On tail behaviour of stationary second-order Galton–Watson processes with immigration

Mátyás Barczy*,◇, Zsuzsanna Bősze**, Gyula Pap***

* MTA-SZTE Analysis and Stochastics Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

** Institute for Mathematical Stochastics, Georg-August-Universität Göttingen, Goldschmidtstr. 7, 37077 Göttingen, Germany.

*** Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H-6720 Szeged, Hungary.

e-mail: barczy@math.u-szeged.hu (M. Barczy), zsuzsanna.boesze@uni-goettingen.de (Zs. Bősze).

◇ Corresponding author.

Abstract

A second-order Galton–Watson process with immigration can be represented as a coordinate process of a 2-type Galton–Watson process with immigration. Sufficient conditions are derived on the offspring and immigration distributions of a second-order Galton–Watson process with immigration under which the corresponding 2-type Galton–Watson process with immigration has a unique stationary distribution such that its common marginals are regularly varying. In the course of the proof sufficient conditions are given under which the distribution of a second-order Galton–Watson process (without immigration) at any fixed time is regularly varying provided that the initial sizes of the population are independent and regularly varying.

1 Introduction

Branching processes have been frequently used in biology, e.g., for modeling the spread of an infectious disease, for gene amplification and deamplification or for modeling telomere shortening, see, e.g., Kimmel and Axelrod [18]. Higher-order Galton–Watson processes with immigration having finite second moment (also called Generalized Integer-valued AutoRegressive (GINAR) processes) have been introduced by Latour [19, equation (1.1)]. Pénisson and Jacob [21] used higher-order Galton–Watson processes (without immigration) for studying the decay phase of an

2010 Mathematics Subject Classifications: 60J80, 60G70.

Key words and phrases: second-order Galton–Watson process with immigration, regularly varying distribution, tail behavior.

Supported by the Hungarian Croatian Intergovernmental S & T Cooperation Programme for 2017-2018 under Grant No. 16-1-2016-0027. Mátyás Barczy is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

epidemic, and, as an application, they investigated the Bovine Spongiform Encephalopathy epidemic in Great Britain after the 1988 feed ban law. As a continuation, Pénisson [20] introduced estimators of the so-called infection parameter in the growth and decay phases of an epidemic. Recently, Kashikar and Deshmukh [16, 17] and Kashikar [15] used second order Galton–Watson processes (without immigration) for modeling the swine flu data for Pune, India and La-Gloria, Mexico. Kashikar and Deshmukh [16] also studied their basic probabilistic properties such as a formula for their probability generator function, probability of extinction, long run behavior and conditional least squares estimation of the offspring means. Higher-order Galton–Watson processes with immigration are special multi-type Galton–Watson processes with immigration, and to give an example for an application of such processes for modeling epidemics, for example, we can mention Dénes et al. [7], where a 17-type Galton–Watson process with immigration has been applied to describe the risk of a major epidemic in connection with the 2012 UEFA European Football Championship took place in Ukraine and Poland between 8 June and 1 July 2012.

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$, and $\mathbb{R}_{--}$ denote the set of non-negative integers, positive integers, real numbers, non-negative real numbers, positive real numbers and negative real numbers, respectively. For functions $f : \mathbb{R}_{++} \to \mathbb{R}_{++}$ and $g : \mathbb{R}_{++} \to \mathbb{R}_{++}$, by the notation $f(x) \sim g(x)$, $f(x) = \mathrm{o}(g(x))$ and $f(x) = \mathrm{O}(g(x))$ as $x \to \infty$, we mean that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$, $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$ and $\limsup_{x\to\infty} \frac{f(x)}{g(x)} < \infty$, respectively. The natural basis of $\mathbb{R}^d$ will be denoted by $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$. For $x \in \mathbb{R}$, the integer part of x is denoted by $\lfloor x \rfloor$. Every random variable will be defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Equality in distributions of random variables or stochastic processes is denoted by $\stackrel{\mathcal{D}}{=}$.

First, we recall the Galton–Watson process with immigration, which assumes that an individual can reproduce only once during its lifetime at age 1, and then it dies immediately. The initial population size at time 0 will be denoted by X_0. For each $n \in \mathbb{N}$, the population consists of the offsprings born at time n and the immigrants arriving at time n. For each $n, i \in \mathbb{N}$, the number of offsprings produced at time n by the i^{th} individual of the $(n-1)^{\mathrm{th}}$ generation will be denoted by $\xi_{n,i}$. The number of immigrants in the n^{th} generation will be denoted by ε_n. Then, for the population size X_n of the n^{th} generation, we have

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_{n,i} + \varepsilon_n, \qquad n \in \mathbb{N}, \tag{1.1}$$

where $\sum_{i=1}^{0} := 0$. Here $\big\{X_0,\, \xi_{n,i},\, \varepsilon_n : n, i \in \mathbb{N}\big\}$ are supposed to be independent non-negative integer-valued random variables, and $\{\xi_{n,i} : n, i \in \mathbb{N}\}$ and $\{\varepsilon_n : n \in \mathbb{N}\}$ are supposed to consist of identically distributed random variables, respectively. If $\varepsilon_n = 0$, $n \in \mathbb{N}$, then we say that $(X_n)_{n\in\mathbb{Z}_+}$ is a Galton–Watson process (without immigration).

Next, we introduce the second-order Galton–Watson branching model with immigration. In this model we suppose that an individual reproduces at age 1 and also at age 2, and then it dies immediately. For each $n \in \mathbb{N}$, the population consists again of the offsprings born at time n and the immigrants arriving at time n. For each $n, i, j \in \mathbb{N}$, the number of offsprings produced at time n by the i^{th} individual of the $(n-1)^{\text{th}}$ generation and by the j^{th} individual of the $(n-2)^{\text{nd}}$ generation will be denoted by $\xi_{n,i}$ and $\eta_{n,j}$, respectively, and ε_n denotes the number of immigrants in the n^{th} generation. Then, for the population size X_n of the n^{th} generation, we have

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_{n,i} + \sum_{j=1}^{X_{n-2}} \eta_{n,j} + \varepsilon_n, \qquad n \in \mathbb{N}, \tag{1.2}$$

where X_{-1} and X_0 are non-negative integer-valued random variables (the initial population sizes). Here $\big\{X_{-1}, X_0, \xi_{n,i}, \eta_{n,j}, \varepsilon_n : n, i, j \in \mathbb{N}\big\}$ are supposed to be non-negative integer-valued random variables such that $\big\{(X_{-1}, X_0), \xi_{n,i}, \eta_{n,j}, \varepsilon_n : n, i, j \in \mathbb{N}\big\}$ are independent, and $\{\xi_{n,i} : n, i \in \mathbb{N}\}$, $\{\eta_{n,j} : n, j \in \mathbb{N}\}$ and $\{\varepsilon_n : n \in \mathbb{N}\}$ are supposed to consist of identically distributed random variables, respectively. Note that the number of individuals alive at time $n \in \mathbb{Z}_+$ is $X_n + X_{n-1}$, which can be larger than the population size X_n of the n^{th} generation, since the individuals of the population at time $n-1$ are still alive at time n, because they can reproduce also at age 2. The stochastic process $(X_n)_{n\geqslant -1}$ given by (1.2) is called a second-order Galton–Watson process with immigration or a Generalized Integer-valued AutoRegressive process of order 2 (GINAR(2) process), see, e.g., Latour [19]. Especially, if $\xi_{1,1}$ and $\eta_{1,1}$ are Bernoulli distributed random variables, then $(X_n)_{n\geqslant -1}$ is also called an Integer-valued AutoRegressive process of order 2 (INAR(2) process), see, e.g., Du and Li [8]. If $\varepsilon_1 = 0$, then we say that $(X_n)_{n\geqslant -1}$ is a second-order Galton–Watson process without immigration, introduced and studied by Kashikar and Deshmukh [16] as well.

The process given in (1.2) with the special choice $\eta_{1,1} = 0$ gives back the process given in (1.1), which will be called a first-order Galton–Watson process with immigration to make a distinction.

For notational convenience, let ξ, η and ε be random variables such that $\xi \stackrel{\mathcal{D}}{=} \xi_{1,1}$, $\eta \stackrel{\mathcal{D}}{=} \eta_{1,1}$ and $\varepsilon \stackrel{\mathcal{D}}{=} \varepsilon_1$, and put $m_\xi := \mathbb{E}(\xi) \in [0,\infty]$, $m_\eta := \mathbb{E}(\eta) \in [0,\infty]$ and $m_\varepsilon := \mathbb{E}(\varepsilon) \in [0,\infty]$.

If $(X_n)_{n\in\mathbb{Z}_+}$ is a (first-order) Galton–Watson process with immigration such that $m_\xi \in (0,1)$, $\mathbb{P}(\varepsilon = 0) < 1$, and $\sum_{j=1}^\infty \mathbb{P}(\varepsilon = j)\log(j) < \infty$, then the Markov process $(X_n)_{n\in\mathbb{Z}_+}$ admits a unique stationary distribution μ, see, e.g., Quine [22]. If ε is regularly varying with index $\alpha \in \mathbb{R}_{++}$, i.e., $\mathbb{P}(\varepsilon > x) \in \mathbb{R}_{++}$ for all $x \in \mathbb{R}_{++}$, and

$$\lim_{x\to\infty} \frac{\mathbb{P}(\varepsilon > qx)}{\mathbb{P}(\varepsilon > x)} = q^{-\alpha} \qquad \text{for all } q \in \mathbb{R}_{++},$$

then, by Lemma E.5, $\sum_{j=1}^\infty \mathbb{P}(\varepsilon = j)\log(j) < \infty$. The content of Theorem 2.1.1 in Basrak et al. [3] is the following statement.

1.1 Theorem. *Let $(X_n)_{n\in\mathbb{Z}_+}$ be a (first-order) Galton–Watson process with immigration such that $m_\xi \in (0,1)$ and ε is regularly varying with index $\alpha \in (0,2)$. In case of $\alpha \in [1,2)$, assume additionally that $\mathbb{E}(\xi^2) < \infty$. Then the tail of the unique stationary distribution μ*

of $(X_n)_{n\in\mathbb{Z}_+}$ *satisfies*

$$\mu((x,\infty)) \sim \sum_{i=0}^{\infty} m_\xi^{i\alpha}\, \mathbb{P}(\varepsilon > x) = \frac{1}{1-m_\xi^\alpha}\, \mathbb{P}(\varepsilon > x) \qquad \text{as } x \to \infty,$$

and hence μ *is also regularly varying with index* α*.*

Note that in case of $\alpha = 1$ and $m_\varepsilon = \infty$ Basrak et al. [3, Theorem 2.1.1] assume additionally that ε is consistently varying (or in other words intermediate varying), but, eventually, it follows from the fact that ε is regularly varying. Basrak et al. [3, Remark 2.2.2] derived the result of Theorem 1.1 also for $\alpha \in [2,3)$ under the additional assumption $\mathbb{E}(\xi^3) < \infty$ (not mentioned in the paper), and they remark that the same applies to all $\alpha \in [3,\infty)$ (possibly under an additional moment assumption $\mathbb{E}(\xi^{\lfloor\alpha\rfloor+1}) < \infty$).

In Barczy et al. [2] we study regularly varying non-stationary (first-order) Galton–Watson processes with immigration.

As the main result of the paper, in Theorem 2.1, in the same spirit as in Theorem 1.1, we present sufficient conditions on the offspring and immigration distributions of a second-order Galton–Watson process with immigration under which its associated 2-type Galton–Watson process with immigration has a unique stationary distribution such that its common marginals are regularly varying. According to our knowledge, such a result has not been established so far, e.g., we could not find any reference which would address regularly varying GINAR(2) processes. Our result and the applied technique might be extended to a p-th order Galton–Watson branching process with immigration, however such an extension is not immediate, for example, it is not clear what would replace the constant $\sum_{i=0}^{\infty} m_i^\alpha$ in Theorem 2.1. More generally, one can pose an open problem, namely, under what conditions on the offspring and immigration distributions of a general p-type Galton–Watson branching process with immigration, its unique (p-dimensional) stationary distribution is jointly regularly varying. We also note that there is a vast literature on tail behavior of regularly varying time series (see, e.g., Hult and Samorodnitsky [12]), however, the available results do not seem to be applicable for describing the tail behavior of the stationary distribution for regularly varying branching processes. The link between GINAR and autoregressive processes is that their autocovariance functions are identical under finite second moment assumptions, but we can not see that it would imply anything for the tail behavior of a GINAR process knowing the tail behaviour of a corresponding autoregressive process. Further, in our situation the second moment is infinite, so the autocovariance function is not defined.

Very recently, Bősze and Pap [5] have studied regularly varying non-stationary second-order Galton–Watson processes with immigration. They have found some sufficient conditions on the initial, the offspring and the immigration distributions of a non-stationary second-order Galton-Watson process with immigration under which the distribution of the process in question is regularly varying at any fixed time. The results in Bősze and Pap [5] can be considered as extensions of the results in Barczy et al. [2] on not necessarily stationary (first-order) Galton–Watson processes with immigration. Concerning the results in Bősze and Pap [5] and in the present paper, there is no overlap, for more details see Remark 2.2.

The paper is organized as follows. In Section 2, first, for a second-order Galton–Watson process with immigration, we give a representation of the unique stationary distribution and its marginals, respectively, then our main result, Theorem 2.1, is formulated. The rest of Section 2 is devoted to the proof of Theorem 2.1. In the course of the proof, we formulate an auxiliary result about the tail behaviour of a second-order Galton–Watson process (without immigration) with a regularly varying initial distribution at time 0 and with value 0 at time -1, see Proposition 2.3. We close the paper with seven appendices which are used throughout the proofs. In Appendix A, we recall a representation of a second-order Galton–Watson process without or with immigration as a (special) 2-type Galton–Watson process without or with immigration, respectively. In Appendix B, we derive an explicit formula for the expectation of a second-order Galton–Watson process with immigration at time n and describe its asymptotic behavior as $n \to \infty$. Appendix C is about the existence and estimation of higher order moments of a second-order Galton–Watson process (without immigration). In Appendix D, we recall a representation of the unique stationary distribution for a 2-type Galton–Watson process with immigration. In Appendix E, we collect several results on regularly varying functions and distributions, to name a few of them: convolution property, Karamata's theorem and Potter's bounds. Appendix F is devoted to recall and (re)prove a result on large deviations for sums of non-negative independent and identically distributed regularly varying random variables due to Tang and Yan [26, part (ii) of Theorem 1]. Finally, in Appendix G, we present a variant of Proposition 2.3, where the initial values X_{-1} and X_0 are independent and regularly varying together with a second type of proof, see Proposition G.1.

2 Tail behavior of the marginals of the stationary distribution of second-order Galton–Watson processes with immigration

Let $(X_n)_{n\geqslant -1}$ be a second order Galton–Watson process with immigration given in (1.2), and let us consider the Markov chain $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ given by

$$\boldsymbol{Y}_n := \begin{bmatrix} Y_{n,1} \\ Y_{n,2} \end{bmatrix} := \begin{bmatrix} X_n \\ X_{n-1} \end{bmatrix} = \sum_{i=1}^{Y_{n-1,1}} \begin{bmatrix} \xi_{n,i} \\ 1 \end{bmatrix} + \sum_{j=1}^{Y_{n-1,2}} \begin{bmatrix} \eta_{n,j} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_n \\ 0 \end{bmatrix}, \qquad n \in \mathbb{N},$$

which is a (special) 2-type Galton–Watson process with immigration, and $(\boldsymbol{e}_1^\top \boldsymbol{Y}_k)_{k\in\mathbb{Z}_+} = (X_k)_{k\in\mathbb{Z}_+}$, $(\boldsymbol{e}_2^\top \boldsymbol{Y}_{k+1})_{k\geqslant -1} = (X_k)_{k\geqslant -1}$ (for more details, see Appendix A). If $m_\xi \in \mathbb{R}_{++}$, $m_\eta \in \mathbb{R}_{++}$, $m_\xi + m_\eta < 1$, $\mathbb{P}(\varepsilon = 0) < 1$ and $\mathbb{E}(\mathbb{1}_{\{\varepsilon\neq 0\}} \log(\varepsilon)) < \infty$, then there exists a unique stationary distribution $\boldsymbol{\pi}$ for $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}_+}$, see Appendix D, since then $\boldsymbol{M}_{\xi,\eta}$ is primitive due to the fact that

$$\boldsymbol{M}_{\xi,\eta}^2 = \begin{bmatrix} m_\xi & m_\eta \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} m_\xi^2 + m_\eta & m_\xi m_\eta \\ m_\xi & m_\eta \end{bmatrix} \in \mathbb{R}_{++}^2.$$

Moreover, the stationary distribution $\boldsymbol{\pi}$ of $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}_+}$ has a representation

$$\boldsymbol{\pi} \stackrel{\mathcal{D}}{=} \sum_{i=0}^{\infty} \boldsymbol{V}_i^{(i)}(\boldsymbol{\varepsilon}_i),$$

where $(\boldsymbol{V}_k^{(i)}(\boldsymbol{\varepsilon}_i))_{k\in\mathbb{Z}_+}$, $i \in \mathbb{Z}_+$, are independent copies of a (special) 2-type Galton–Watson process $(\boldsymbol{V}_k(\boldsymbol{\varepsilon}))_{k\in\mathbb{Z}_+}$ (without immigration) with initial vector $\boldsymbol{V}_0(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}$ and with the same offspring distributions as $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$, and the series $\sum_{i=0}^{\infty} \boldsymbol{V}_i^{(i)}(\boldsymbol{\varepsilon})$ converges with probability 1, see Appendix D. Using the considerations for the backward representation in Appendix A, we have $(\boldsymbol{e}_1^\top \boldsymbol{V}_k(\boldsymbol{\varepsilon}))_{k\in\mathbb{Z}_+} = (V_k(\varepsilon))_{k\in\mathbb{Z}_+}$ and $(\boldsymbol{e}_2^\top \boldsymbol{V}_{k+1}(\boldsymbol{\varepsilon}))_{k\geqslant -1} = (V_k(\varepsilon))_{k\geqslant -1}$, where $(V_k(\varepsilon))_{k\geqslant -1}$ is a second-order Galton–Watson process (without immigration) with initial values $V_0(\varepsilon) = \varepsilon$ and $V_{-1}(\varepsilon) = 0$, and with the same offspring distributions as $(X_k)_{k\geqslant -1}$. Consequently, the marginals of the stationary distribution $\boldsymbol{\pi}$ are the same distributions π, and it admits the representation

$$\pi \stackrel{\mathcal{D}}{=} \sum_{i=0}^{\infty} V_i^{(i)}(\varepsilon_i),$$

where $(V_k^{(i)}(\varepsilon_i))_{k\in\mathbb{Z}_+}$, $i \in \mathbb{Z}_+$, are independent copies of $(V_k(\varepsilon))_{k\geqslant -1}$. This follows also from the fact that the stationary distribution $\boldsymbol{\pi}$ is the limit in distribution of $\boldsymbol{Y}_n$ as $n \to \infty$ and

$$\boldsymbol{Y}_n = \begin{bmatrix} X_n \\ X_{n-1} \end{bmatrix}, \qquad n \in \mathbb{Z}_+,$$

thus the coordinates of $\boldsymbol{Y}_n$ converge in distribution to the same distribution π as $n \to \infty$.

