1}^{n} a_{j;i,k} = \widetilde{c}_d \sum_{i=1}^{d} \sum_{k=1}^{d} \left| \sum_{j=1}^{n} a_{j;i,k} \right| \leq \frac{\widetilde{c}_d}{c_d} \left\| \sum_{j=1}^{n} \boldsymbol{A}_j \right\|. \end{aligned}$$

□

Proof of Proposition 1. Condition (2.6) follows from (I1), (I2) and (I3), as in the general case. As we already mentioned (2.7) is automatic, since

$$\sum_{j=1}^{n} \boldsymbol{A}_{j,n} = \sum_{j=1}^{n} \left(\boldsymbol{B}_{[j,n]} - \boldsymbol{B}_{[j-1,n]} \right) = \boldsymbol{I} - \boldsymbol{B}_{[0,n]} \to \boldsymbol{I} \qquad \text{as } n \to \infty$$

by condition (I3).

If

$$C := \sup_{j \geq 1} \frac{\|\boldsymbol{I} - \boldsymbol{B}_j\|}{1 - \|\boldsymbol{B}_j\|} < \infty$$

then for all $n \geq j \geq 1$,

$$\|\boldsymbol{A}_{j,n}\| \leq \|\boldsymbol{I} - \boldsymbol{B}_j\| \cdot \|\boldsymbol{B}_{[j,n]}\| \leq C(1 - \|\boldsymbol{B}_j\|) \, \|\boldsymbol{B}_{[j,n]}\|.$$

Since $\boldsymbol{B}_{[n,n]} = \boldsymbol{I}$ we have $\|\boldsymbol{B}_{[n,n]}\| = 1$, and

$$\|\boldsymbol{B}_{[j,n]}\| \le \prod_{k=j+1}^{n} \|\boldsymbol{B}_k\|$$

for all $n > j$, thus

$$\begin{aligned}\sum_{j=1}^{n} \|\boldsymbol{A}_{j,n}\| &\le C(1 - \|\boldsymbol{B}_n\|) + C\sum_{j=1}^{n-1}(1 - \|\boldsymbol{B}_j\|) \prod_{k=j+1}^{n} \|\boldsymbol{B}_k\| \\ &= C\left(1 - \prod_{k=1}^{n} \|\boldsymbol{B}_k\|\right) \le C,\end{aligned}$$

and we deduce (2.8).

Otherwise, if $\boldsymbol{A}_{j,n} \in \mathbb{R}_+^{d\times d}$ for all $n \ge 1$ and all $j \in \{1, \dots, n\}$ then by Lemma 5

$$\sum_{j=1}^{n} \|\boldsymbol{A}_{j,n}\| \le C_d \left\| \sum_{j=1}^{n} \boldsymbol{A}_{j,n} \right\|,$$

and (2.7) implies (2.8). □

Proof of Proposition 2. We have to check only (2.8), since (2.6) and (2.7) follow from conditions (I1)–(I3). In this case

$$\boldsymbol{B}_{[j,n]} = \boldsymbol{B}_{j+1} \dots \boldsymbol{B}_n = \boldsymbol{U} \operatorname{diag}\left(\varrho_{[j,n],1}, \dots, \varrho_{[j,n],d}\right) \boldsymbol{U}^\top,$$

where $\varrho_{[j,n],i} = \varrho_{j+1,i} \dots \varrho_{n,i}$. Using again that the norm of a normal element in a C^*-algebra equal to its spectral radius, we have

$$\begin{aligned}\|\boldsymbol{A}_{j,n}\| &= \|\boldsymbol{U}(\operatorname{diag}(\varrho_{[j,n],1}, \dots, \varrho_{[j,n],d}) - \operatorname{diag}(\varrho_{[j-1,n],1}, \dots, \varrho_{[j-1,n],d}))\boldsymbol{U}^\top\| \\ &= \| \operatorname{diag}((1 - \varrho_{j,1})\varrho_{[j,n],1}, \dots, (1 - \varrho_{j,d})\varrho_{[j,n],d})\| \\ &= \max_{1\le i\le d} (1 - \varrho_{j,i})\varrho_{[j,n],i}.\end{aligned}$$

Thus (2.8) follows from $\sum_{j=1}^{n}(1 - \varrho_{j,i})\varrho_{[j,n],i} = 1 - \varrho_{1,i} \dots \varrho_{n,i} \le 1$, $i \in \{1, \dots, d\}$. □

Next we turn to the proofs when $\boldsymbol{B}_n = \varrho_n \boldsymbol{B}$. A slight modification of the proof of Theorem 5.2.1 in Doob [5] gives

Lemma 6. *Assume that $(a_{j,n})$ satisfies (3.1), $\sum_{j=1}^{n-1} |a_{j+1,n} - a_{j,n}| \to 0$, and let $\boldsymbol{B}$ be a matrix such that $\|\boldsymbol{B}\| \le 1$. Then there exists a matrix $\boldsymbol{A}$, such that*

$$\lim_{n\to\infty} \sum_{j=1}^{n} a_{j,n} \boldsymbol{B}^j = \boldsymbol{A}.$$

Moreover $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{A} = \boldsymbol{A}^2$.