Note that $(X_n)_{n\geqslant -1}$ is only a second-order Markov chain, but not a Markov chain. Moreover, $(X_n)_{n\geqslant -1}$ is strictly stationary if and only if the distribution of the initial population sizes $(X_0, X_{-1})^\top$ coincides with the stationary distribution $\boldsymbol{\pi}$ of the Markov chain $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$. Indeed, if $(X_0, X_{-1})^\top \stackrel{\mathcal{D}}{=} \boldsymbol{\pi}$, then $\boldsymbol{Y}_0 \stackrel{\mathcal{D}}{=} \boldsymbol{\pi}$, thus $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ is strictly stationary, and hence for each $n, m \in \mathbb{Z}_0$, $(\boldsymbol{Y}_0, \ldots, \boldsymbol{Y}_n) \stackrel{\mathcal{D}}{=} (\boldsymbol{Y}_m, \ldots, \boldsymbol{Y}_{n+m})$, yielding

$$(X_0, X_{-1}, X_1, X_0, \ldots, X_n, X_{n-1}) \stackrel{\mathcal{D}}{=} (X_m, X_{m-1}, X_{m+1}, X_m, \ldots, X_{n+m}, X_{n+m-1}).$$

Especially, $(X_{-1}, X_0, X_1, \ldots, X_n) \stackrel{\mathcal{D}}{=} (X_{m-1}, X_m, X_{m+1}, \ldots, X_{n+m})$, hence $(X_n)_{n\geqslant -1}$ is strictly stationary. Since $(X_m, X_{m-1}, X_{m+1}, X_m, \ldots, X_{n+m}, X_{n+m-1})$ is a continuous function of $(X_{m-1}, X_m, X_{m+1}, \ldots, X_{n+m})$, these considerations work backwards as well. Consequently, $\boldsymbol{\pi}$ is the unique stationary distribution of the second-order Markov chain $(X_n)_{n\geqslant -1}$.

2.1 Theorem. *Let $(X_n)_{n\geqslant -1}$ be a second-order Galton–Watson process with immigration such that $m_\xi \in \mathbb{R}_{++}$, $m_\eta \in \mathbb{R}_{++}$, $m_\xi + m_\eta < 1$ and ε is regularly varying with index $\alpha \in (0, 2)$. In case of $\alpha \in [1, 2)$, assume additionally that $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{E}(\eta^2) < \infty$. Then the tail of the marginals π of the unique stationary distribution $\boldsymbol{\pi}$ of $(X_n)_{n\geqslant -1}$ satisfies*

$$\pi((x, \infty)) \sim \sum_{i=0}^{\infty} m_i^\alpha \, \mathbb{P}(\varepsilon > x) \qquad \text{as } x \to \infty,$$

where $m_0 := 1$ *and*

$$(2.1) \qquad m_k := \frac{\lambda_+^{k+1} - \lambda_-^{k+1}}{\lambda_+ - \lambda_-}, \qquad \lambda_+ := \frac{m_\xi + \sqrt{m_\xi^2 + 4m_\eta}}{2}, \qquad \lambda_- := \frac{m_\xi - \sqrt{m_\xi^2 + 4m_\eta}}{2}$$

for $k \in \mathbb{N}$. *Consequently,* π *is also regularly varying with index* α.

Note that λ_+ and λ_- are the eigenvalues of the offspring mean matrix $\boldsymbol{M}_{\xi,\eta}$ given in (B.2) related to the recursive formula (B.1) for the expectations $\mathbb{E}(X_n)$, $n \in \mathbb{N}$. For each $k \in \mathbb{Z}_+$, the assumptions $m_\xi \in \mathbb{R}_{++}$ and $m_\eta \in \mathbb{R}_{++}$ imply $m_k \in \mathbb{R}_{++}$. Further, by (B.4), for all $k \in \mathbb{Z}_+$, we have $m_k = \mathbb{E}(V_{k,0})$, where $(V_{n,0})_{n\geqslant -1}$ is a second-order Galton–Watson process (without immigration) with initial values $V_{0,0} = 1$ and $V_{-1,0} = 0$, and with the same offspring distributions as $(X_n)_{n\geqslant -1}$. Consequently, the series $\sum_{i=0}^\infty m_i^\alpha$ appearing in Theorem 2.1 is convergent, since for each $i \in \mathbb{N}$, we have $m_i = \mathbb{E}(V_{i,0}) \leqslant \lambda_+^i < 1$ by (B.5) and the assumption $m_\xi + m_\eta < 1$.

We point out that in Theorem 2.1 only the regular variation of the marginals π of $\boldsymbol{\pi}$ is proved, the question of the joint regular variation of $\boldsymbol{\pi}$ remains open.

2.2 Remark. Note that there is no overlap between the results in the recent paper of Bősze and Pap [5] on non-stationary second-order Galton-Watson processes with immigration and in the present paper. In [5] the authors always suppose that the initial values X_0 and X_{-1} of a second-order Galton-Watson process with immigration $(X_n)_{n\geqslant -1}$ are independent, so in the results of [5] the distribution of (X_0, X_{-1}) can not be chosen as the unique stationary distribution $\boldsymbol{\pi}$, since the marginals of $\boldsymbol{\pi}$ are not independent in general. □

For the proof of Theorem 2.1, we need an auxiliary result on the tail behaviour of second-order Galton–Watson processes (without immigration) having regularly varying initial distributions.

2.3 Proposition. *Let* $(X_n)_{n\geqslant -1}$ *be a second-order Galton–Watson process (without immigration) such that* X_0 *is regularly varying with index* $\beta_0 \in \mathbb{R}_+$, $X_{-1} = 0$, $m_\xi \in \mathbb{R}_{++}$ *and* $m_\eta \in \mathbb{R}_+$. *In case of* $\beta_0 \in [1, \infty)$, *assume additionally that there exists* $r \in (\beta_0, \infty)$ *with* $\mathbb{E}(\xi^r) < \infty$ *and* $\mathbb{E}(\eta^r) < \infty$. *Then for all* $n \in \mathbb{N}$,

$$\mathbb{P}(X_n > x) \sim m_n^{\beta_0}\, \mathbb{P}(X_0 > x) \qquad \text{as } x \to \infty,$$

where m_i, $i \in \mathbb{Z}_+$, *are given in Theorem 2.1, and hence,* X_n *is also regularly varying with index* β_0 *for each* $n \in \mathbb{N}$.

Proof of Proposition 2.3. Let us fix $n \in \mathbb{N}$. In view of the additive property (A.4), it is sufficient to prove

$$\mathbb{P}\left(\sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} > x\right) \sim m_n^{\beta_0}\, \mathbb{P}(X_0 > x) \qquad \text{as } x \to \infty.$$

This relation follows from Proposition E.13, since $\mathbb{E}(\zeta_{1,0}^{(n)}) = m_n \in \mathbb{R}_{++}$, $n \in \mathbb{N}$, by (B.4). $\Box$

In Appendix G, we present a variant of Proposition 2.3, where the initial values X_{-1} and X_0 are independent and regularly varying together with a second type of proof, see Proposition G.1.

Proof of Theorem 2.1. First, note that, by Lemma E.5, $\mathbb{E}(\mathbb{1}_{\{\varepsilon\neq 0\}} \log(\varepsilon)) < \infty$. We will use the ideas of the proof of Theorem 2.1.1 in Basrak et al. [3]. Due to the representation (A.4), for each $i \in \mathbb{Z}_+$, we have

$$V_i^{(i)}(\varepsilon_i) \stackrel{\mathcal{D}}{=} \sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)},$$

where $\{\varepsilon_i, \zeta_{j,0}^{(i)} : j \in \mathbb{N}\}$ are independent random variables such that $\{\zeta_{j,0}^{(i)} : j \in \mathbb{N}\}$ are independent copies of $V_{i,0}$, where $(V_{k,0})_{k\geqslant -1}$ is a second-order Galton–Watson process (without immigration) with initial values $V_{0,0} = 1$ and $V_{-1,0} = 0$, and with the same offspring distributions as $(X_k)_{k\geqslant -1}$. For each $i \in \mathbb{Z}_+$, by Proposition 2.3, we obtain $\mathbb{P}(V_i^{(i)}(\varepsilon_i) > x) \sim m_i^\alpha \, \mathbb{P}(\varepsilon > x)$ as $x \to \infty$, yielding that random variables $V_i^{(i)}(\varepsilon_i)$, $i \in \mathbb{Z}_+$, are also regularly varying with index α. Since $V_i^{(i)}(\varepsilon_i)$, $i \in \mathbb{Z}_+$, are independent, for each $n \in \mathbb{Z}_+$, by Lemma E.10, we have

$$\mathbb{P}\left(\sum_{i=0}^{n} V_i^{(i)}(\varepsilon_i) > x\right) \sim \sum_{i=0}^{n} m_i^\alpha \, \mathbb{P}(\varepsilon > x) \qquad \text{as } x \to \infty, \tag{2.2}$$

and hence the random variables $\sum_{i=0}^{n} V_i^{(i)}(\varepsilon_i)$, $n \in \mathbb{Z}_+$, are also regularly varying with index α. For each $n \in \mathbb{N}$, using that $V_i^{(i)}(\varepsilon_i)$, $i \in \mathbb{Z}_+$, are non-negative, we have

$$\liminf_{x\to\infty} \frac{\pi((x,\infty))}{\mathbb{P}(\varepsilon > x)} = \liminf_{x\to\infty} \frac{\mathbb{P}(\sum_{i=0}^{\infty} V_i^{(i)}(\varepsilon_i) > x)}{\mathbb{P}(\varepsilon > x)} \geqslant \liminf_{x\to\infty} \frac{\mathbb{P}(\sum_{i=0}^{n} V_i^{(i)}(\varepsilon_i) > x)}{\mathbb{P}(\varepsilon > x)} = \sum_{i=0}^{n} m_i^\alpha,$$

hence, letting $n \to \infty$, we obtain

$$\liminf_{x\to\infty} \frac{\pi((x,\infty))}{\mathbb{P}(\varepsilon > x)} \geqslant \sum_{i=0}^{\infty} m_i^\alpha. \tag{2.3}$$

Moreover, for each $n \in \mathbb{N}$ and $q \in (0,1)$, we have

$$\begin{aligned}
\limsup_{x\to\infty} \frac{\pi((x,\infty))}{\mathbb{P}(\varepsilon > x)} &= \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=0}^{n-1} V_i^{(i)}(\varepsilon_i) + \sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > x\big)}{\mathbb{P}(\varepsilon > x)} \\
&\leqslant \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=0}^{n-1} V_i^{(i)}(\varepsilon_i) > (1-q)x\big) + \mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > x)} \leqslant L_{1,n}(q) + L_{2,n}(q)
\end{aligned}$$

with

$$L_{1,n}(q) := \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=0}^{n-1} V_i^{(i)}(\varepsilon_i) > (1-q)x\big)}{\mathbb{P}(\varepsilon > x)}, \qquad L_{2,n}(q) := \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > x)}.$$

Since ε is regularly varying with index α, by (2.2), we obtain

$$L_{1,n}(q) = \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=0}^{n-1} V_i^{(i)}(\varepsilon_i) > (1-q)x\big)}{\mathbb{P}(\varepsilon > (1-q)x)} \cdot \frac{\mathbb{P}(\varepsilon > (1-q)x)}{\mathbb{P}(\varepsilon > x)} = (1-q)^{-\alpha} \sum_{i=0}^{n-1} m_i^\alpha$$

and

$$L_{2,n}(q) = \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > qx)} \cdot \frac{\mathbb{P}(\varepsilon > qx)}{\mathbb{P}(\varepsilon > x)} = q^{-\alpha} \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > qx)},$$

and hence

$$\lim_{n\to\infty} L_{1,n}(q) = (1-q)^{-\alpha} \sum_{i=0}^{\infty} m_i^\alpha,$$

$$\lim_{n\to\infty} L_{2,n}(q) = q^{-\alpha} \lim_{n\to\infty} \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > qx)}.$$

The aim of the following discussion is to show

$$\lim_{n\to\infty} \limsup_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > qx\big)}{\mathbb{P}(\varepsilon > qx)} = 0, \qquad q \in (0,1). \tag{2.4}$$

First, we consider the case $\alpha \in (0,1)$. For each $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$, we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=n}^{\infty} V_i^{(i)}(\varepsilon_i) > x\right) \\
&= \mathbb{P}\left(\sum_{i\geqslant n} V_i^{(i)}(\varepsilon_i) > x,\ \sup_{i\geqslant n} \varrho^i \varepsilon_i > (1-\delta)x\right) + \mathbb{P}\left(\sum_{i\geqslant n} V_i^{(i)}(\varepsilon_i) > x,\ \sup_{i\geqslant n} \varrho^i \varepsilon_i \leqslant (1-\delta)x\right) \\
&= \mathbb{P}\left(\sum_{i\geqslant n} V_i^{(i)}(\varepsilon_i) > x,\ \sup_{i\geqslant n} \varrho^i \varepsilon_i > (1-\delta)x\right) \\
&\quad + \mathbb{P}\left(\sum_{i\geqslant n} V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}} > x,\ \sup_{i\geqslant n} \varrho^i \varepsilon_i \leqslant (1-\delta)x\right) \\
&\leqslant \mathbb{P}\left(\sup_{i\geqslant n} \varrho^i \varepsilon_i > (1-\delta)x\right) + \mathbb{P}\left(\sum_{i\geqslant n} V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}} > x\right) =: P_{1,n}(x,\delta) + P_{2,n}(x,\delta),
\end{aligned}$$

where ϱ is given in (B.6). By subadditivity of probability,

$$P_{1,n}(x,\delta) \leqslant \sum_{i\geqslant n} \mathbb{P}(\varrho^i \varepsilon_i > (1-\delta)x) = \sum_{i\geqslant n} \mathbb{P}(\varepsilon > (1-\delta)\varrho^{-i}x).$$

Using Potter's upper bound (see Lemma E.12), for $\delta \in (0, \frac{\alpha}{2})$, there exists $x_0 \in \mathbb{R}_{++}$ such that

$$\frac{\mathbb{P}(\varepsilon > (1-\delta)\varrho^{-i}x)}{\mathbb{P}(\varepsilon > x)} < (1+\delta)[(1-\delta)\varrho^{-i}]^{-\alpha+\delta} < (1+\delta)[(1-\delta)\varrho^{-i}]^{-\frac{\alpha}{2}} \tag{2.5}$$

if $x \in [x_0, \infty)$ and $(1-\delta)\varrho^{-i} \in [1, \infty)$, which holds for sufficiently large $i \in \mathbb{N}$ due to $\varrho \in (0,1)$. Consequently, if $\delta \in (0, \frac{\alpha}{2})$, then

$$\lim_{n\to\infty} \limsup_{x\to\infty} \frac{P_{1,n}(x,\delta)}{\mathbb{P}(\varepsilon > x)} \leqslant \lim_{n\to\infty} \sum_{i\geqslant n} (1+\delta)[(1-\delta)\varrho^{-i}]^{-\frac{\alpha}{2}} = 0,$$

since $\varrho^{\frac{\alpha}{2}} < 1$ (due to $\varrho \in (0,1)$) yields $\sum_{i=0}^{\infty} (\varrho^{-i})^{-\alpha/2} < \infty$. Now we turn to prove that $\lim_{n\to\infty} \limsup_{x\to\infty} \frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon_1 > x)} = 0$. By Markov's inequality,

$$P_{2,n}(x,\delta) \leqslant \frac{1}{x} \sum_{i\geqslant n} \mathbb{E}\big(V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}}\big).$$

By the representation $V_i^{(i)}(\varepsilon_i) \stackrel{\mathcal{D}}{=} \sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)}$, we have

$$\begin{aligned}
\mathbb{E}\big(V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}}\big) &= \mathbb{E}\left(\sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)} \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}} \right) = \mathbb{E}\left[\mathbb{E}\left(\sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)} \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}} \,\middle|\, \varepsilon_i \right) \right] \\
&= \mathbb{E}\left(\sum_{j=1}^{\varepsilon_i} \mathbb{E}(\zeta_{1,0}^{(i)}) \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}} \right) = \mathbb{E}(\zeta_{1,0}^{(i)})\, \mathbb{E}\big(\varepsilon_i \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}}\big),
\end{aligned}$$

since $\{\zeta_{j,0}^{(i)} : j \in \mathbb{N}\}$ and ε_i are independent. Moreover,

$$\begin{aligned}
\mathbb{E}\big(\varepsilon_i \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)\varrho^{-i}x\}}\big) &= \mathbb{E}\big(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant (1-\delta)\varrho^{-i}x\}}\big) = \int_0^\infty \mathbb{P}\big(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant (1-\delta)\varrho^{-i}x\}} > t\big)\, \mathrm{d}t \\
&= \int_0^{(1-\delta)\varrho^{-i}x} \mathbb{P}(t < \varepsilon \leqslant (1-\delta)\varrho^{-i}x)\, \mathrm{d}t \leqslant \int_0^{(1-\delta)\varrho^{-i}x} \mathbb{P}(\varepsilon > t)\, \mathrm{d}t.
\end{aligned}$$

By Karamata's theorem (see, Theorem E.11), we have

$$\lim_{y\to\infty} \frac{\int_0^y \mathbb{P}(\varepsilon > t)\, \mathrm{d}t}{y\, \mathbb{P}(\varepsilon > y)} = \frac{1}{1-\alpha},$$

thus there exists $y_0 \in \mathbb{R}_{++}$ such that

$$\int_0^y \mathbb{P}(\varepsilon > t)\, \mathrm{d}t \leqslant \frac{2y\, \mathbb{P}(\varepsilon > y)}{1-\alpha}, \qquad y \in [y_0, \infty),$$

hence

$$\int_0^{(1-\delta)\varrho^{-i}x} \mathbb{P}(\varepsilon > t)\, \mathrm{d}t \leqslant \frac{2(1-\delta)\varrho^{-i}x\, \mathbb{P}(\varepsilon > (1-\delta)\varrho^{-i}x)}{1-\alpha}$$

whenever $(1-\delta)\varrho^{-i}x \in [y_0, \infty)$, which holds for $i \geqslant n$ with sufficiently large $n \in \mathbb{N}$ and $x \in [(1-\delta)^{-1}\varrho^n y_0, \infty)$ due to $\varrho \in (0,1)$. Thus, for sufficiently large $n \in \mathbb{N}$ and $x \in [(1-\delta)^{-1}\varrho^n y_0, \infty)$, we obtain

$$\begin{aligned}
\frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon > x)} &\leqslant \frac{1}{x\, \mathbb{P}(\varepsilon > x)} \sum_{i\geqslant n} \mathbb{E}(\zeta_{1,0}^{(i)}) \int_0^{(1-\delta)\varrho^{-i}x} \mathbb{P}(\varepsilon > t)\, \mathrm{d}t \\
&\leqslant \frac{2(1-\delta)}{1-\alpha} \sum_{i\geqslant n} \frac{\mathbb{P}(\varepsilon > (1-\delta)\varrho^{-i}x)}{\mathbb{P}(\varepsilon > x)},
\end{aligned}$$

since $\mathbb{E}(\zeta_{1,0}^{(i)}) \leqslant \varrho^i$, $i \in \mathbb{Z}_+$, by (B.5) and $\zeta_{1,0}^{(0)} = 1$. Using (2.5), we get

$$\frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon > x)} \leqslant \frac{2(1-\delta)}{1-\alpha} \sum_{i \geqslant n} (1+\delta)[(1-\delta)\varrho^{-i}]^{-\frac{\alpha}{2}}$$

for $\delta \in (0, \frac{\alpha}{2})$, for sufficiently large $n \in \mathbb{N}$ and for all $x \in [\max(x_0, (1-\delta)^{-1}\varrho^n y_0), \infty)$. Hence for $\delta \in (0, \frac{\alpha}{2})$ we have

$$\lim_{n\to\infty} \limsup_{x\to\infty} \frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon > x)} \leqslant \lim_{n\to\infty} \frac{2(1-\delta^2)}{1-\alpha} \sum_{i \geqslant n} [(1-\delta)\varrho^{-i}]^{-\frac{\alpha}{2}} = 0,$$

where the last step follows by the fact that the series $\sum_{i=0}^\infty (\varrho^i)^{\frac{\alpha}{2}}$ is convergent, since $\varrho \in (0,1)$.