Proof. Since for every j the matrix $\|\boldsymbol{B}^j\| \leq 1$, the sequence $\sum_{j=1}^{n} a_{j,n}\boldsymbol{B}^j$ is bounded, so there is a subsequence n_k and a limit $\boldsymbol{A}$ such that

$$\sum_{j=1}^{n_k} a_{j,n_k}\boldsymbol{B}^j \to \boldsymbol{A} \quad \text{as } k \to \infty.$$

Multiplying by $\boldsymbol{B}$ we obtain

$$\sum_{j=1}^{n_k} a_{j,n_k}\boldsymbol{B}^{j+1} \to \boldsymbol{B}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{B} \quad \text{as } k \to \infty.$$

Writing n instead of n_k, the difference between the two limits is

$$\sum_{j=1}^{n} a_{j,n}\boldsymbol{B}^{j+1} - \sum_{j=1}^{n} a_{j,n}\boldsymbol{B}^j = a_{n,n}\boldsymbol{B}^{n+1} - a_{1,n}\boldsymbol{B} + \sum_{j=2}^{n}(a_{j-1,n} - a_{j,n})\boldsymbol{B}^j.$$

Using that $\boldsymbol{B}^j$ is bounded, $a_{1,n} \to 0, a_{n,n} \to 0$ and that $\sum_{j=1}^{n-1} |a_{j+1,n} - a_{j,n}| \to 0$ we obtain that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{A}$. And so

$$\left(\sum_{j=1}^{n} a_{j,n}\boldsymbol{B}^j\right)\boldsymbol{A} = \sum_{j=1}^{n} a_{j,n}\boldsymbol{A}$$

gives that for any other subsequential limit $\boldsymbol{C}$, $\boldsymbol{A}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{A} = \boldsymbol{A}$. Since the roles are interchangeable, we obtain that there is only one limit matrix, which is idempotent. □

Using the lemma above it is easy to prove Proposition 3.

Proof of Proposition 3. We only have to check that for $a_{j,n} = \varrho_{[j,n]}(1 - \varrho_j)$ the condition $\sum_{j=1}^{n-1} |a_{j+1,n} - a_{j,n}| \to 0$ satisfied. We have

$$a_{j+1,n} - a_{j,n} = \varrho_{[j+1,n]}\left[(1 - \varrho_{j+1}) - \varrho_{j+1}(1 - \varrho_j)\right],$$

thus

$$\sum_{j=1}^{n-1} |a_{j+1,n} - a_{j,n}| = \sum_{j=1}^{n-1} \frac{|(1 - \varrho_{j+1}) - \varrho_{j+1}(1 - \varrho_j)|}{1 - \varrho_{j+1}}\varrho_{[j+1,n]}(1 - \varrho_{j+1}),$$

which goes to 0, since

$$\frac{|(1 - \varrho_{n+1}) - \varrho_{n+1}(1 - \varrho_n)|}{1 - \varrho_{n+1}} = \left|1 - \varrho_{n+1}\frac{1 - \varrho_n}{1 - \varrho_{n+1}}\right| \to 0.$$

□

Proof of Theorem 6. To prove the theorem we only have to show that condition (2.7) holds for $\boldsymbol{A}_{j,n} = (\boldsymbol{B} - \boldsymbol{B}_j)\boldsymbol{B}_{[j,n]}$.

By the monotonicity assumptions

$$\boldsymbol{B}_{[j,n]} = \boldsymbol{B}_{j+1} \dots \boldsymbol{B}_n \le \varrho_{j+1}\boldsymbol{B} \dots \varrho_n \boldsymbol{B} = \varrho_{[j,n]}\boldsymbol{B}^{n-j}$$

and similarly

$$\boldsymbol{B}_{[j,n]} \ge \vartheta_{[j,n]}\boldsymbol{B}^{n-j}.$$

Keeping in mind that each element of $\boldsymbol{B} - \boldsymbol{B}_j$ is non-negative, we have

$$(1-\varrho_j)\vartheta_{[j,n]}\boldsymbol{B}^{n-j+1} \le \boldsymbol{A}_{j,n} \le (1-\vartheta_j)\varrho_{[j,n]}\boldsymbol{B}^{n-j+1}.$$

After summation

$$\sum_{j=1}^{n} \boldsymbol{B}^{n-j+1}\vartheta_{[j,n]}(1-\varrho_j) \le \sum_{j=1}^{n} \boldsymbol{A}_{j,n} \le \sum_{j=1}^{n} \boldsymbol{B}^{n-j+1}\varrho_{[j,n]}(1-\vartheta_j). \tag{4.2}$$

First we show that the sequences $(\vartheta_{[j,n]}(1-\varrho_j))$ and $(\varrho_{[j,n]}(1-\vartheta_j))$ satisfy conditions (3.1). According to the assumptions

$$\sum_{j=1}^{n} \vartheta_{[j,n]}(1-\vartheta_j) \to 1 \quad \text{and} \quad \sum_{j=1}^{n} \varrho_{[j,n]}(1-\varrho_j) \to 1. \tag{4.3}$$

Since $\varrho_n \ge \vartheta_n$ we have

$$0 \le \sum_{j=1}^{n} \vartheta_{[j,n]}(\varrho_j - \vartheta_j) = \sum_{j=1}^{n} \frac{\varrho_j - \vartheta_j}{1-\vartheta_j}(1-\vartheta_j)\vartheta_{[j,n]} \to 0, \tag{4.4}$$

as

$$\frac{\varrho_j - \vartheta_j}{1-\vartheta_j} \le \frac{\varrho_j - \vartheta_j}{1-\varrho_j} \to 0.$$

Similarly

$$0 \le \sum_{j=1}^{n} \varrho_{[j,n]}(\varrho_j - \vartheta_j) = \sum_{j=1}^{n} \frac{\varrho_j - \vartheta_j}{1-\varrho_j}(1-\varrho_j)\varrho_{[j,n]} \to 0. \tag{4.5}$$

Noting that $\vartheta_{[j,n]}(1-\varrho_j) = \vartheta_{[j,n]}[(1-\vartheta_j) - (\varrho_j - \vartheta_j)]$ and $\varrho_{[j,n]}(1-\vartheta_j) = \varrho_{[j,n]}[(1-\varrho_j) + (\varrho_j - \vartheta_j)]$, (4.3) combined with (4.4) and with (4.5) shows that conditions (3.1) indeed hold.

When the convergence $\boldsymbol{B}^n \to \boldsymbol{A}$ holds, both the upper and the lower estimation in (4.2) tends to $\boldsymbol{A}$, and the statement follows.

In case (b) the extra condition assures the convergence of the bounds in (4.2) by Lemma 6, and the equality of the limits readily follows. □

Proof of Theorem 3. By Lemma 3 we have to check that

$$\sum_{j=1}^{n} (\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) - \boldsymbol{1})\boldsymbol{m}_j^\top \to (\boldsymbol{x} - \boldsymbol{1})\boldsymbol{\lambda}^\top.$$

By (4.1) we have

$$\mathbf{1} - \overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) \le (\mathbf{1} - \boldsymbol{x}) \nabla \overline{\boldsymbol{G}}_{j+1,n}(\mathbf{1})^\top = (\mathbf{1} - \boldsymbol{x}) \boldsymbol{B}_{[j,n]}^\top, \tag{4.6}$$

therefore

$$\sum_{j=1}^{n} (\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) - \mathbf{1}) \boldsymbol{m}_j^\top \ge \sum_{j=1}^{n} (\boldsymbol{x} - \mathbf{1}) \boldsymbol{B}_{[j,n]}^\top \boldsymbol{m}_j^\top \to (\boldsymbol{x} - \mathbf{1}) \boldsymbol{\lambda}^\top,$$

where the last convergence holds under the assumptions of the theorem.

According to (4.6) $\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) \in [\mathbf{1} - \mathbf{1}\boldsymbol{B}_{[j,n]}^\top, \mathbf{1}]$, for all $\boldsymbol{x} \in [0,1]^d$. Again by the mean value theorem and by the monotonicity of the derivatives

$$\mathbf{1} - \boldsymbol{G}_j(\boldsymbol{y}) \ge (\mathbf{1} - \boldsymbol{y}) \, \nabla \boldsymbol{G}_j(\mathbf{1} - \mathbf{1}\boldsymbol{B}_{[j,n]}^\top)^\top =: (\mathbf{1} - \boldsymbol{y}) \, \boldsymbol{\Theta}_{j,n}^\top,$$