Consequently, due to the fact that $\mathbb{P}(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i) > x) \leqslant P_{1,n}(x,\delta) + P_{2,n}(x,\delta)$, $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$, $\delta \in (0,1)$, we obtain (2.4), and we conclude $\lim_{n\to\infty} L_{2,n}(q) = 0$ for all $q \in (0,1)$. Thus we obtain

$$\limsup_{x\to\infty} \frac{\pi((x,\infty))}{\mathbb{P}(\varepsilon > x)} \leqslant \lim_{n\to\infty} L_{1,n}(q) + \lim_{n\to\infty} L_{2,n}(q) = (1-q)^{-\alpha} \sum_{i=0}^\infty m_i^\alpha$$

for all $q \in (0,1)$. Letting $q \downarrow 0$, this yields

$$\limsup_{x\to\infty} \frac{\pi((x,\infty))}{\mathbb{P}(\varepsilon > x)} \leqslant \sum_{i=0}^\infty m_i^\alpha.$$

Taking into account (2.3), the proof of (2.4) is complete in case of $\alpha \in (0,1)$.

Next, we consider the case $\alpha \in [1,2)$. Note that (2.4) is equivalent to

$$\lim_{n\to\infty} \limsup_{x\to\infty} \frac{\mathbb{P}(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i) > \sqrt{x})}{\mathbb{P}(\varepsilon > \sqrt{x})} = \lim_{n\to\infty} \limsup_{x\to\infty} \frac{\mathbb{P}\big((\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i))^2 > x\big)}{\mathbb{P}(\varepsilon^2 > x)} = 0.$$

Repeating a similar argument as for $\alpha \in (0,1)$, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\left(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\right)^2 > x\right) \\
&= \mathbb{P}\left(\left(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\right)^2 > x,\ \sup_{i\geqslant n} \varrho^{2i}\varepsilon_i^2 > (1-\delta)x\right) \\
&\quad + \mathbb{P}\left(\left(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\right)^2 > x,\ \sup_{i\geqslant n} \varrho^{2i}\varepsilon_i^2 \leqslant (1-\delta)x\right) \\
&= \mathbb{P}\left(\left(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\right)^2 > x,\ \sup_{i\geqslant n} \varrho^{2i}\varepsilon_i^2 > (1-\delta)x\right) \\
&\quad + \mathbb{P}\left(\left(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i}x\}}\right)^2 > x,\ \sup_{i\geqslant n} \varrho^{2i}\varepsilon_i^2 \leqslant (1-\delta)x\right)
\end{aligned}$$

$$\leqslant \mathbb{P}\biggl(\sup_{i\geqslant n} \varrho^{2i}\varepsilon_i^2 > (1-\delta)x\biggr) + \mathbb{P}\Biggl(\Biggl(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\mathbb{1}_{\{\varepsilon_i^2\leqslant(1-\delta)\varrho^{-2i}x\}}\Biggr)^2 > x\Biggr)$$

$$=: P_{1,n}(x,\delta) + P_{2,n}(x,\delta)$$

for each $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$. By the subadditivity of probability,

$$P_{1,n}(x,\delta) \leqslant \sum_{i=n}^\infty \mathbb{P}(\varrho^{2i}\varepsilon_i^2 > (1-\delta)x) = \sum_{i=n}^\infty \mathbb{P}(\varepsilon^2 > (1-\delta)\varrho^{-2i}x)$$

for each $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$. Since ε^2 is regularly varying with index $\frac{\alpha}{2}$ (see Lemma E.3), using Potter's upper bound (see Lemma E.12) for $\delta \in \bigl(0, \frac{\alpha}{4}\bigr)$, there exists $x_0 \in \mathbb{R}_{++}$ such that

$$\frac{\mathbb{P}(\varepsilon^2 > (1-\delta)\varrho^{-2i}x)}{\mathbb{P}(\varepsilon^2 > x)} < (1+\delta)[(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{2}+\delta} < (1+\delta)[(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{4}} \tag{2.6}$$

if $x \in [x_0,\infty)$ and $(1-\delta)\varrho^{-2i} \in [1,\infty)$, which holds for sufficiently large $i \in \mathbb{N}$ (due to $\varrho \in (0,1)$). Consequently, if $\delta \in (0, \frac{\alpha}{4})$, then

$$\lim_{n\to\infty}\limsup_{x\to\infty} \frac{P_{1,n}(x,\delta)}{\mathbb{P}(\varepsilon^2 > x)} \leqslant \lim_{n\to\infty}\sum_{i=n}^\infty (1+\delta)[(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{4}} = 0,$$

since $\varrho^{\frac{\alpha}{2}} < 1$ (due to $\varrho \in (0,1)$). By Markov's inequality, for $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$, we have

$$\begin{aligned}
\frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon^2 > x)} &\leqslant \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)}\mathbb{E}\Biggl(\Biggl(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)\mathbb{1}_{\{\varepsilon_i^2\leqslant(1-\delta)\varrho^{-2i}x\}}\Biggr)^2\Biggr)\\
&= \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)}\mathbb{E}\Biggl(\sum_{i=n}^\infty V_i^{(i)}(\varepsilon_i)^2\mathbb{1}_{\{\varepsilon_i^2\leqslant(1-\delta)\varrho^{-2i}x\}}\Biggr)\\
&\quad + \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)}\mathbb{E}\Biggl(\sum_{i,j=n,\,i\neq j}^\infty V_i^{(i)}(\varepsilon_i)V_j^{(j)}(\varepsilon_j)\mathbb{1}_{\{\varepsilon_i^2\leqslant(1-\delta)\varrho^{-2i}x\}}\mathbb{1}_{\{\varepsilon_j^2\leqslant(1-\delta)\varrho^{-2j}x\}}\Biggr)\\
&=: J_{2,1,n}(x,\delta) + J_{2,2,n}(x,\delta)
\end{aligned}$$

for each $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$. By Lemma C.2, (B.4) and (B.5) with $X_0 = 1$ and $X_{-1} = 0$, we have

$$\begin{aligned}
\mathbb{E}(V_i^{(i)}(n)^2) &= \mathbb{E}\Biggl(\Biggl(\sum_{j=1}^n \zeta_{j,0}^{(i)}\Biggr)^2\Biggr) = \sum_{j=1}^n \mathbb{E}\bigl((\zeta_{j,0}^{(i)})^2\bigr) + \sum_{j,\ell=1,\,j\neq\ell}^n \mathbb{E}\bigl(\zeta_{j,0}^{(i)}\bigr)\,\mathbb{E}\bigl(\zeta_{\ell,0}^{(i)}\bigr)\\
&\leqslant c_{\mathrm{sub}}\sum_{j=1}^n \varrho^i + \sum_{j,\ell=1,\,j\neq\ell}^n \varrho^i\varrho^i \leqslant c_{\mathrm{sub}} n\varrho^i + (n^2-n)\varrho^{2i} \leqslant c_{\mathrm{sub}}\varrho^i n + \varrho^{2i}n^2
\end{aligned}$$

for $i, n \in \mathbb{N}$. Hence, using that $(\varepsilon_i, V_i^{(i)}(\varepsilon_i)) \stackrel{\mathcal{D}}{=} \big(\varepsilon_i, \sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)}\big)$ and that ε_i and $\{\zeta_{j,0}^{(i)} : j \in \mathbb{N}\}$ are independent, we have

$$\begin{aligned}
J_{2,1,n}(x,\delta) &= \sum_{i=n}^{\infty} \frac{\mathbb{E}\big(V_i^{(i)}(\varepsilon_i)^2 \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i}x\}}\big)}{x\,\mathbb{P}(\varepsilon^2 > x)} \\
&= \sum_{i=n}^{\infty} \frac{\mathbb{E}\Big(\big(\sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)}\big)^2 \mathbb{1}_{\{\varepsilon_i \leqslant (1-\delta)^{\frac{1}{2}}\varrho^{-i}x^{\frac{1}{2}}\}}\Big)}{x\,\mathbb{P}(\varepsilon^2 > x)} \\
&= \sum_{i=n}^{\infty} \frac{\sum_{0 \leqslant \ell \leqslant (1-\delta)^{\frac{1}{2}}\varrho^{-i}x^{\frac{1}{2}}} \mathbb{E}\Big(\big(\sum_{j=1}^{\ell} \zeta_{j,0}^{(i)}\big)^2\Big)\,\mathbb{P}(\varepsilon_i = \ell)}{x\,\mathbb{P}(\varepsilon^2 > x)} \\
&\leqslant \sum_{i=n}^{\infty} \frac{\sum_{0 \leqslant \ell \leqslant (1-\delta)^{\frac{1}{2}}\varrho^{-i}x^{\frac{1}{2}}} (c_{\mathrm{sub}}\varrho^i \ell + \varrho^{2i}\ell^2)\,\mathbb{P}(\varepsilon = \ell)}{x\,\mathbb{P}(\varepsilon^2 > x)} \\
&= \sum_{i=n}^{\infty} c_{\mathrm{sub}}\varrho^i \frac{\mathbb{E}\big(\varepsilon \mathbb{1}_{\{\varepsilon^2 \leqslant (1-\delta)\varrho^{-2i}x\}}\big)}{x\,\mathbb{P}(\varepsilon^2 > x)} + \sum_{i=n}^{\infty} \varrho^{2i} \frac{\mathbb{E}\big(\varepsilon^2 \mathbb{1}_{\{\varepsilon^2 \leqslant (1-\delta)\varrho^{-2i}x\}}\big)}{x\,\mathbb{P}(\varepsilon^2 > x)} \\
&=: J_{2,1,1,n}(x,\delta) + J_{2,1,2,n}(x,\delta).
\end{aligned}$$

Since ε^2 is regularly varying with index $\frac{\alpha}{2} \in [\frac{1}{2}, 1)$ (see Lemma E.3), by Karamata's theorem (see, Theorem E.11), we have

$$\lim_{y \to \infty} \frac{\int_0^y \mathbb{P}(\varepsilon^2 > t)\,\mathrm{d}t}{y\,\mathbb{P}(\varepsilon^2 > y)} = \frac{1}{1 - \frac{\alpha}{2}},$$

thus there exists $y_0 \in \mathbb{R}_{++}$ such that

$$\int_0^y \mathbb{P}(\varepsilon^2 > t)\,\mathrm{d}t \leqslant \frac{2y\,\mathbb{P}(\varepsilon^2 > y)}{1 - \frac{\alpha}{2}}, \qquad y \in [y_0, \infty),$$

hence

$$\begin{aligned}
\mathbb{E}(\varepsilon^2 \mathbb{1}_{\{\varepsilon^2 \leqslant (1-\delta)\varrho^{-2i}x\}}) &= \int_0^\infty \mathbb{P}(\varepsilon^2 \mathbb{1}_{\{\varepsilon^2 \leqslant (1-\delta)\varrho^{-2i}x\}} > y)\,\mathrm{d}y \\
&= \int_0^{(1-\delta)\varrho^{-2i}x} \mathbb{P}(y < \varepsilon^2 \leqslant (1-\delta)\varrho^{-2i}x)\,\mathrm{d}y \\
&\leqslant \int_0^{(1-\delta)\varrho^{-2i}x} \mathbb{P}(\varepsilon^2 > t)\,\mathrm{d}t \leqslant \frac{2(1-\delta)\varrho^{-2i}x\,\mathbb{P}(\varepsilon^2 > (1-\delta)\varrho^{-2i}x)}{1 - \frac{\alpha}{2}}
\end{aligned}$$

whenever $(1-\delta)\varrho^{-2i}x \in [y_0, \infty)$, which holds for $i \geqslant n$ with sufficiently large $n \in \mathbb{N}$, and $x \in [(1-\delta)^{-1}\varrho^{2n}y_0, \infty)$ due to $\varrho \in (0,1)$. Thus for $\delta \in (0, \frac{\alpha}{4})$, for sufficiently large $n \in \mathbb{N}$ (satisfying $(1-\delta)\varrho^{-2n} \in (1, \infty)$ as well) and for all $x \in [\max(x_0, (1-\delta)^{-1}\varrho^{2n}y_0), \infty)$, using

(2.6), we obtain

$$J_{2,1,2,n}(x,\delta) \leqslant \frac{2(1-\delta)}{1-\frac{\alpha}{2}} \sum_{i=n}^{\infty} \frac{\mathbb{P}(\varepsilon^2 > (1-\delta)\varrho^{-2i}x)}{\mathbb{P}(\varepsilon^2 > x)} \leqslant \frac{2(1-\delta)}{1-\frac{\alpha}{2}} \sum_{i=n}^{\infty} (1+\delta)[(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{4}}$$
$$= \frac{2(1-\delta^2)}{1-\frac{\alpha}{2}} \sum_{i=n}^{\infty} [(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{4}}.$$

Hence for $\delta \in (0, \frac{\alpha}{4})$, we have

$$\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,2,n}(x,\delta) \leqslant \frac{2(1-\delta^2)}{1-\frac{\alpha}{2}} \lim_{n\to\infty} \sum_{i=n}^{\infty} [(1-\delta)\varrho^{-2i}]^{-\frac{\alpha}{4}} = 0,$$

yielding $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,2,n}(x,\delta) = 0$ for $\delta \in (0, \frac{\alpha}{4})$. Further, if $\alpha \in (1,2)$, or $\alpha = 1$ and $m_\varepsilon < \infty$, we have

$$J_{2,1,1,n}(x,\delta) \leqslant c_{\mathrm{sub}} \sum_{i=n}^{\infty} \varrho^i \frac{m_\varepsilon}{x\,\mathbb{P}(\varepsilon^2 > x)},$$

and hence, using that $\lim_{x\to\infty} x\,\mathbb{P}(\varepsilon^2 > x) = \infty$ (see Lemma E.4),

$$\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,1,n}(x,\delta) \leqslant c_{\mathrm{sub}} m_\varepsilon \lim_{n\to\infty} \left(\sum_{i=n}^{\infty} \varrho^i \right) \limsup_{x\to\infty} \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)} = 0,$$

yielding $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,1,n}(x,\delta) = 0$ for $\delta \in (0,1)$.

If $\alpha = 1$ and $m_\varepsilon = \infty$, then we have

$$J_{2,1,1,n}(x,\delta) = \sum_{i=n}^{\infty} c_{\mathrm{sub}} \varrho^i \frac{\mathbb{E}\left(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant (1-\delta)^{\frac{1}{2}} \varrho^{-i} x^{\frac{1}{2}}\}}\right)}{x\,\mathbb{P}(\varepsilon^2 > x)}$$

for $x \in \mathbb{R}_{++}$, $n \in \mathbb{N}$ and $\delta \in (0,1)$. Note that

$$\mathbb{E}(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant y\}}) \leqslant \int_0^\infty \mathbb{P}(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant y\}} > t)\,\mathrm{d}t = \int_0^y \mathbb{P}(t < \varepsilon \leqslant y)\,\mathrm{d}t \leqslant \int_0^y \mathbb{P}(t < \varepsilon)\,\mathrm{d}t =: \widetilde{L}(y)$$

for $y \in \mathbb{R}_+$. Because of $\alpha = 1$, Proposition 1.5.9a in Bingham et al. [4] yields that $\widetilde{L}$ is a slowly varying function (at infinity). By Potter's bounds (see Lemma E.12), for every $\delta \in \mathbb{R}_{++}$, there exists $z_0 \in \mathbb{R}_{++}$ such that

$$\frac{\widetilde{L}(y)}{\widetilde{L}(z)} < (1+\delta) \left(\frac{y}{z}\right)^\delta$$

for $z \geqslant z_0$ and $y \geqslant z$. Hence, for $x \geqslant z_0^2$, we have

$$\mathbb{E}\big(\varepsilon \mathbb{1}_{\{\varepsilon \leqslant (1-\delta)^{\frac{1}{2}} \varrho^{-i} x^{\frac{1}{2}}\}}\big) \leqslant \widetilde{L}\big((1-\delta)^{\frac{1}{2}} \varrho^{-i} x^{\frac{1}{2}}\big) \leqslant \widetilde{L}(\varrho^{-i} x^{\frac{1}{2}}) \leqslant (1+\delta)\varrho^{-i\delta} \widetilde{L}(x^{\frac{1}{2}}), \qquad i \geqslant n,$$

where we also used that $\widetilde{L}$ is monotone increasing. Using this, we conclude that for every $\delta \in \mathbb{R}_{++}$, there exists $z_0 \in \mathbb{R}_{++}$ such that for $x \geqslant z_0^2$, we have

$$J_{2,1,1,n}(x,\delta) \leqslant (1+\delta) c_{\mathrm{sub}} \frac{\widetilde{L}(x^{\frac{1}{2}})}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i=n}^{\infty} \varrho^{-i\delta}.$$

Here, since $\varrho \in (0,1)$ and $\delta \in \mathbb{R}_{++}$, we have $\lim_{n\to\infty} \sum_{i=n}^{\infty} \varrho^{-i\delta} = 0$, and

$$\frac{\widetilde{L}(\sqrt{x})}{x\,\mathbb{P}(\varepsilon^2 > x)} = \frac{\widetilde{L}(\sqrt{x})}{x^{1/4}} \cdot \frac{1}{x^{3/4}\,\mathbb{P}(\varepsilon > \sqrt{x})} \to 0 \qquad \text{as } x \to \infty,$$

by Lemma E.4, due to the fact that $\widetilde{L}$ is slowly varying and the function $\mathbb{R}_{++} \ni x \mapsto \mathbb{P}(\varepsilon > \sqrt{x})$ is regularly varying with index $-1/2$. Hence $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,1,n}(x,\delta) = 0$ for $\delta \in (0,1)$ in case of $\alpha = 1$ and $m_\varepsilon = \infty$.