for $\boldsymbol{y} \in [\mathbf{1} - \mathbf{1}\boldsymbol{B}_{[j,n]}^\top, \mathbf{1}]$, in particular

$$\mathbf{1} - \boldsymbol{G}_j(\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) \ge (\mathbf{1} - \overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x})) \, \boldsymbol{\Theta}_{j,n}^\top,$$

and so induction gives

$$\mathbf{1} - \overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) \ge (\mathbf{1} - \boldsymbol{x}) \boldsymbol{\Theta}_{n,n}^\top \boldsymbol{\Theta}_{n-1,n}^\top \dots \boldsymbol{\Theta}_{j+1,n}^\top =: (\mathbf{1} - \boldsymbol{x}) \boldsymbol{\Theta}_{[j,n]}^\top,$$

so

$$\sum_{j=1}^{n} (\overline{\boldsymbol{G}}_{j+1,n}(\boldsymbol{x}) - \mathbf{1}) \boldsymbol{m}_j^\top \le \sum_{j=1}^{n} (\boldsymbol{x} - \mathbf{1}) \boldsymbol{\Theta}_{[j,n]}^\top \boldsymbol{m}_j^\top.$$

We have to check under what conditions

$$\sum_{j=1}^{n} \boldsymbol{m}_j \boldsymbol{\Theta}_{[j,n]} \to \boldsymbol{\lambda}.$$

Clearly $\boldsymbol{\Theta}_{j,n} \uparrow \boldsymbol{B}_j$ as $n \to \infty$. Introduce $\boldsymbol{C}_{j,n} = (\boldsymbol{I} - \boldsymbol{B}_j) \boldsymbol{\Theta}_{[j,n]}$. Since by definition the elements of $\boldsymbol{C}_{j,n}$ are less then or equal to the elements of $\boldsymbol{A}_{j,n}$, so the only assumption we have to check in order to guarantee the convergence above is

$$\sum_{j=1}^{n} \boldsymbol{C}_{j,n} \to \boldsymbol{I}.$$

We have

$$\begin{aligned} \sum_{j=1}^{n} \boldsymbol{C}_{j,n} &= (\boldsymbol{I} - \boldsymbol{B}_1) \boldsymbol{\Theta}_{[1,n]} + (\boldsymbol{I} - \boldsymbol{B}_2) \boldsymbol{\Theta}_{[2,n]} + \dots + (\boldsymbol{I} - \boldsymbol{B}_{n-1}) \boldsymbol{\Theta}_{[n-1,n]} + \boldsymbol{I} - \boldsymbol{B}_n \\ &= \boldsymbol{I} - \Big[(\boldsymbol{B}_n - \boldsymbol{\Theta}_{n,n}) + (\boldsymbol{B}_{n-1} - \boldsymbol{\Theta}_{n-1,n}) \boldsymbol{\Theta}_{[n-1,n]} + \dots + (\boldsymbol{B}_2 - \boldsymbol{\Theta}_{2,n}) \boldsymbol{\Theta}_{[2,n]} + \boldsymbol{B}_1 \boldsymbol{\Theta}_{[1,n]} \Big]. \end{aligned}$$

We show that the sum in the brackets (note that every term is nonnegative) converge to the 0 matrix. Let us estimate the (i,k)-th element of $\boldsymbol{B}_j - \boldsymbol{\Theta}_{j,n}$. The mean value theorem and the monotonicity of the derivatives imply

$$\begin{aligned}(\boldsymbol{B}_j - \boldsymbol{\Theta}_{j,n})_{i,k} &= \frac{\partial}{\partial x_k} G_{j,i}(\mathbf{1}) - \frac{\partial}{\partial x_k} G_{j,i}(\mathbf{1} - \mathbf{1}\boldsymbol{B}_{[j,n]}^\top) \\ &\le (\mathbf{1}\boldsymbol{B}_{[j,n]}^\top)\left(\frac{\partial^2}{\partial x_k \partial x_1} G_{j,i}(\mathbf{1}), \ldots, \frac{\partial^2}{\partial x_k \partial x_d} G_{j,i}(\mathbf{1})\right)^\top,\end{aligned}$$

thus

$$(\boldsymbol{B}_j - \boldsymbol{\Theta}_{j,n})_{i,k} \le d\,\|\mathbf{1}\|\, m_2(j).$$

So finally we obtain

$$\sum_{j=1}^{n} (\boldsymbol{B}_j - \boldsymbol{\Theta}_{j,n})\, \boldsymbol{\Theta}_{[j,n]} \le d\,\|\mathbf{1}\| \sum_{j=1}^{n} m_2(j) \begin{bmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{bmatrix} \boldsymbol{B}_{[j,n]},$$

which goes to the 0 matrix, whenever $\|(\boldsymbol{I} - \boldsymbol{B}_n)^{-1}\|\, m_2(n) \to 0$. □

Acknowledgement

The work of the first author was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TÁMOP-4.2.2.C-11/1/KONV-2012-0013). The second, third and fourth authors have been partially supported by the Hungarian Chinese Intergovernmental S & T Cooperation Programme 2011–2013. The second author has been 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. The third and fourth author have been supported by the European Union and co-funded by the European Social Fund under the project 'Telemedicine-focused research activities on the field of Mathematics, Informatics and Medical sciences' of project number 'TÁMOP-4.2.2.A-11/1/KONV-2012-0073'. The third author's research was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program Elaborating and operating an inland student and researcher personal support system'. The project was subsidized by the European Union and co-financed by the European Social Fund. The third author was partially supported by the Hungarian Scientific Research Fund OTKA PD106181. The fourth author has been supported by the Hungarian Scientific Research Fund under Grant No. OTKA T-079128.

References

[1] Altman, E. and Fiems, D. (2012). Branching processes and their generalization applied to wireless networking. In *Paradigms for Biologically- Inspired Autonomic Networks and Services. The BIONETS Project eBook*, Eitan Altman, Paolo Dini, Daniele Miorandi, eds. pp. 54–66.

[2] Boon, M.A.A. (2012). A polling model with reneging at polling instants. *Annals of Operations Research*, 198:5–23.

[3] Boon, M.A.A., Adan, I.J.B.F. and Boxma, O.J. (2010). A polling model with multiple priority levels. *Performance Evaluation*, 67:468–484.

[4] Böckenholt, U. (1999). Mixed INAR(1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data. *Journal of Econometrics*, 89:317–338.

[5] Doob, J.L. (1990). *Stochastic Processes*. Wiley.

[6] Du, J.G. and Li, Y. (1991). The integer valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.*, 12:129–142.

[7] Franke, J. and Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*, Math. Dep. UMIST, England.

[8] Fritz, J. (1975). Learning from an ergodic training sequence. In: *Limit theorems of probability theory (Colloq., Keszthely, 1974)*, pp. 79–91. Colloq. Math. Soc. János Bolyai, Vol. **11**, North-Holland, Amsterdam.

[9] Gladstien, K. and Lange, K. (1978). Equilibrium distributions for deleterious genes in large stationary populations. *Theoretical Population Biology*, 14:322–328.

[10] Gourieroux, C. and Jasiak, J. (2004). Heterogeneous INAR(1) model with application to car insurance. *Insurance: Mathematics and Economics*, 34:177–192.

[11] Györfi, L., Ispány, M., Pap G. and Varga, K. (2007). Poisson limit of an inhomogeneous nearly critical INAR(1) model. *Acta Sci. Math. (Szeged)*, **73** (3-4), 789–815.

[12] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. John Wiley & Sons, Inc., New York.

[13] Haccou, P., Jagers, P. and Vatutin, V.A. (2007). *Branching processes: variation, growth, and extinction of populations*. Cambridge University Press, Cambridge.

[14] Kaplan, N. (1973). The multitype Galton-Watson process with immigration. *Ann. Probab.* Vol 1, Number 6, 947–953.

[15] Kevei, P. (2012). Asymptotics of nearly critical Galton–Watson process with immigration. *Acta Sci. Math. (Szeged)*, **77**, 681–702.

[16] Kocherlakota, S. and Kocherlakota, K. (1992). *Bivariate Discrete Distributions.* Dekker, New York.

[17] Krummenauer, F. (1998). Limit theorems for multivariate discrete distributions. *Metrika*, 47:47–69.

[18] Lange, K. and Fan, R. (1997). Branching process models for mutant genes in nonstationary populations. *Theoretical Population Biology*, 51:118–133.

[19] van der Mei, R.D. (2007). Towards a unifying theory on branching-type polling systems in heavy traffic. *Queueing Systems*, 57:29-46.

[20] Mode, C.J. (1971). *Multitype Branching Processes. Theory and Applications.* Elsevier, New York.

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

[22] Quine, M.P. (1972). The multi-type Galton–Watson process with ϱ near to 1. *Adv. Appl. Prob.* 4:429–452.

[23] Resing, J.A.C. (1993). Polling systems and multitype branching processes. *Queueing Systems*, 13:409-426.

[24] Weiß, Ch. H. (2008). Thinning operations for modeling time series of counts — a survey. *AStA Adv. Stat. Anal.*, 92:319–341.