Consequently, we have $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,1,n}(x,\delta) = 0$ for $\delta \in (0, \frac{\alpha}{4})$.

Now we turn to prove $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,2,n}(x,\delta) = 0$ for $\delta \in (0,1)$. Using that $\{(\varepsilon_i, V_i^{(i)}(\varepsilon_i) : i \in \mathbb{N}\}$ are independent, we have

$$J_{2,2,n}(x,\delta) \leqslant \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i,j=n,\, i\neq j}^{\infty} \mathbb{E}(V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i} x\}})\, \mathbb{E}(V_j^{(j)}(\varepsilon_j) \mathbb{1}_{\{\varepsilon_j^2 \leqslant (1-\delta)\varrho^{-2j} x\}}).$$

Here, using that $\Big(\varepsilon_i, V_i^{(i)}(\varepsilon_i)\Big) \stackrel{\mathcal{D}}{=} \big(\varepsilon_i, \sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)}\big)$, where ε_i and $\{\zeta_{j,0}^{(i)} : j \in \mathbb{N}\}$ are independent, and (B.5) with $X_0 = 1$ and $X_{-1} = 0$, we have

$$\begin{aligned}
\mathbb{E}(V_i^{(i)}(\varepsilon_i) \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i} x\}}) &= \mathbb{E}\left(\left(\sum_{j=1}^{\varepsilon_i} \zeta_{j,0}^{(i)}\right) \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i} x\}}\right) \\
&= \sum_{\ell=0}^{\lfloor (1-\delta)^{\frac{1}{2}} \varrho^{-i} x^{\frac{1}{2}} \rfloor} \mathbb{E}\left(\sum_{j=1}^{\ell} \zeta_{j,0}^{(i)}\right) \mathbb{P}(\varepsilon_i = \ell) \\
&\leq \sum_{\ell=0}^{\lfloor (1-\delta)^{\frac{1}{2}} \varrho^{-i} x^{\frac{1}{2}} \rfloor} \ell \varrho^i\, \mathbb{P}(\varepsilon_i = \ell) = \varrho^i\, \mathbb{E}(\varepsilon_i \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i} x\}})
\end{aligned}$$

for $x \in \mathbb{R}_{++}$ and $\delta \in (0,1)$. If $\alpha \in (1,2)$, or $\alpha = 1$ and $m_\varepsilon < \infty$, then

$$\begin{aligned}
J_{2,2,n}(x,\delta) &\leqslant \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i,j=n,\, i\neq j}^{\infty} \varrho^{i+j}\, \mathbb{E}(\varepsilon_i \mathbb{1}_{\{\varepsilon_i^2 \leqslant (1-\delta)\varrho^{-2i} x\}})\, \mathbb{E}(\varepsilon_j \mathbb{1}_{\{\varepsilon_j^2 \leqslant (1-\delta)\varrho^{-2j} x\}}) \\
&\leqslant \frac{m_\varepsilon^2}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i,j=n,\, i\neq j}^{\infty} \varrho^{i+j} \leqslant \frac{m_\varepsilon^2}{x\,\mathbb{P}(\varepsilon^2 > x)} \left(\sum_{i=n}^{\infty} \varrho^i\right)^2
\end{aligned}$$

for $x \in \mathbb{R}_{++}$ and $\delta \in (0,1)$, and then, by Lemma E.4,

$$\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,2,n}(x,\delta) \leqslant m_\varepsilon^2 \lim_{n\to\infty} \left(\sum_{i=n}^\infty \varrho^i\right)^2 \limsup_{x\to\infty} \frac{1}{x\,\mathbb{P}(\varepsilon^2 > x)}$$
$$= m_\varepsilon^2 \left(\lim_{n\to\infty} \frac{\varrho^{2n}}{(1-\varrho)^2}\right) \cdot 0 = 0,$$

yielding that $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,2,n}(x,\delta) = 0$.

If $\alpha = 1$ and $m_\varepsilon = \infty$, then we can apply the same argument as for $J_{2,1,1,n}(x,\delta)$. Namely,

$$J_{2,2,n}(x,\delta) \leqslant \frac{(1+\delta)^2}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i,j=n,\, i\neq j}^\infty \varrho^{(1-\delta)(i+j)} (\widetilde{L}(x^{\frac{1}{2}}))^2$$
$$\leqslant (1+\delta)^2 \frac{(\widetilde{L}(x^{\frac{1}{2}}))^2}{x\,\mathbb{P}(\varepsilon^2 > x)} \sum_{i,j=n,\, i\neq j}^\infty \varrho^{(1-\delta)(i+j)} = (1+\delta)^2 \frac{(\widetilde{L}(x^{\frac{1}{2}}))^2}{x\,\mathbb{P}(\varepsilon^2 > x)} \left(\sum_{i=n}^\infty \varrho^{(1-\delta)i}\right)^2$$

for $x \in \mathbb{R}_{++}$ and $\delta \in (0,1)$, where

$$\frac{(\widetilde{L}(x^{\frac{1}{2}}))^2}{x\,\mathbb{P}(\varepsilon^2 > x)} = \left(\frac{\widetilde{L}(x^{\frac{1}{2}})}{x^{\frac{1}{2}}}\right)^2 \frac{1}{x^{\frac{3}{4}}\,\mathbb{P}(\varepsilon > \sqrt{x})} \to 0 \qquad \text{as } x \to \infty,$$

yielding that $\lim_{n\to\infty} \limsup_{x\to\infty} J_{2,2,n}(x,\delta) = 0$ for $\delta \in (0,1)$ in case of $\alpha = 1$ and $m_\varepsilon = \infty$ as well.

Consequently, $\lim_{n\to\infty} \limsup_{x\to\infty} \frac{P_{2,n}(x,\delta)}{\mathbb{P}(\varepsilon^2 > x)} = 0$ for $\delta \in (0, \frac{\alpha}{4})$ yielding (2.4) in case of $\alpha \in [1,2)$ as well, and we conclude $\lim_{n\to\infty} L_{2,n}(q) = 0$ for all $q \in (0,1)$. The proof can be finished as in case of $\alpha \in (0,1)$. □

2.4 Remark. The statement of Theorem 2.1 remains true in the case when $m_\xi \in (0,1)$ and $m_\eta = 0$. In this case we get the statement for classical Galton–Watson processes, see Theorem 2.1.1 in Basrak et al. [3] or Theorem 1.1. However, note that this is not a special case of Theorem 2.1, since in this case the mean matrix $\boldsymbol{M}_{\xi,\eta}$ is not primitive. □

Appendices

A Representations of second-order Galton–Watson processes without or with immigration

First, we recall a representation of a second-order Galton–Watson process without or with immigration as a (special) 2-type Galton–Watson process without or with immigration, respectively. Let $(X_n)_{n\geqslant -1}$ be a second-order Galton–Watson process with immigration given in

(1.2), and let us introduce the random vectors

$$\text{(A.1)} \qquad \boldsymbol{Y}_n := \begin{bmatrix} Y_{n,1} \\ Y_{n,2} \end{bmatrix} := \begin{bmatrix} X_n \\ X_{n-1} \end{bmatrix}, \qquad n \in \mathbb{Z}_+.$$

Then we have

$$\text{(A.2)} \qquad \boldsymbol{Y}_n = \sum_{i=1}^{Y_{n-1,1}} \begin{bmatrix} \xi_{n,i} \\ 1 \end{bmatrix} + \sum_{j=1}^{Y_{n-1,2}} \begin{bmatrix} \eta_{n,j} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_n \\ 0 \end{bmatrix}, \qquad n \in \mathbb{N},$$

hence $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}_+}$ is a (special) 2-type Galton–Watson process with immigration and with initial vector

$$\boldsymbol{Y}_0 = \begin{bmatrix} X_0 \\ X_{-1} \end{bmatrix}.$$

In fact, the type 1 and 2 individuals are identified with individuals of age 0 and 1, respectively, and for each $n, i, j \in \mathbb{N}$, at time n, the i^{th} individual of type 1 of the $(n-1)^{\mathrm{th}}$ generation produces $\xi_{n,i}$ individuals of type 1 and exactly one individual of type 2, and the j^{th} individual of type 2 of the $(n-1)^{\mathrm{th}}$ generation produces $\eta_{n,j}$ individuals of type 1 and no individual of type 2.

The representation (A.2) works backwards as well, namely, let $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ be a special 2-type Galton–Watson process with immigration given by

$$\text{(A.3)} \qquad \boldsymbol{Y}_k = \sum_{j=1}^{Y_{k-1,1}} \begin{bmatrix} \xi_{k,j,1,1} \\ 1 \end{bmatrix} + \sum_{j=1}^{Y_{k-1,2}} \begin{bmatrix} \xi_{k,j,2,1} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ 0 \end{bmatrix}, \qquad k \in \mathbb{N},$$

where $\boldsymbol{Y}_0$ is a 2-dimensional integer-valued random vector. Here, for each $k, j \in \mathbb{N}$ and $i \in \{1, 2\}$, $\xi_{k,j,i,1}$ denotes the number of type 1 offsprings in the k^{th} generation produced by the j^{th} offspring of the $(k-1)^{\mathrm{th}}$ generation of type i, and ε_k denotes the number of type 1 immigrants in the k^{th} generation. For the second coordinate process of $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$, we get $Y_{k,2} = Y_{k-1,1}$, $k \in \mathbb{N}$, and substituting this into (A.3), the first coordinate process of $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ satisfies

$$Y_{k,1} = \sum_{j=1}^{Y_{k-1,1}} \xi_{k,j,1,1} + \sum_{j=1}^{Y_{k-2,1}} \xi_{k,j,2,1} + \varepsilon_{k,1}, \qquad k \geqslant 2.$$

Thus, the first coordinate process of $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ given by (A.3) satisfies equation (1.2) with $X_n := Y_{n,1}$, $\xi_{n,i} := \xi_{n,i,1,1}$, $\eta_{n,j} := \xi_{n,j,2,1}$, $\varepsilon_n := \varepsilon_{n,1}$, $n, i, j \in \mathbb{N}$, and with initial values $X_0 := Y_{0,1}$ and $X_{-1} := Y_{0,2}$, i.e., it is a second-order Galton–Watson process with immigration. Moreover, the second coordinate process of $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ also satisfies equation (1.2) with $X_n := Y_{n+1,2}$, $\xi_{n,i} := \xi_{n,i,1,1}$, $\eta_{n,j} := \xi_{n,j,2,1}$, $\varepsilon_n := \varepsilon_{n,1}$, $n, i, j \in \mathbb{N}$, and with initial values $X_0 := Y_{0,1}$ and $X_{-1} := Y_{0,2}$, i.e., it is also a second-order Galton–Watson process with immigration.

Note that, for a second-order Galton–Watson process $(X_n)_{n\geqslant -1}$ (without immigration), the additive (or branching) property of a 2-type Galton–Watson process (without immigration) (see, e.g. in Athreya and Ney [1, Chapter V, Section 1]), together with the law of total probability, for each $n \in \mathbb{N}$, imply

$$X_n \stackrel{\mathcal{D}}{=} \sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} + \sum_{j=1}^{X_{-1}} \zeta_{j,-1}^{(n)}, \tag{A.4}$$

where $\big\{(X_0, X_{-1}), \zeta_{i,0}^{(n)}, \zeta_{j,-1}^{(n)} : i, j \in \mathbb{N}\big\}$ are independent random variables such that $\{\zeta_{i,0}^{(n)} : i \in \mathbb{N}\}$ are independent copies of $V_{n,0}$ and $\{\zeta_{j,-1}^{(n)} : j \in \mathbb{N}\}$ are independent copies of $V_{n,-1}$, where $(V_{k,0})_{k\geqslant -1}$ and $(V_{k,-1})_{k\geqslant -1}$ are second-order Galton–Watson processes (without immigration) with initial values $V_{0,0} = 1$, $V_{-1,0} = 0$, $V_{0,-1} = 0$ and $V_{-1,-1} = 1$, and with the same offspring distributions as $(X_k)_{k\geqslant -1}$.

Moreover, if $(X_n)_{n\geqslant -1}$ is a second-order Galton–Watson process with immigration, then for each $n \in \mathbb{N}$, we have

$$X_n = V_0^{(n)}(X_0, X_{-1}) + \sum_{i=1}^{n} V_i^{(n-i)}(\varepsilon_i, 0), \tag{A.5}$$

where $\big\{V_0^{(n)}(X_0, X_{-1}), V_i^{(n-i)}(\varepsilon_i, 0) : i \in \{1, \ldots, n\}\big\}$ are independent random variables such that $V_0^{(n)}(X_0, X_{-1})$ represents the number of newborns at time n, resulting from the initial individuals X_0 at time 0 and X_{-1} at time -1, and for each $i \in \{1, \ldots, n\}$, $V_i^{(n-i)}(\varepsilon_i, 0)$ represents the number of newborns at time n, resulting from the immigration ε_i at time i. Indeed, considering the (special) 2-type Galton–Watson process $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$ with immigration given in (A.1) and applying formula (1.1) in Kaplan [14], we obtain

$$\boldsymbol{Y}_n = \boldsymbol{V}_0^{(n)}(\boldsymbol{Y}_0) + \sum_{i=1}^{n} \boldsymbol{V}_i^{(n-i)}(\boldsymbol{\varepsilon}_i) \qquad \text{with} \qquad \boldsymbol{\varepsilon}_i := \begin{bmatrix} \varepsilon_i \\ 0 \end{bmatrix}, \quad i \in \mathbb{N}, \tag{A.6}$$

for each $n \in \mathbb{N}$, where $\big\{\boldsymbol{V}_0^{(n)}(\boldsymbol{Y}_0), \boldsymbol{V}_i^{(n-i)}(\boldsymbol{\varepsilon}_i) : i \in \{1, \ldots, n\}\big\}$ are independent random vectors such that $\boldsymbol{V}_0^{(n)}(\boldsymbol{Y}_0)$ represents the number of individuals alive at time n, resulting from the initial individuals $\boldsymbol{Y}_0$ at time 0, and for each $i \in \{1, \ldots, n\}$, $\boldsymbol{V}_i^{(n-i)}(\boldsymbol{\varepsilon}_i)$ represents the number of individuals alive at time n, resulting from the immigration $\boldsymbol{\varepsilon}_i$ at time i. Clearly, $(\boldsymbol{V}_0^{(k)}(\boldsymbol{Y}_0))_{k\in\mathbb{Z}_+}$ and $(\boldsymbol{V}_i^{(k)}(\boldsymbol{\varepsilon}_i))_{k\in\mathbb{Z}_+}$, $i \in \{1, \ldots, n\}$, are (special) 2-type Galton–Watson processes (without immigration) of the form (A.3) with initial vectors $\boldsymbol{V}_0^{(0)}(\boldsymbol{Y}_0) = \boldsymbol{Y}_0$ and $\boldsymbol{V}_i^{(0)}(\boldsymbol{\varepsilon}_i) = \boldsymbol{\varepsilon}_i$, $i \in \{1, \ldots, n\}$, respectively, and with the same offspring distributions as $(\boldsymbol{Y}_k)_{k\in\mathbb{Z}_+}$. Using the considerations for the backward representation presented before, the first coordinates in (A.6) gives (A.5), where $(V_0^{(k)}(X_0, X_{-1}))_{k\geqslant -1}$ and $(V_i^{(k)}(\varepsilon_i, 0))_{k\geqslant -1}$, $i \in \{1, \ldots, n\}$, are second-order Galton–Watson processes (without immigration) with initial values $V_0^{(0)}(X_0, X_{-1}) = X_0$, $V_0^{(-1)}(X_0, X_{-1}) = X_{-1}$, $V_i^{(0)}(\varepsilon_i, 0) = \varepsilon_i$ and $V_i^{(-1)}(\varepsilon_i, 0) = 0$, $i \in \{1, \ldots, n\}$, and with the same offspring distributions as $(X_k)_{k\geqslant -1}$.

B On the expectation of second-order Galton–Watson processes with immigration

Our aim is to derive an explicit formula for the expectation of a second-order Galton–Watson process with immigration at time n and to describe its asymptotic behavior as $n \to \infty$.

Recall that ξ, η and ε are random variables such that $\xi \stackrel{\mathcal{D}}{=} \xi_{1,1}$, $\eta \stackrel{\mathcal{D}}{=} \eta_{1,1}$ and $\varepsilon \stackrel{\mathcal{D}}{=} \varepsilon_1$, and we put $m_\xi = \mathbb{E}(\xi) \in [0, \infty]$, $m_\eta = \mathbb{E}(\eta) \in [0, \infty]$ and $m_\varepsilon = \mathbb{E}(\varepsilon) \in [0, \infty]$. If $m_\xi \in \mathbb{R}_+$, $m_\eta \in \mathbb{R}_+$, $m_\varepsilon \in \mathbb{R}_+$, $\mathbb{E}(X_0) \in \mathbb{R}_+$ and $\mathbb{E}(X_{-1}) \in \mathbb{R}_+$, then (1.2) implies

$$\mathbb{E}(X_n \,|\, \mathcal{F}_{n-1}^X) = X_{n-1} m_\xi + X_{n-2} m_\eta + m_\varepsilon, \qquad n \in \mathbb{N},$$

where $\mathcal{F}_n^X := \sigma(X_{-1}, X_0, \ldots, X_n)$, $n \in \mathbb{Z}_+$. Consequently,

$$\mathbb{E}(X_n) = m_\xi \, \mathbb{E}(X_{n-1}) + m_\eta \, \mathbb{E}(X_{n-2}) + m_\varepsilon, \qquad n \in \mathbb{N},$$

which can be written in the matrix form

$$\text{(B.1)} \qquad \begin{bmatrix} \mathbb{E}(X_n) \\ \mathbb{E}(X_{n-1}) \end{bmatrix} = \boldsymbol{M}_{\xi,\eta} \begin{bmatrix} \mathbb{E}(X_{n-1}) \\ \mathbb{E}(X_{n-2}) \end{bmatrix} + \begin{bmatrix} m_\varepsilon \\ 0 \end{bmatrix}, \qquad n \in \mathbb{N},$$

with

$$\text{(B.2)} \qquad \boldsymbol{M}_{\xi,\eta} := \begin{bmatrix} m_\xi & m_\eta \\ 1 & 0 \end{bmatrix}.$$

Note that $\boldsymbol{M}_{\xi,\eta}$ is the mean matrix of the 2-type Galton–Watson process $(\boldsymbol{Y}_n)_{n \in \mathbb{Z}_+}$ given in (A.1). Thus, we conclude

$$\text{(B.3)} \qquad \begin{bmatrix} \mathbb{E}(X_n) \\ \mathbb{E}(X_{n-1}) \end{bmatrix} = \boldsymbol{M}_{\xi,\eta}^n \begin{bmatrix} \mathbb{E}(X_0) \\ \mathbb{E}(X_{-1}) \end{bmatrix} + \sum_{k=1}^n \boldsymbol{M}_{\xi,\eta}^{n-k} \begin{bmatrix} m_\varepsilon \\ 0 \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Hence, the asymptotic behavior of the sequence $(\mathbb{E}(X_n))_{n \in \mathbb{N}}$ depends on the asymptotic behavior of the powers $(\boldsymbol{M}_{\xi,\eta}^n)_{n \in \mathbb{N}}$, which is related to the spectral radius ϱ of $\boldsymbol{M}_{\xi,\eta}$, see Lemma B.1 and (B.6). If $(X_n)_{n \geqslant -1}$ is a second-order Galton–Watson process with immigration such that $m_\xi \in \mathbb{R}_+$ and $m_\eta \in \mathbb{R}_+$, then $(X_n)_{n \geqslant -1}$ is called subcritical, critical or supercritical if $\varrho < 1$, $\varrho = 1$ or $\varrho > 1$, respectively. It is easy to check that a second-order Galton–Watson process with immigration is subcritical, critical or supercritical if and only if $m_\xi + m_\eta < 1$, $m_\xi + m_\eta = 1$ or $m_\xi + m_\eta > 1$, respectively. We call the attention that for the classification of second-order Galton–Watson process with immigration we do not suppose the finiteness of the expectation of X_0, X_{-1} or ε.

B.1 Lemma. *Let $(X_n)_{n \geqslant -1}$ be a second-order Galton–Watson process with immigration such that $m_\xi \in \mathbb{R}_+$, $m_\eta \in \mathbb{R}_+$, $m_\varepsilon \in \mathbb{R}_+$, $\mathbb{E}(X_0) \in \mathbb{R}_+$ and $\mathbb{E}(X_{-1}) \in \mathbb{R}_+$.*

If $m_\xi = 0$ and $m_\eta = 0$, then, for all $n \in \mathbb{N}$, we have $\mathbb{E}(X_n) = m_\varepsilon$.

If $m_\xi + m_\eta > 0$, *then, for all* $n \in \mathbb{N}$, *we have*

$$\mathbb{E}(X_n) = \frac{\lambda_+^{n+1} - \lambda_-^{n+1}}{\lambda_+ - \lambda_-}\,\mathbb{E}(X_0) + \frac{\lambda_+^n - \lambda_-^n}{\lambda_+ - \lambda_-} m_\eta\,\mathbb{E}(X_{-1}) + \frac{C_n(\lambda_+, \lambda_-)}{\lambda_+ - \lambda_-} m_\varepsilon \tag{B.4}$$

with

$$C_n(\lambda_+, \lambda_-) := \begin{cases} \lambda_+ \frac{1-\lambda_+^n}{1-\lambda_+} - \lambda_- \frac{1-\lambda_-^n}{1-\lambda_-} & \text{if } \lambda_+ \neq 1, \\ n - \lambda_- \frac{1-\lambda_-^n}{1-\lambda_-} & \text{if } \lambda_+ = 1, \end{cases}$$

where λ_+ *and* λ_- *are given in* (2.1), *and hence*

$$\mathbb{E}(X_n) = \begin{cases} \frac{m_\varepsilon}{(1-\lambda_+)(1-\lambda_-)} + \mathrm{O}(\lambda_+^n) & \text{if } \lambda_+ \in (0,1), \\ \frac{m_\varepsilon}{1-\lambda_-} n + \mathrm{O}(1) & \text{if } \lambda_+ = 1, \\ \frac{1}{\lambda_+ - \lambda_-}\big(\lambda_+\,\mathbb{E}(X_0) + m_\eta\,\mathbb{E}(X_{-1}) + \frac{\lambda_+}{\lambda_+ - 1} m_\varepsilon\big)\lambda_+^n + \mathrm{O}(1 + |\lambda_-|^n) & \text{if } \lambda_+ \in (1,\infty) \end{cases}$$

as $n \to \infty$. *Moreover,* λ_+ *is the spectral radius* ϱ *of* $\boldsymbol{M}_{\xi,\eta}$.

Further, in case of $m_\varepsilon = 0$, *we have the following more precise statements:*

If $m_\xi = 0$, $m_\eta > 0$ *and* $m_\varepsilon = 0$, *then, for all* $k \in \mathbb{N}$, *we have* $\mathbb{E}(X_{2k-1}) = \mathbb{E}(X_{-1})\lambda_+^{2k}$ *and* $\mathbb{E}(X_{2k}) = \mathbb{E}(X_0)\lambda_+^{2k}$.

If $m_\xi > 0$, $m_\eta = 0$ *and* $m_\varepsilon = 0$, *then, for all* $n \in \mathbb{N}$, *we have* $\mathbb{E}(X_n) = \mathbb{E}(X_0)\lambda_+^n$.

If $m_\xi > 0$, $m_\eta > 0$ *and* $m_\varepsilon = 0$, *then*

$$\mathbb{E}(X_n) = \frac{\lambda_+\,\mathbb{E}(X_0) + m_\eta\,\mathbb{E}(X_{-1})}{\lambda_+ - \lambda_-}\lambda_+^n + \mathrm{O}(|\lambda_-|^n) \qquad \text{as } n \to \infty.$$

If $m_\varepsilon = 0$, *i.e., there is no immigration, then*

$$\mathbb{E}(X_n) \leqslant \varrho^n\,\mathbb{E}(X_0) + \varrho^{n-1} m_\eta\,\mathbb{E}(X_{-1}), \qquad n \in \mathbb{N}. \tag{B.5}$$

Proof. We are going to use (B.3). The matrix $\boldsymbol{M}_{\xi,\eta}$ has eigenvalues

$$\lambda_+ = \frac{m_\xi + \sqrt{m_\xi^2 + 4m_\eta}}{2}, \qquad \lambda_- = \frac{m_\xi - \sqrt{m_\xi^2 + 4m_\eta}}{2},$$

satisfying $\lambda_+ \in \mathbb{R}_+$ and $\lambda_- \in [-\lambda_+, 0]$, hence the spectral radius of $\boldsymbol{M}_{\xi,\eta}$ is

$$\varrho = \lambda_+ = \frac{m_\xi + \sqrt{m_\xi^2 + 4m_\eta}}{2}. \tag{B.6}$$

In what follows, we suppose that $m_\xi + m_\eta > 0$, which yields that $\lambda_+ \in \mathbb{R}_{++}$ and $\lambda_- \in (-\lambda_+, 0]$. One can easily check that the powers of $\boldsymbol{M}_{\xi,\eta}$ can be written in the form

$$\boldsymbol{M}_{\xi,\eta}^n = \frac{\lambda_+^n}{\lambda_+ - \lambda_-}\begin{bmatrix} \lambda_+ & m_\eta \\ 1 & -\lambda_- \end{bmatrix} + \frac{\lambda_-^n}{\lambda_+ - \lambda_-}\begin{bmatrix} -\lambda_- & -m_\eta \\ -1 & \lambda_+ \end{bmatrix}, \qquad n \in \mathbb{Z}_+. \tag{B.7}$$

Consequently,

$$\varrho^{-n}\boldsymbol{M}_{\xi,\eta}^{n} \to \frac{1}{\lambda_+ - \lambda_-}\begin{bmatrix}\lambda_+ & m_\eta \\ 1 & -\lambda_-\end{bmatrix} \qquad \text{as} \quad n \to \infty.$$

Moreover, (B.3) and (B.7) yield

$$\begin{bmatrix}\mathbb{E}(X_n) \\ \mathbb{E}(X_{n-1})\end{bmatrix} = \frac{\mathbb{E}(X_0)}{\lambda_+ - \lambda_-}\begin{bmatrix}\lambda_+^{n+1} - \lambda_-^{n+1} \\ \lambda_+^{n} - \lambda_-^{n}\end{bmatrix} + \frac{\mathbb{E}(X_{-1})}{\lambda_+ - \lambda_-}\begin{bmatrix}m_\eta(\lambda_+^{n} - \lambda_-^{n}) \\ -\lambda_-\lambda_+^{n} + \lambda_+\lambda_-^{n}\end{bmatrix}$$

$$+ \frac{m_\varepsilon}{\lambda_+ - \lambda_-}\sum_{k=1}^{n}\begin{bmatrix}\lambda_+^{n-k+1} - \lambda_-^{n-k+1} \\ \lambda_+^{n-k} - \lambda_-^{n-k}\end{bmatrix}, \qquad n \in \mathbb{Z}_+,$$

and hence, we obtain (B.4) and (B.5). Indeed, by (B.7) and by $\lambda_+ \in \mathbb{R}_{++}$ and $-\lambda_+ < \lambda_- \leqslant 0$, for each $k \in \mathbb{Z}_+$, we have

$$\frac{\lambda_+^{2k+1} - \lambda_-^{2k+1}}{\lambda_+ - \lambda_-} = \sum_{i=0}^{2k}\lambda_-^{i}\lambda_+^{2k-i} = \lambda_+^{2k} + \sum_{j=1}^{k}(\lambda_-^{2j-1}\lambda_+^{2k-2j+1} + \lambda_-^{2j}\lambda_+^{2k-2j}) \leqslant \lambda_+^{2k},$$

since $\lambda_-^{2j-1}\lambda_+^{2k-2j+1} + \lambda_-^{2j}\lambda_+^{2k-2j} = \lambda_-^{2j-1}\lambda_+^{2k-2j}(\lambda_+ + \lambda_-) \leqslant 0$, and, in a similar way,

$$\frac{\lambda_+^{2k+2} - \lambda_-^{2k+2}}{\lambda_+ - \lambda_-} = \sum_{i=0}^{2k+1}\lambda_-^{i}\lambda_+^{2k+1-i} = \lambda_+^{2k+1} + \sum_{j=1}^{k}(\lambda_-^{2j-1}\lambda_+^{2k-2j+1} + \lambda_-^{2j}\lambda_+^{2k-2j}) + \lambda_-^{2k+1} \leqslant \lambda_+^{2k+1}.$$

Further, if $\lambda_+ \in (0, 1)$, then

$$\frac{m_\varepsilon}{\lambda_+ - \lambda_-}C_n(\lambda_+, \lambda_-) = \frac{m_\varepsilon}{\lambda_+ - \lambda_-}\frac{\lambda_+ - \lambda_- + \lambda_+^{n+1}(\lambda_- - 1) + \lambda_-^{n+1}(1 - \lambda_+)}{(1 - \lambda_+)(1 - \lambda_-)}$$

$$= \frac{m_\varepsilon}{(1 - \lambda_+)(1 - \lambda_-)} + \mathrm{O}(\lambda_+^{n})$$

as $n \to \infty$. The other statements easily follow from (B.4). □

C Moment estimations

The first moment of a second-order Galton–Watson process $(X_n)_{n\geqslant -1}$ (without immigration) can be estimated by (B.5). Next, we present an auxiliary lemma on higher moments of $(X_n)_{n\geqslant -1}$.

C.1 Lemma. *Let $(X_n)_{n\geqslant -1}$ be a second-order Galton–Watson process (without immigration) such that $\mathbb{E}(X_{-1}^r) < \infty$, $\mathbb{E}(X_0^r) < \infty$, $\mathbb{E}(\xi^r) < \infty$ and $\mathbb{E}(\eta^r) < \infty$ with some $r > 1$. Then $\mathbb{E}(X_n^r) < \infty$ for all $n \in \mathbb{N}$.*

Proof. By power means inequality, we have

$$\begin{aligned}
\mathbb{E}(X_n^r \,|\, \mathcal{F}_{n-1}^X) &= \mathbb{E}\left(\left(\sum_{i=1}^{X_{n-1}} \xi_{n,i} + \sum_{j=1}^{X_{n-2}} \eta_{n,j}\right)^r \,\middle|\, \mathcal{F}_{n-1}^X\right) \\
&\leqslant 2^{r-1}\,\mathbb{E}\left(\left(\sum_{i=1}^{X_{n-1}} \xi_{n,i}\right)^r + \left(\sum_{j=1}^{X_{n-2}} \eta_{n,j}\right)^r \,\middle|\, \mathcal{F}_{n-1}^X\right) \\
&\leqslant 2^{r-1}\,\mathbb{E}\left(X_{n-1}^{r-1}\sum_{i=1}^{X_{n-1}} \xi_{n,i}^r + X_{n-2}^{r-1}\sum_{j=1}^{X_{n-2}} \eta_{n,j}^r \,\middle|\, \mathcal{F}_{n-1}^X\right) \\
&= 2^{r-1}\left(X_{n-1}^r\,\mathbb{E}(\xi^r) + X_{n-2}^r\,\mathbb{E}(\eta^r)\right) < \infty
\end{aligned}$$

for all $n \in \mathbb{N}$. Hence $\mathbb{E}(X_n^r) \leqslant 2^{r-1}\big(\mathbb{E}(X_{n-1}^r)\,\mathbb{E}(\xi^r) + \mathbb{E}(X_{n-2}^r)\,\mathbb{E}(\eta^r)\big)$, $n \in \mathbb{N}$. By induction we obtain the statement. □

Moreover, we present an auxiliary lemma on an estimation of the second moment of a second-order Galton–Watson process (without immigration). This lemma is valid for the subcritical, critical and supercritical cases as well, however, in the proofs we only use it for the subcritical case.

C.2 Lemma. *Let $(X_n)_{n\geqslant -1}$ be a second-order Galton–Watson process (without immigration) such that $X_0 = 1$, $X_{-1} = 0$, $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{E}(\eta^2) < \infty$. Then for all $n \in \mathbb{N}$,*

$$\mathbb{E}(X_n^2) \leqslant \begin{cases} c_{\mathrm{sub}}\,\varrho^n, & \text{if } \varrho \in (0,1), \\ c_{\mathrm{crit}}\,n, & \text{if } \varrho = 1, \\ c_{\mathrm{sup}}\,\varrho^{2n}, & \text{if } \varrho \in (1,\infty), \end{cases} \tag{C.1}$$

where

$$c_{\mathrm{sub}} := 1 + \frac{\mathrm{Var}(\xi)}{\varrho(1-\varrho)} + \frac{\mathrm{Var}(\eta)}{\varrho^2(1-\varrho)}, \qquad c_{\mathrm{crit}} := 1 + \mathrm{Var}(\xi) + \mathrm{Var}(\eta), \qquad c_{\mathrm{sup}} := 1 + \frac{\mathrm{Var}(\xi)}{\varrho(\varrho-1)} + \frac{\mathrm{Var}(\eta)}{\varrho^3(\varrho-1)}.$$

Proof. By formula (A2) in Lemma A.1 in Ispány and Pap [13], we have

$$\mathrm{Var}(\boldsymbol{Y}_n) = \sum_{j=0}^{n-1} \boldsymbol{M}_{\xi,\eta}^j \big[(\boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1)\boldsymbol{V}_\xi + (\boldsymbol{e}_2^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1)\boldsymbol{V}_\eta\big](\boldsymbol{M}_{\xi,\eta}^\top)^j,$$

where $(\boldsymbol{Y}_n)_{n\in\mathbb{Z}_+}$ is given by (A.1) with $\boldsymbol{Y}_0 = [1\ 0]^\top$, and

$$\boldsymbol{V}_\xi := \mathrm{Var}\left(\begin{bmatrix}\xi\\1\end{bmatrix}\right) = \mathrm{Var}(\xi)\boldsymbol{e}_1\boldsymbol{e}_1^\top, \qquad \boldsymbol{V}_\eta := \mathrm{Var}\left(\begin{bmatrix}\eta\\0\end{bmatrix}\right) = \mathrm{Var}(\eta)\boldsymbol{e}_1\boldsymbol{e}_1^\top,$$

where ξ and η are random variables such that $\xi \stackrel{\mathcal{D}}{=} \xi_{1,1}$ and $\eta \stackrel{\mathcal{D}}{=} \eta_{1,1}$. Here we note that formula (A2) in Lemma A.1 in Ispány and Pap [13] is stated only for critical processes, but it also holds in the subcritical and supercritical cases as well; the proof is the very same. Consequently,

$$
\begin{aligned}
\operatorname{Var}(X_n) &= \operatorname{Var}(\boldsymbol{e}_1^\top \boldsymbol{Y}_n) = \boldsymbol{e}_1^\top \operatorname{Var}(\boldsymbol{Y}_n)\boldsymbol{e}_1 \\
&= \boldsymbol{e}_1^\top \sum_{j=0}^{n-1} \boldsymbol{M}_{\xi,\eta}^j \big[(\boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1)\operatorname{Var}(\xi)\boldsymbol{e}_1\boldsymbol{e}_1^\top + (\boldsymbol{e}_2^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1)\operatorname{Var}(\eta)\boldsymbol{e}_1\boldsymbol{e}_1^\top\big](\boldsymbol{M}_{\xi,\eta}^\top)^j \boldsymbol{e}_1 \\
&= \sum_{j=0}^{n-1} (\boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^j \boldsymbol{e}_1)^2 \big[\operatorname{Var}(\xi)(\boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1) + \operatorname{Var}(\eta)(\boldsymbol{e}_2^\top \boldsymbol{M}_{\xi,\eta}^{n-j-1}\boldsymbol{e}_1)\big],
\end{aligned}
$$

where we used that $\boldsymbol{e}_1^\top (\boldsymbol{M}_{\xi,\eta}^\top)^j \boldsymbol{e}_1 = \boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^j \boldsymbol{e}_1$. Using (B.3) with $X_0 = 1$ and $X_{-1} = 0$, we have $\boldsymbol{e}_1^\top \boldsymbol{M}_{\xi,\eta}^j \boldsymbol{e}_1 = \mathbb{E}(X_j)$ and $\boldsymbol{e}_2^\top \boldsymbol{M}_{\xi,\eta}^j \boldsymbol{e}_1 = \mathbb{E}(X_{j-1})$ for each $j \in \mathbb{Z}_+$, hence

$$
\operatorname{Var}(X_n) = \operatorname{Var}(\xi)\sum_{j=0}^{n-1} [\mathbb{E}(X_j)]^2\, \mathbb{E}(X_{n-j-1}) + \operatorname{Var}(\eta)\sum_{j=0}^{n-2} [\mathbb{E}(X_j)]^2\, \mathbb{E}(X_{n-j-2}),
$$

where we used that $X_{-1} = 0$. We note that the above formula for $\operatorname{Var}(X_n)$ can also be found in Kashikar and Deshmukh [16, page 562]. Using (B.5) with $X_0 = 1$ and $X_{-1} = 0$, we obtain

$$
\begin{aligned}
\mathbb{E}(X_n^2) = \operatorname{Var}(X_n) + [\mathbb{E}(X_n)]^2 &\leqslant \operatorname{Var}(\xi)\sum_{j=0}^{n-1} \varrho^{n+j-1} + \operatorname{Var}(\eta)\sum_{j=0}^{n-2} \varrho^{n+j-2} + \varrho^{2n} \\
&= \begin{cases} n\operatorname{Var}(\xi) + (n-1)\operatorname{Var}(\eta) + 1, & \text{if } \varrho = 1, \\ \operatorname{Var}(\xi)\,\frac{\varrho^{n-1}-\varrho^{2n-1}}{1-\varrho} + \operatorname{Var}(\eta)\,\frac{\varrho^{n-2}-\varrho^{2n-3}}{1-\varrho} + \varrho^{2n}, & \text{if } \varrho \neq 1, \end{cases}
\end{aligned}
$$

yielding (C.1). Indeed, for example, if $\varrho \in (1, \infty)$, then

$$
\frac{\varrho^{n-2} - \varrho^{2n-3}}{1-\varrho} = \frac{\varrho^{2n-3}(1-\varrho^{-n+1})}{\varrho - 1} \leqslant \frac{\varrho^{2n-3}}{\varrho-1} = \frac{\varrho^{2n}}{\varrho^3(\varrho-1)}, \qquad n \in \mathbb{N}.
$$

□

D Representation of the unique stationary distribution for 2-type Galton–Watson processes with immigration

First, we introduce 2-type Galton–Watson processes with immigration. For each $k, j \in \mathbb{Z}_+$ and $i, \ell \in \{1, 2\}$, the number of individuals of type i born or arrived as immigrants in the k^{th} generation will be denoted by $X_{k,i}$, the number of type ℓ offsprings produced by the j^{th} individual who is of type i belonging to the $(k-1)^{\mathrm{th}}$ generation will be denoted by $\xi_{k,j,i,\ell}$, and the number of type i immigrants in the k^{th} generation will be denoted by $\varepsilon_{k,i}$. Then we have

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

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

$$\boldsymbol{X}_0 := \begin{bmatrix} X_{0,1} \\ X_{0,2} \end{bmatrix}, \qquad \boldsymbol{X}_k := \begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix}, \qquad \boldsymbol{\xi}_{k,j,i} := \begin{bmatrix} \xi_{k,j,i,1} \\ \xi_{k,j,i,2} \end{bmatrix}, \qquad \boldsymbol{\varepsilon}_k := \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}.$$

For notational convenience, let $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ and $\boldsymbol{\varepsilon}$ be random vectors such that $\boldsymbol{\xi}_1 \stackrel{\mathcal{D}}{=} \boldsymbol{\xi}_{1,1,1}$, $\boldsymbol{\xi}_2 \stackrel{\mathcal{D}}{=} \boldsymbol{\xi}_{1,1,2}$ and $\boldsymbol{\varepsilon} \stackrel{\mathcal{D}}{=} \boldsymbol{\varepsilon}_1$, and put $\boldsymbol{m}_{\boldsymbol{\xi}_1} := \mathbb{E}(\boldsymbol{\xi}_1) \in [0,\infty]^2$, $\boldsymbol{m}_{\boldsymbol{\xi}_2} := \mathbb{E}(\boldsymbol{\xi}_2) \in [0,\infty]^2$, and $\boldsymbol{m}_{\boldsymbol{\varepsilon}} := \mathbb{E}(\boldsymbol{\varepsilon}) \in [0,\infty]^2$, and put

$$\boldsymbol{M}_{\boldsymbol{\xi}} := \begin{bmatrix} \boldsymbol{m}_{\boldsymbol{\xi}_1} & \boldsymbol{m}_{\boldsymbol{\xi}_2} \end{bmatrix} \in [0,\infty]^{2\times 2}.$$

We call $\boldsymbol{M}_{\boldsymbol{\xi}}$ the offspring mean matrix, and note that many authors define the offspring mean matrix as $\boldsymbol{M}_{\boldsymbol{\xi}}^\top$. If $\boldsymbol{m}_{\boldsymbol{\xi}_1} \in \mathbb{R}_+^2$, $\boldsymbol{m}_{\boldsymbol{\xi}_2} \in \mathbb{R}_+^2$, and $\boldsymbol{m}_{\boldsymbol{\varepsilon}} \in \mathbb{R}_+^2$, then for each $n \in \mathbb{Z}_+$, (D.1) implies

$$\mathbb{E}(\boldsymbol{X}_n \,|\, \mathcal{F}_{n-1}^{\boldsymbol{X}}) = X_{n-1,1}\, \boldsymbol{m}_{\boldsymbol{\xi}_1} + X_{n-1,2}\, \boldsymbol{m}_{\boldsymbol{\xi}_2} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} = \boldsymbol{M}_{\boldsymbol{\xi}}\, \boldsymbol{X}_{n-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad n \in \mathbb{N},$$

where $\mathcal{F}_n^{\boldsymbol{X}} := \sigma\big(\boldsymbol{X}_0, \ldots, \boldsymbol{X}_n\big)$, $n \in \mathbb{Z}_+$. Consequently, $\mathbb{E}(\boldsymbol{X}_n) = \boldsymbol{M}_{\boldsymbol{\xi}}\, \mathbb{E}(\boldsymbol{X}_{n-1}) + \boldsymbol{m}_{\boldsymbol{\varepsilon}}$, $n \in \mathbb{N}$, which implies

$$\mathbb{E}(\boldsymbol{X}_n) = \boldsymbol{M}_{\boldsymbol{\xi}}^n\, \mathbb{E}(\boldsymbol{X}_0) + \sum_{k=1}^{n} \boldsymbol{M}_{\boldsymbol{\xi}}^{n-k} \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad n \in \mathbb{N}.$$

Hence, the asymptotic behavior of the sequence $(\mathbb{E}(\boldsymbol{X}_n))_{n\in\mathbb{Z}_+}$ depends on the asymptotic behavior of the powers $(\boldsymbol{M}_{\boldsymbol{\xi}}^n)_{n\in\mathbb{N}}$ of the offspring mean matrix, which is related to the spectral radius $r(\boldsymbol{M}_{\boldsymbol{\xi}}) \in \mathbb{R}_+$ of $\boldsymbol{M}_{\boldsymbol{\xi}}$ (see the Frobenius–Perron theorem, e.g., Horn and Johnson [11, Theorems 8.2.8 and 8.5.1]). A 2-type Galton–Watson process $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ with immigration is referred to respectively as *subcritical*, *critical* or *supercritical* if $r(\boldsymbol{M}_{\boldsymbol{\xi}}) < 1$, $r(\boldsymbol{M}_{\boldsymbol{\xi}}) = 1$ or $r(\boldsymbol{M}_{\boldsymbol{\xi}}) > 1$ (see, e.g., Athreya and Ney [1, V.3] or Quine [22]). We extend this classification for all 2-type Galton–Watson processes with immigration.

If $\boldsymbol{m}_{\boldsymbol{\xi}_1} \in \mathbb{R}_+^2$, $\boldsymbol{m}_{\boldsymbol{\xi}_2} \in \mathbb{R}_+^2$, $r(\boldsymbol{M}_{\boldsymbol{\xi}}) < 1$, $\boldsymbol{M}_{\boldsymbol{\xi}}$ is primitive, i.e., there exists $m \in \mathbb{N}$ such that $\boldsymbol{M}_{\boldsymbol{\xi}}^m \in \mathbb{R}_{++}^{2\times 2}$, $\mathbb{P}(\boldsymbol{\varepsilon} = \boldsymbol{0}) < 1$ and $\mathbb{E}(\mathbb{1}_{\{\boldsymbol{\varepsilon}\neq\boldsymbol{0}\}} \log((\boldsymbol{e}_1 + \boldsymbol{e}_2)^\top \boldsymbol{\varepsilon})) < \infty$, then, by the Theorem in Quine [22], there exists a unique stationary distribution $\boldsymbol{\pi}$ for $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$. As a consequence of formula (16) for the probability generating function of $\boldsymbol{\pi}$ in Quine [22], we have

$$\sum_{i=0}^{n} \boldsymbol{V}_i^{(i)}(\boldsymbol{\varepsilon}_i) \xrightarrow{\mathcal{D}} \boldsymbol{\pi} \qquad \text{as } n \to \infty,$$

where $(\boldsymbol{V}_k^{(i)}(\boldsymbol{\varepsilon}_i))_{k\in\mathbb{Z}_+}$, $i \in \mathbb{Z}_+$, are independent copies of a 2-type Galton–Watson process $(\boldsymbol{V}_k(\boldsymbol{\varepsilon}))_{k\in\mathbb{Z}_+}$ (without immigration) with initial vector $\boldsymbol{V}_0(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}$ and with the same offspring distributions as $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$. Consequently, we have

$$\sum_{i=0}^{\infty} \boldsymbol{V}_i^{(i)}(\boldsymbol{\varepsilon}_i) \stackrel{\mathcal{D}}{=} \boldsymbol{\pi},$$

where the series $\sum_{i=0}^{\infty} \boldsymbol{V}_i^{(i)}(\boldsymbol{\varepsilon}_i)$ converges with probability 1, see, e.g., Heyer [10, Theorem 3.1.6]. The above representation of the stationary distribution $\boldsymbol{\pi}$ for $(\boldsymbol{X}_n)_{n\in\mathbb{Z}_+}$ can be interpreted in a way that we consider independent 2-type Galton–Watson processes without immigration such that the i^{th} one admits initial vector $\boldsymbol{\varepsilon}_i$, $i \in \mathbb{Z}_+$, evaluate the i^{th} 2-type Galton-Watson processes at time point i, and then sum up all these random variables.

E Regularly varying distributions

First, we recall the notions of slowly varying and regularly varying functions, respectively.

E.1 Definition. *A measurable function $U : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is called regularly varying at infinity with index $\rho \in \mathbb{R}$ if for all $q \in \mathbb{R}_{++}$,*

$$\lim_{x\to\infty} \frac{U(qx)}{U(x)} = q^{\rho}.$$

In case of $\rho = 0$, U is called slowly varying at infinity.

Next, we recall the notion of regularly varying random variables.

E.2 Definition. *A non-negative random variable X is called regularly varying with index $\alpha \in \mathbb{R}_+$ if $U(x) := \mathbb{P}(X > x) \in \mathbb{R}_{++}$ for all $x \in \mathbb{R}_{++}$, and U is regularly varying at infinity with index $-\alpha$.*

E.3 Lemma. *If ζ is a non-negative regularly varying random variable with index $\alpha \in \mathbb{R}_+$, then for each $c \in \mathbb{R}_{++}$, ζ^c is regularly varying with index $\frac{\alpha}{c}$.*

Proof. For any $q \in \mathbb{R}_{++}$, we have

$$\lim_{x\to\infty} \frac{\mathbb{P}(\zeta^c > qx)}{\mathbb{P}(\zeta^c > x)} = \lim_{x\to\infty} \frac{\mathbb{P}(\zeta > q^{1/c}x^{1/c})}{\mathbb{P}(\zeta > x^{1/c})} = q^{-\alpha/c},$$

as desired. □

E.4 Lemma. *If $L : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is a slowly varying function (at infinity), then*

$$\lim_{x\to\infty} x^{\delta} L(x) = \infty, \qquad \lim_{x\to\infty} x^{-\delta} L(x) = 0, \qquad \delta \in \mathbb{R}_{++}.$$

For Lemma E.4, see, Bingham et al. [4, Proposition 1.3.6. (v)].

E.5 Lemma. *If* ε *is a non-negative regularly varying random variable with index* $\alpha \in \mathbb{R}_{++}$*, then* $\mathbb{E}(\mathbb{1}_{\{\varepsilon \neq 0\}} \log(\varepsilon)) < \infty$ *and* $\mathbb{E}(\log(\varepsilon + 1)) < \infty$*.*

Proof. Since $\mathbb{E}(\mathbb{1}_{\{\varepsilon \neq 0\}} \log(\varepsilon)) \leqslant \mathbb{E}(\log(\varepsilon + 1))$, it is enough to prove that $\mathbb{E}(\log(\varepsilon+1)) < \infty$. Since $\log(\varepsilon + 1) \geqslant 0$, we have

$$\begin{aligned}
\mathbb{E}(\log(\varepsilon + 1)) &= \int_0^\infty \mathbb{P}(\log(\varepsilon + 1) \geqslant x) \, \mathrm{d}x = \int_0^\infty \mathbb{P}(\varepsilon \geqslant \mathrm{e}^x - 1) \, \mathrm{d}x \\
&= \int_0^1 \mathbb{P}(\varepsilon \geqslant \mathrm{e}^x - 1) \, \mathrm{d}x + \int_1^\infty \mathbb{P}(\varepsilon \geqslant \mathrm{e}^x - 1) \, \mathrm{d}x := I_1 + I_2.
\end{aligned}$$

Here $I_1 \leqslant 1$, and, by substitution $y = \mathrm{e}^x - 1$,

$$I_2 = \int_{\mathrm{e}-1}^\infty y^{-\alpha} L(y) \frac{1}{1+y} \, \mathrm{d}y,$$

where $L(y) := y^\alpha \, \mathbb{P}(\varepsilon > y)$, $y \in \mathbb{R}_{++}$, is a slowly varying function. By Lemma E.4, there exists $y_0 \in (\mathrm{e} - 1, \infty)$ such that $y^{-\frac{\alpha}{2}} L(y) \leqslant 1$ for all $y \in [y_0, \infty)$. Hence

$$\begin{aligned}
I_2 &= \int_{\mathrm{e}-1}^{y_0} y^{-\alpha} L(y) \frac{1}{1+y} \, \mathrm{d}y + \int_{y_0}^\infty y^{-\alpha} L(y) \frac{1}{1+y} \, \mathrm{d}y \\
&\leqslant \int_{\mathrm{e}-1}^{y_0} y^{-\alpha} L(y) \frac{1}{1+y} \, \mathrm{d}y + \int_{y_0}^\infty y^{-\frac{\alpha}{2}} \frac{1}{1+y} \, \mathrm{d}y \\
&\leqslant \int_{\mathrm{e}-1}^{y_0} y^{-\alpha} L(y) \frac{1}{1+y} \, \mathrm{d}y + \int_{y_0}^\infty y^{-\frac{\alpha}{2}-1} \, \mathrm{d}y \\
&\leqslant \int_{\mathrm{e}-1}^{y_0} \frac{1}{1+y} \, \mathrm{d}y + \int_{y_0}^\infty y^{-\frac{\alpha}{2}-1} \, \mathrm{d}y < \infty,
\end{aligned}$$

since $y^{-\alpha} L(y) = \mathbb{P}(\varepsilon > y) \leqslant 1$ for all $y \in \mathbb{R}_{++}$. □

E.6 Lemma. *If* η *is a non-negative regularly varying random variable with index* $\alpha \in (1, 2)$*, then for every* $\varrho \in (\alpha, \infty)$*, there exist* $y_0 \in \mathbb{R}_{++}$ *and* $B \in \mathbb{R}_{++}$ *such that*

$$\frac{\mathbb{P}(\eta > z)}{\mathbb{P}(\eta > y)} \leqslant B \left(\frac{z}{y} \right)^{-\varrho}, \qquad y \geqslant z \geqslant y_0,$$

or equivalently,

$$\frac{\mathbb{P}(\eta > \theta y)}{\mathbb{P}(\eta > y)} \leqslant B \theta^{-\varrho}, \qquad \theta \in (0, 1], \qquad y \geqslant \frac{y_0}{\theta}.$$

For Lemma E.6, see Proposition 2.2.1 in Bingham et al. [4].

For the next lemma, see Faÿ et al. [9, Lemma 4.4]. Here we present a proof as well, since we state their result in a little bit extended form.

E.7 Lemma. *Let* $h : \mathbb{R}_+ \to \mathbb{R}_{++}$ *be a function such that* $\lim_{x\to\infty} h(x) = 0$. *Then there exists a monotone increasing, left-continuous, slowly varying (at infinity) function* L *such that* $L(x) \geqslant 1$, $x \in \mathbb{R}_+$, $\lim_{x\to\infty} L(x) = \infty$ *and* $\lim_{x\to\infty} L(x)h(x) = 0$. *One can also choose a version of* L *which is right-continuous with all the other properties remaining true.*

Proof. We can construct L as follows. Let $L(x) := 1$ for $x \in [0, x_0]$, where $x_0 := \sup\{y \in \mathbb{R}_+ : h(y) > 1\}$, and we define $\sup\emptyset := 0$. Since $\lim_{x\to\infty} h(x) = 0$, we have $x_0 \in \mathbb{R}_+$. Let $L(x) := 2$ for $x \in (x_0, x_1]$, where $x_1 := \max\{2x_0, \sup\{y \in \mathbb{R}_+ : h(y) > 2^{-2}\}\}$. Let $L(x) := 3$ for $x \in (x_1, x_2]$, where $x_2 := \max\{3x_1, \sup\{y \in \mathbb{R}_+ : h(y) > 3^{-2}\}\}$, and continue this construction in the straightforward way: $L(x) := k+1$ for $x \in (x_{k-1}, x_k]$, where $x_k := \max\{(k+1)x_{k-1}, \sup\{y \in \mathbb{R}_+ : h(y) > (k+1)^{-2}\}\}$, $k \in \mathbb{N}$. Since h takes positive values and $\lim_{x\to\infty} h(x) = 0$, we have $\lim_{x\to\infty} L(x) = \infty$, and, since for all $k \in \mathbb{Z}_+$ and $x > x_k$,

$$L(x)h(x) = \sum_{i=k}^{\infty} L(x)h(x)\mathbb{1}_{(x_i,x_{i+1}](x)} \leqslant \sum_{i=k}^{\infty}(i+2)\frac{1}{(i+1)^2}\mathbb{1}_{(x_i,x_{i+1}](x)} \leqslant \frac{k+2}{(k+1)^2},$$

we have $\lim_{x\to\infty} L(x)h(x) = 0$. It remains to check that L is slowly varying (at infinity). For this it is enough to verify that for any $q \in \mathbb{R}_{++}$ and sufficiently large $x \in \mathbb{R}_{++}$, we have x and qx are either in the same interval of type $(x_{k-1}, x_k]$ or in two neighbouring intervals of this type, since in this case for sufficiently large $x \in \mathbb{R}_{++}$:

$$\frac{L(qx)}{L(x)} \in \left\{1, \frac{k_x}{k_x+1}, \frac{k_x+1}{k_x}\right\}$$

with some $k_x \in \mathbb{N}$, and for sufficiently large $x \in \mathbb{R}_{++}$ and for $y \geqslant x$,

$$\left|\frac{L(qy)}{L(y)} - 1\right| \in \left\{0, \left|\frac{k_y+1}{k_y} - 1\right|, \left|\frac{k_y}{k_y+1} - 1\right|\right\} = \left\{0, \frac{1}{k_y}, \frac{1}{k_y+1}\right\},$$

where $k_y \geqslant k_x$ and $\lim_{y\to\infty} k_y = \infty$, yielding that $\lim_{x\to\infty} \frac{L(qx)}{L(x)} = 1$. To finish the proof, if $x \in (x_{k-1}, x_k]$ with some $k \in \mathbb{N}$, then in case of $q \geqslant 1$, we have $qx \in (x_{k-1}, x_k] \cup (x_k, x_{k+1}]$ provided that $k + 2 \geqslant q$, and in case of $q \in (0, 1)$, we have $qx \in (x_{k-2}, x_{k-1}] \cup (x_{k-1}, x_k]$ provided that $k > \frac{1}{q}$. Indeed, $x_k \geqslant (k+1)x_{k-1}$, $k \in \mathbb{N}$, and if $k + 2 \geqslant q$, then $qx \leqslant qx_k \leqslant (k+2)x_k \leqslant x_{k+1}$, as desired, and if $k > \frac{1}{q}$, then $qx > qx_{k-1} > \frac{1}{k}x_{k-1} \geqslant x_{k-2}$, as desired. □

E.8 Lemma. *If* X *and* Y *are non-negative random variables such that* X *is regularly varying with index* $\alpha \in \mathbb{R}_+$ *and there exists* $r \in (\alpha, \infty)$ *with* $\mathbb{E}(Y^r) < \infty$, *then* $\mathbb{P}(Y > x) = \mathrm{o}(\mathbb{P}(X > x))$ *as* $x \to \infty$.

For a proof of Lemma E.8, see, e.g., Barczy et al. [2, Lemma C.6].

E.9 Lemma. *If* X_1 *and* X_2 *are non-negative regularly varying random variables with index* $\alpha_1 \in \mathbb{R}_+$ *and* $\alpha_2 \in \mathbb{R}_+$, *respectively, such that* $\alpha_1 < \alpha_2$, *then* $\mathbb{P}(X_2 > x) = \mathrm{o}(\mathbb{P}(X_1 > x))$ *as* $x \to \infty$.

For a proof of Lemma E.9, see, e.g., Barczy et al. [2, Lemma C.7].

E.10 Lemma. (Convolution property) *If* X_1 *and* X_2 *are non-negative random variables such that* X_1 *is regularly varying with index* $\alpha_1 \in \mathbb{R}_+$ *and* $\mathbb{P}(X_2 > x) = \mathrm{o}(\mathbb{P}(X_1 > x))$ *as* $x \to \infty$, *then* $\mathbb{P}(X_1 + X_2 > x) \sim \mathbb{P}(X_1 > x)$ *as* $x \to \infty$, *and hence* $X_1 + X_2$ *is regularly varying with index* α_1.

If X_1 *and* X_2 *are independent non-negative regularly varying random variables with index* $\alpha_1 \in \mathbb{R}_+$ *and* $\alpha_2 \in \mathbb{R}_+$, *respectively, then*

$$\mathbb{P}(X_1 + X_2 > x) \sim \begin{cases} \mathbb{P}(X_1 > x) & \text{if } \alpha_1 < \alpha_2, \\ \mathbb{P}(X_1 > x) + \mathbb{P}(X_2 > x) & \text{if } \alpha_1 = \alpha_2, \\ \mathbb{P}(X_2 > x) & \text{if } \alpha_1 > \alpha_2, \end{cases}$$

as $x \to \infty$, *and hence* $X_1 + X_2$ *is regularly varying with index* $\min\{\alpha_1, \alpha_2\}$.

The statements of Lemma E.10 follow, e.g., from parts 1 and 3 of Lemma B.6.1 of Buraczewski et al. [6] and Lemma E.9 together with the fact that the sum of two slowly varying functions is slowly varying.

E.11 Theorem. (Karamata's theorem) *Let* $U : \mathbb{R}_{++} \to \mathbb{R}_{++}$ *be a locally integrable function such that it is integrable on intervals including* 0 *as well.*

(i) If U *is regularly varying (at infinity) with index* $-\alpha \in [-1, \infty)$, *then* $\mathbb{R}_{++} \ni x \mapsto \int_0^x U(t)\,\mathrm{d}t$ *is regularly varying (at infinity) with index* $1 - \alpha$, *and*

$$\lim_{x\to\infty} \frac{xU(x)}{\int_0^x U(t)\,\mathrm{d}t} = 1 - \alpha.$$

(ii) If U *is regularly varying (at infinity) with index* $-\alpha \in (-\infty, -1)$, *then* $\mathbb{R}_{++} \ni x \mapsto \int_x^\infty U(t)\,\mathrm{d}t$ *is regularly varying (at infinity) with index* $1 - \alpha$, *and*

$$\lim_{x\to\infty} \frac{xU(x)}{\int_x^\infty U(t)\,\mathrm{d}t} = -1 + \alpha.$$

For Theorem E.11, see, e.g., Resnick [23, Theorem 2.1].

E.12 Lemma. (Potter's bounds) *If* $U : \mathbb{R}_{++} \to \mathbb{R}_{++}$ *is a regularly varying function (at infinity) with index* $-\alpha \in \mathbb{R}$, *then for every* $\delta \in \mathbb{R}_{++}$, *there exists* $x_0 \in \mathbb{R}_+$ *such that*

$$(1-\delta)q^{-\alpha-\delta} < \frac{U(qx)}{U(x)} < (1+\delta)q^{-\alpha+\delta}, \qquad x \in [x_0, \infty), \quad q \in [1, \infty).$$

For Lemma E.12, see, e.g., Resnick [23, Proposition 2.6].

Finally, we recall a result on the tail behaviour of regularly varying random sums.

E.13 Proposition. *Let* τ *be a non-negative integer-valued random variable and let* $\{\zeta, \zeta_i : i \in \mathbb{N}\}$ *be independent and identically distributed non-negative random variables, independent of* τ*, such that* τ *is regularly varying with index* $\beta \in \mathbb{R}_+$ *and* $\mathbb{E}(\zeta) \in \mathbb{R}_{++}$*. In case of* $\beta \in [1, \infty)$*, assume additionally that there exists* $r \in (\beta, \infty)$ *with* $\mathbb{E}(\zeta^r) < \infty$*. Then we have*

$$\mathbb{P}\left(\sum_{i=1}^{\tau} \zeta_i > x\right) \sim \mathbb{P}\left(\tau > \frac{x}{\mathbb{E}(\zeta)}\right) \sim (\mathbb{E}(\zeta))^{\beta}\, \mathbb{P}(\tau > x) \qquad \text{as} \quad x \to \infty,$$

and hence $\sum_{i=1}^{\tau} \zeta_i$ *is also regularly varying with index* β*.*

For a proof of Proposition E.13, see, e.g., Barczy et al. [2, Proposition F.3].

F Large deviations

We recall a result about large deviations for sums of non-negative independent and identically distributed regularly varying random variables, see, Tang and Yan [26, part (ii) of Theorem 1]. We use it in the second proof of Theorem G.1 in case of $\alpha \in (1, 2)$. Here we present a complete proof as well, since the one in Tang and Yan [26, part (ii) of Theorem 1] contains a gap.

F.1 Theorem. (Large deviations) *If* $(\eta_j)_{j \in \mathbb{N}}$ *are independent, identically distributed non-negative regularly varying random variables with index* $\alpha \in (1, 2)$*, then for each* $\gamma \in (\mathbb{E}(\eta_1), \infty)$*, there exists a constant* $C \in \mathbb{R}_{++}$ *such that*

$$\mathbb{P}(\eta_1 + \cdots + \eta_n > y) \leqslant Cn\, \mathbb{P}(\eta_1 > y)$$

for all $n \in \mathbb{N}$ *and* $y \in [\gamma n, \infty)$*.*

Proof. We will follow the proof of part (ii) of Theorem 1 in Tang and Yan [26]. Let $q \in (0, 1)$ and

$$\widetilde{\eta}_j := \eta_j \mathbb{1}_{\{\eta_j \leqslant qy\}}, \quad j \in \mathbb{N}, \qquad \widetilde{S}_j := \sum_{i=1}^{j} \widetilde{\eta}_i, \quad j \in \mathbb{N}.$$

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

$$\begin{aligned}
&\mathbb{P}(\eta_1 + \cdots + \eta_n > y) \\
&\quad = \mathbb{P}(\eta_1 + \cdots + \eta_n > y, \max_{1 \leqslant j \leqslant n} \eta_j > qy) + \mathbb{P}(\eta_1 + \cdots + \eta_n > y, \max_{1 \leqslant j \leqslant n} \eta_j \leqslant qy) \\
&\quad \leqslant \mathbb{P}(\max_{1 \leqslant j \leqslant n} \eta_j > qy) + \mathbb{P}(\widetilde{\eta}_1 + \cdots + \widetilde{\eta}_n > y, \max_{1 \leqslant j \leqslant n} \eta_j \leqslant qy) \\
&\quad \leqslant \mathbb{P}(\max_{1 \leqslant j \leqslant n} \eta_j > qy) + \mathbb{P}(\widetilde{S}_n > y) \qquad \text{(F.1)} \\
&\quad \leqslant \sum_{j=1}^{n} \mathbb{P}(\eta_j > qy) + \mathbb{P}(\widetilde{S}_n > y) \\
&\quad = n\, \mathbb{P}(\eta_1 > qy) + \mathbb{P}(\widetilde{S}_n > y), \qquad y \in \mathbb{R}_+.
\end{aligned}$$

Here

$$\mathbb{P}(\eta_1 > qy) = \frac{\mathbb{P}(\eta_1 > qy)}{\mathbb{P}(\eta_1 > y)} \cdot \mathbb{P}(\eta_1 > y), \qquad y \in \mathbb{R}_{++},$$

and since $\lim_{y\to\infty} \frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)} = q^{-\alpha}$, there exists an $y_* \in \mathbb{R}_{++}$ such that $\frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)} \leqslant 2q^{-\alpha}$ for all $y \geqslant y_*$. Now we check that $\frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)}$ is bounded on the interval $[0, y_*]$. Since $\lim_{y\to 0} \frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)} = 1$, there exists an $y_1 \in \mathbb{R}_{++}$ such that $y_1 < y_*$ and $\frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)} \leqslant 2$ on the interval $[0, y_1]$. On the interval $[y_1, y_*]$ the quantity $\frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)}$ can be bounded from above by $\frac{\mathbb{P}(\eta_1>qy_1)}{\mathbb{P}(\eta_1>y_*)}$. Hence the function $\mathbb{R}_+ \ni y \mapsto \frac{\mathbb{P}(\eta_1>qy)}{\mathbb{P}(\eta_1>y)}$ is bounded, and consequently, there exists a constant $C_1(q) \in \mathbb{R}_{++}$ (depending possibly on the distribution of η_1 as well) such that

$$\text{(F.2)} \qquad n\,\mathbb{P}(\eta_1 > qy) \leqslant C_1(q) n\,\mathbb{P}(\eta_1 > y), \qquad y \in \mathbb{R}_{++}, \quad n \in \mathbb{N}.$$

Let $a(n,y) := \max\{-\log(n\,\mathbb{P}(\eta_1 > y), 1\}$, $n \in \mathbb{N}$, $y \in \mathbb{R}_{++}$. Then $a(n,y)$ tends to ∞ uniformly for $y \geqslant \gamma n$ as $n \to \infty$, i.e., $\lim_{n\to\infty} \inf_{y\geqslant\gamma n} a(n,y) = \infty$, since, by Lemma E.4,

(F.3)

$$n\,\mathbb{P}(\eta_1 > y) \leqslant n\,\mathbb{P}(\eta_1 > \gamma n) = n(\gamma n)^{-\alpha} L_{\eta_1}(\gamma n) = \gamma^{-\alpha} n^{1-\alpha} L_{\eta_1}(n) \frac{L_{\eta_1}(\gamma n)}{L_{\eta_1}(n)} \to \gamma^{-\alpha} \cdot 0 \cdot 1 = 0$$

as $n \to \infty$, where $L_{\eta_1}(y) := y^\alpha\,\mathbb{P}(\eta_1 > y)$, $y \in \mathbb{R}_{++}$, is a slowly varying (at infinity) function. For any $y \in \mathbb{R}_{++}$, $h \in \mathbb{R}_{++}$ and $n \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{\mathbb{P}(\widetilde{S}_n > y)}{n\,\mathbb{P}(\eta_1 > y)} &\leqslant \frac{\mathrm{e}^{-hy}\,\mathbb{E}(\mathrm{e}^{h\widetilde{S}_n})}{n\,\mathbb{P}(\eta_1 > y)} = \frac{\mathrm{e}^{-hy}(\mathbb{E}(\mathrm{e}^{h\widetilde{\eta}_1}))^n}{n\,\mathbb{P}(\eta_1 > y)} \\
&= \frac{\mathrm{e}^{-hy}\left(\int_0^{qy} \mathrm{e}^{ht}\ \ F_{\eta_1}(\mathrm{d}t)\right)^n}{n\,\mathbb{P}(\eta_1 > y)} = \frac{\mathrm{e}^{-hy}\left(\int_0^{qy} (\mathrm{e}^{ht} - 1)\ \ F_{\eta_1}(\mathrm{d}t) + 1\right)^n}{n\,\mathbb{P}(\eta_1 > y)} \\
&\leqslant \frac{\mathrm{e}^{-hy} \exp\left\{n \int_0^{qy} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t)\right\}}{\mathrm{e}^{-a(n,y)}},
\end{aligned}$$

where the last step follows from $(1 + y)^n \leqslant \mathrm{e}^{ny}$, $y \in \mathbb{R}_+$, $n \in \mathbb{N}$, and from $a(n,y) \geqslant -\log(n\,\mathbb{P}(\eta_1 > y))$, yielding $\mathrm{e}^{-a(n,y)} \leqslant n\,\mathbb{P}(\eta_1 > y)$. Using that $a(n,y) \geqslant 1$, $n \in \mathbb{N}$, $y \in \mathbb{R}_{++}$, let us consider the decomposition

$$\int_0^{qy} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t) = \int_0^{\frac{qy}{a(n,y)}} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t) + \int_{\frac{qy}{a(n,y)}}^{qy} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t) =: I_1 + I_2.$$

Using the inequality $\mathrm{e}^y - 1 \leqslant y\mathrm{e}^y$, $y \in \mathbb{R}_+$, we have

$$\begin{aligned}
I_1 = \int_0^{\frac{qy}{a(n,y)}} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t) &\leqslant \int_0^{\frac{qy}{a(n,y)}} ht\mathrm{e}^{ht}\, F_{\eta_1}(\mathrm{d}t) \\
&\leqslant \mathrm{e}^{\frac{hqy}{a(n,y)}} \int_0^{\frac{qy}{a(n,y)}} ht\, F_{\eta_1}(\mathrm{d}t) \leqslant h\mathrm{e}^{\frac{hqy}{a(n,y)}}\,\mathbb{E}(\eta_1).
\end{aligned}$$

Now we turn to treat I_2. Applying Lemma E.6, for all $\varrho > \alpha$, there exist $y_0 \in \mathbb{R}_{++}$ and $B \in \mathbb{R}_{++}$ (possibly depending on ϱ and on the distribution of η_1) such that

$$\frac{\mathbb{P}\left(\eta_1 > \frac{qy}{a(n,y)}\right)}{\mathbb{P}(\eta_1 > y)} \leqslant B\left(\frac{q}{a(n,y)}\right)^{-\varrho} \qquad \text{whenever } \ y \geqslant \frac{qy}{a(n,y)} \geqslant y_0.$$

The aim of the following discussion is to show that for each $n \in \mathbb{N}$, there exists $\widetilde{y}_0(n) \in \mathbb{R}_{++}$ such that $y \geqslant \frac{qy}{a(n,y)} \geqslant y_0$ holds for all $y \geqslant \widetilde{y}_0(n)$. For each $n \in \mathbb{N}$, the first inequality holds for sufficiently large y, since $\lim_{y\to\infty} a(n,y) = \infty$. Moreover, for each $n \in \mathbb{N}$, the second inequality holds for sufficiently large y, since $\lim_{y\to\infty} \frac{a(n,y)}{y} = 0$. Indeed, for each $n \in \mathbb{N}$ we have $a(n,y) = -\log(n\,\mathbb{P}(\eta_1 > y))$ for sufficiently large y, and hence

$$\frac{a(n,y)}{y} = \frac{-\log(n\,\mathbb{P}(\eta_1 > y))}{y} = \frac{-\log(ny^{-\alpha}L_{\eta_1}(y))}{y} = \frac{-\log(n) + \alpha\log(y) - \log(L_{\eta_1}(y))}{y}.$$

By Lemma E.4, for any $\delta \in \mathbb{R}_{++}$, we have $y^{-\delta} \leqslant L_{\eta_1}(y) \leqslant y^{\delta}$ for sufficiently large y. Taking logarithm, dividing by y, and using that $\lim_{y\to\infty} \frac{\log(y)}{y} = 0$, one concludes $\lim_{y\to\infty} \frac{a(n,y)}{y} = 0$.

Set

$$h := h(n,y,K) := \frac{a(n,y) - K\varrho\log(a(n,y))}{Kqy},$$

where $\varrho > \alpha$ and $K > 1$ will be chosen later. We show that there exists $N_1 \in \mathbb{N}$ such that $h > 0$ and $a(n,y) > 1$ for all $y \geqslant \gamma n$ and $n \geqslant N_1$. Since $\lim_{x\to\infty} \frac{\log(x)}{x} = 0$, there exists $M > 0$ such that $\frac{\log(x)}{x} < \frac{1}{K\varrho}$ for all $x \geqslant M$. Since $\lim_{n\to\infty} \inf_{y\geqslant\gamma n} a(n,y) = \infty$, there exists $n_0(M) \in \mathbb{N}$ such that $a(n,y) \geqslant M$ for all $y \geqslant \gamma n$ with $n \geqslant n_0(M)$. Hence $\frac{\log(a(n,y))}{a(n,y)} < \frac{1}{K\varrho}$ for all $y \geqslant \gamma n$ with $n \geqslant n_0(M)$, as desired. Hence for all $\varrho > \alpha$ and $y \geqslant \max\{\widetilde{y}_0(n), \gamma n\}$ with $n \geqslant N_1$, we have

$$\begin{aligned}
I_2 &= \int_{\frac{qy}{a(n,y)}}^{qy} (\mathrm{e}^{ht} - 1)\, F_{\eta_1}(\mathrm{d}t) \leqslant \mathrm{e}^{hqy}\,\mathbb{P}\left(\eta_1 > \frac{qy}{a(n,y)}\right) \\
&\leqslant \exp\left\{\frac{a(n,y) - K\varrho\log(a(n,y))}{K}\right\} B\left(\frac{q}{a(n,y)}\right)^{-\varrho} \mathbb{P}(\eta_1 > y) \\
&= Bq^{-\varrho}\mathrm{e}^{\frac{a(n,y)}{K}}\,\mathbb{P}(\eta_1 > y) = Bq^{-\varrho}(n\,\mathbb{P}(\eta_1 > y))^{-\frac{1}{K}}\,\mathbb{P}(\eta_1 > y),
\end{aligned}$$

where we used that $1 < a(n,y) = -\log(n\,\mathbb{P}(\eta_1 > y))$. Putting together the bounds for I_1 and I_2 and using that $\frac{hqy}{a(n,y)} \leqslant \frac{1}{K}$ for $y \geqslant \gamma n$ with $n \geqslant N_1$, we obtain that

$$\text{(F.4)} \qquad \frac{\mathbb{P}(\widetilde{S}_n > y)}{n\,\mathbb{P}(\eta_1 > y)} \leqslant \exp\left\{nh\,\mathbb{E}(\eta_1)\mathrm{e}^{\frac{1}{K}} + Bq^{-\varrho}(n\,\mathbb{P}(\eta_1 > y))^{1-\frac{1}{K}} - hy + a(n,y)\right\}$$

for $y \geqslant \max\{\widetilde{y}_0(n), \gamma n\}$ with $n \geqslant N_1$. Noting that $n\,\mathbb{P}(\eta_1 > y) \to 0$ uniformly for $y \geqslant \gamma n$ as $n \to \infty$ (see, (F.3)), we obtain that there exists $C_2 \in \mathbb{R}_{++}$ such that the right-hand side of (F.4) can be bounded by

$$
\begin{aligned}
&C_2 \exp\Big\{ nh\,\mathbb{E}(\eta_1)\mathrm{e}^{\frac{1}{K}} - hy + a(n,y) \Big\} \\
&= C_2 \exp\left\{ hy\left(\frac{\mathrm{e}^{\frac{1}{K}} n\,\mathbb{E}(\eta_1)}{y} - 1 \right) + a(n,y) \right\} \\
&\leqslant C_2 \exp\left\{ \frac{a(n,y) - K\varrho \log(a(n,y))}{Kq} \left(\frac{\mathrm{e}^{\frac{1}{K}}\,\mathbb{E}(\eta_1)}{\gamma} - 1 \right) + a(n,y) \right\}
\end{aligned}
$$

for all $\varrho > \alpha$, sufficiently large $n \in \mathbb{N}$ (greater than N_1) and $y \geqslant \max\{\widetilde{y}_0(n), \gamma n\}$. Since $\gamma > \mathbb{E}(\eta_1)$, we can choose $K > 1$ sufficiently large such that $\frac{1}{\gamma}\mathrm{e}^{\frac{1}{K}}\,\mathbb{E}(\eta_1) < 1$, then we choose $q > 0$ sufficiently small such that

$$
\frac{1}{Kq}\left(\frac{\mathrm{e}^{\frac{1}{K}}\,\mathbb{E}(\eta_1)}{\gamma} - 1 \right) + 2 < 0,
$$

i.e., $q < \frac{1}{2K}(1 - \frac{1}{\gamma}\mathrm{e}^{\frac{1}{K}}\,\mathbb{E}(\eta_1))$. Then we have

$$
\begin{aligned}
\frac{\mathbb{P}(\widetilde{S}_n > y)}{n\,\mathbb{P}(\eta_1 > y)} &\leqslant C_2 \exp\left\{ (a(n,y) - K\varrho \log(a(n,y)))(-2) + a(n,y) \right\} \\
&= C_2 \exp\left\{ 2K\varrho \log(a(n,y)) - a(n,y) \right\}
\end{aligned}
$$

for all $\varrho > \alpha$, sufficiently large $n \in \mathbb{N}$ (greater than N_1) and $y \geqslant \max\{\widetilde{y}_0(n), \gamma n\}$, where we used that $a(n,y) - K\varrho\log(a(n,y)) > 0$ for $y \geqslant \gamma n$ with $n \geqslant N_1$. Here $C_2 \exp\{2K\varrho\log(a(n,y)) - a(n,y)\}$ tends to 0 uniformly for $y \geqslant \gamma n$ as $n \to \infty$, i.e.,

$$
\sup_{y \geqslant \gamma n} \exp\left\{ 2K\varrho \log(a(n,y)) - a(n,y) \right\} = \exp\left\{ \sup_{y \geqslant \gamma n} (2K\varrho \log(a(n,y)) - a(n,y)) \right\} \to 0
$$

as $n \to \infty$. Indeed, this will be a consequence of $\sup_{y \geqslant \gamma n}(2K\varrho\log(a(n,y)) - a(n,y)) \to -\infty$ as $n \to \infty$. We have

$$
\sup_{y \geqslant \gamma n} (2K\varrho \log(a(n,y)) - a(n,y)) \leqslant S_1(n) + S_2(n), \tag{F.5}
$$

where

$$
\begin{aligned}
S_1(n) &:= \sup_{y \geqslant \gamma n} \left(2K\varrho \log(a(n,y)) - \frac{1}{2}a(n,y) \right), \\
S_2(n) &:= \sup_{y \geqslant \gamma n} \left(-\frac{1}{2}a(n,y) \right) = -\frac{1}{2} \inf_{y \geqslant \gamma n} a(n,y) \to -\infty \qquad \text{as } n \to \infty.
\end{aligned}
$$

Moreover, $\lim_{x \to \infty} \frac{\log(x)}{x} = 0$ implies that there exists $\widetilde{M} > 0$ such that $\frac{\log(x)}{x} < \frac{1}{4K\varrho}$ for all $x \geqslant \widetilde{M}$. Since $\lim_{n \to \infty} \inf_{y \geqslant \gamma n} a(n,y) = \infty$, there exists $n_0(\widetilde{M}) \in \mathbb{N}$ such that $a(n,y) \geqslant \widetilde{M}$ for all $y \geqslant \gamma n$ with $n \geqslant n_0(\widetilde{M})$. Hence $\frac{\log(a(n,y))}{a(n,y)} < \frac{1}{4K\varrho}$ for all $y \geqslant \gamma n$ with $n \geqslant n_0(\widetilde{M})$, thus $2K\varrho\log(a(n,y)) < \frac{1}{2}a(n,y)$. Consequently, we obtain $S_1(n) \leqslant 0$ for all $n \geqslant n_0(\widetilde{M})$, and hence, by (F.5), we conclude $\sup_{y\geqslant\gamma n}(2K\varrho\log(a(n,y)) - a(n,y)) \to -\infty$ as $n \to \infty$, as desired. So we have

$$\lim_{n\to\infty} \sup_{y\geqslant\gamma n} \frac{\mathbb{P}(\widetilde{S}_n > y)}{n\,\mathbb{P}(\eta_1 > y)} = 0.$$

Consequently, there exists an $N \in \mathbb{N}$ such that

$$\sup_{n\geqslant N,\, y\geqslant\gamma n} \frac{\mathbb{P}(\widetilde{S}_n > y)}{n\,\mathbb{P}(\eta_1 > y)} < \infty.$$

This, together with (F.1) and (F.2) yield that

$$\sup_{n\geqslant N,\, y\geqslant\gamma n} \frac{\mathbb{P}(S_n > y)}{n\,\mathbb{P}(\eta_1 > y)} < \infty. \tag{F.6}$$

Finally, using the convolution property (see, Lemma E.10),

$$\sup_{1\leqslant n\leqslant N,\, y\geqslant\gamma n} \frac{\mathbb{P}(S_n > y)}{n\,\mathbb{P}(\eta_1 > y)} \leqslant \sum_{n=1}^{N} \sup_{y\geqslant\gamma n} \frac{\mathbb{P}(S_n > y)}{n\,\mathbb{P}(\eta_1 > y)} < \infty. \tag{F.7}$$

The desired statement readily follows from (F.6) and (F.7). □

G Tail behavior of second-order Galton–Watson processes (without immigration) having regularly varying initial distributions

G.1 Proposition. *Let $(X_n)_{n\geqslant -1}$ be a second-order Galton–Watson process (without immigration) such that X_0 and X_{-1} are independent, X_0 is regularly varying with index $\beta_0 \in \mathbb{R}_+$, X_{-1} is regularly varying with index $\beta_{-1} \in \mathbb{R}_+$ and $m_\xi, m_\eta \in \mathbb{R}_{++}$. In case of* $\max\{\beta_0, \beta_{-1}\} \in [1, \infty)$, *assume additionally that there exists $r \in (\max\{\beta_0, \beta_{-1}\}, \infty)$ with* $\mathbb{E}(\xi^r) < \infty$ *and* $\mathbb{E}(\eta^r) < \infty$. *Then for each $n \in \mathbb{N}$,*

$$\mathbb{P}(X_n > x) \sim \begin{cases} m_n^{\beta_0}\,\mathbb{P}(X_0 > x) & \text{if } 0 \leqslant \beta_0 < \beta_{-1}, \\ m_n^{\beta_0}\,\mathbb{P}(X_0 > x) + m_{n-1}^{\beta_{-1}} m_\eta^{\beta_{-1}}\,\mathbb{P}(X_{-1} > x) & \text{if } \beta_0 = \beta_{-1}, \\ m_{n-1}^{\beta_{-1}} m_\eta^{\beta_{-1}}\,\mathbb{P}(X_{-1} > x) & \text{if } \beta_{-1} < \beta_0 \end{cases}$$

as $x \to \infty$, where m_i, $i \in \mathbb{Z}_+$, are given in Theorem 2.1 and hence, X_n is regularly varying with index $\min\{\beta_0, \beta_{-1}\}$ *for each $n \in \mathbb{N}$.*

First proof of Proposition G.1. Let us fix $n \in \mathbb{N}$. In view of the additive property (A.4), the independence of X_0 and X_{-1}, and the convolution property of regularly varying distributions described in Lemma E.10, it is sufficient to prove

$$\mathbb{P}\left(\sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} > x\right) \sim m_n^{\beta_0}\,\mathbb{P}(X_0 > x), \qquad \mathbb{P}\left(\sum_{j=1}^{X_{-1}} \zeta_{j,-1}^{(n)} > x\right) \sim m_{n-1}^{\beta_{-1}} m_\eta^{\beta_{-1}}\,\mathbb{P}(X_{-1} > x) \tag{G.1}$$

as $x \to \infty$. These relations follow from Proposition E.13, since $\mathbb{E}(\zeta_{1,0}^{(n)}) = m_n \in \mathbb{R}_{++}$ and $\mathbb{E}(\zeta_{1,-1}^{(n)}) = m_{n-1} m_\eta \in \mathbb{R}_{++}$, $n \in \mathbb{N}$, by (B.4). $\Box$

Second proof of Proposition G.1. Let us fix $n \in \mathbb{N}$. In view of the additive property (A.4), the independence of X_0 and X_{-1}, and the convolution property of regularly varying distributions described in Lemma E.10, it is sufficient to prove (G.1). We show only the first relation in (G.1), since the second one can be proven in the same way. Note that $\mathbb{E}(\zeta_{1,0}^{(n)}) = m_n$ by (B.4). First, we prove

$$\liminf_{x\to\infty} \frac{\mathbb{P}\big(\sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} > x\big)}{\mathbb{P}(X_0 > x)} \geqslant m_n^{\beta_0}. \tag{G.2}$$

Let $q \in (0,1)$ be arbitrary. For sufficiently large $x \in \mathbb{R}_{++}$, we have $\lfloor (1+q)x/m_n \rfloor \geqslant 1$, since $m_n > 0$. Using that for each $i \in \mathbb{N}$, $\zeta_{i,0}^{(n)}$ is non-negative, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} > x\right) \geqslant \sum_{k=\lfloor (1+q)x/m_n \rfloor}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} > x\right) \mathbb{P}(X_0 = k) \\
&\geqslant \mathbb{P}\left(\sum_{i=1}^{\lfloor (1+q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) \sum_{k=\lfloor (1+q)x/m_n \rfloor}^{\infty} \mathbb{P}(X_0 = k) \\
&= \mathbb{P}\left(\frac{1}{\lfloor (1+q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1+q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{x}{\lfloor (1+q)x/m_n \rfloor}\right) \mathbb{P}(X_0 \geqslant \lfloor (1+q)x/m_n \rfloor) \\
&\geqslant \mathbb{P}\left(\frac{1}{\lfloor (1+q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1+q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{x}{\lfloor (1+q)x/m_n \rfloor}\right) \mathbb{P}(X_0 > (1+q)x/m_n)
\end{aligned}$$

for sufficiently large $x \in \mathbb{R}_{++}$. For sufficiently large $x \in \mathbb{R}_{++}$, we have $\frac{x}{\lfloor (1+q)x/m_n \rfloor} \leqslant \frac{m_n}{1+(q/2)}$, since $\frac{x}{\lfloor (1+q)x/m_n \rfloor} \to \frac{m_n}{1+q}$ as $x \to \infty$ and $\frac{m_n}{1+q} < \frac{m_n}{1+(q/2)}$. Hence, for sufficiently large $x \in \mathbb{R}_{++}$, we have

$$\mathbb{P}\left(\sum_{i=1}^{X_0} \zeta_{i,0}^{(n)} > x\right) \geqslant \mathbb{P}\left(\frac{1}{\lfloor (1+q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1+q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{m_n}{1+(q/2)}\right) \mathbb{P}\left(X_0 > \frac{(1+q)x}{m_n}\right).$$

We have

$$\frac{1}{N} \sum_{i=1}^{N} \zeta_{i,0}^{(n)} \xrightarrow{\text{a.s.}} \mathbb{E}(\zeta_{1,0}^{(n)}) = m_n \qquad \text{as } N \to \infty \tag{G.3}$$

by the strong law of large numbers, hence $\frac{m_n}{1+(q/2)} < m_n$ yields

$$\mathbb{P}\left(\frac{1}{\lfloor (1+q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1+q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{m_n}{1+(q/2)}\right) \to 1 \qquad \text{as } x \to \infty.$$

Thus, using that X_0 is regularly varying with index β_0, we have

$$\mathbb{P}\left(\frac{1}{\lfloor(1+q)x/m_n\rfloor}\sum_{i=1}^{\lfloor(1+q)x/m_n\rfloor}\zeta_{i,0}^{(n)}>\frac{m_n}{1+(q/2)}\right)\mathbb{P}\left(X_0>\frac{(1+q)x}{m_n}\right)\sim\mathbb{P}\left(X_0>\frac{(1+q)x}{m_n}\right)$$
$$\sim\left(\frac{m_n}{1+q}\right)^{\beta_0}\mathbb{P}(X_0>x)$$

as $x\to\infty$. Consequently,

$$\liminf_{x\to\infty}\frac{\mathbb{P}\big(\sum_{i=1}^{X_0}\zeta_{i,0}^{(n)}>x\big)}{\mathbb{P}(X_0>x)}\geqslant\left(\frac{m_n}{1+q}\right)^{\beta_0},\qquad q\in(0,1),$$

and, by $q\downarrow 0$, we conclude (G.2).

Next, we prove

$$\limsup_{x\to\infty}\frac{\mathbb{P}\big(\sum_{i=1}^{X_0}\zeta_{i,0}^{(n)}>x\big)}{\mathbb{P}(X_0>x)}\leqslant m_n^{\beta_0}. \tag{G.4}$$

Let $q\in(0,1)$ be arbitrary. For sufficiently large $x\in\mathbb{R}_{++}$, we have $\lfloor(1-q)x/m_n\rfloor\geqslant 1$, and hence

$$\mathbb{P}\left(\sum_{i=1}^{X_0}\zeta_{i,0}^{(n)}>x\right)\leqslant\mathbb{P}(X_0>\lfloor(1-q)x/m_n\rfloor)+\sum_{k=1}^{\lfloor(1-q)x/m_n\rfloor}\mathbb{P}\left(\sum_{i=1}^{k}\zeta_{i,0}^{(n)}>x\right)\mathbb{P}(X_0=k)$$
$$=\mathbb{P}\left(X_0>\frac{(1-q)x}{m_n}\right)+\sum_{k=1}^{\lfloor(1-q)x/m_n\rfloor}\mathbb{P}\left(\sum_{i=1}^{k}\zeta_{i,0}^{(n)}>x\right)\mathbb{P}(X_0=k).$$

Since X_0 is regularly varying with index β_0, we have

$$\mathbb{P}\left(X_0>\frac{(1-q)x}{m_n}\right)\sim\left(\frac{m_n}{1-q}\right)^{\beta_0}\mathbb{P}(X_0>x)\qquad\text{as } x\to\infty,$$

hence, by taking the limit $q\downarrow 0$, we get (G.4) provided we check

$$p(x,q):=\sum_{k=1}^{\lfloor(1-q)x/m_n\rfloor}\mathbb{P}\left(\sum_{i=1}^{k}\zeta_{i,0}^{(n)}>x\right)\mathbb{P}(X_0=k)=\mathrm{o}\left(\mathbb{P}\left(X_0>x\right)\right)\qquad\text{as } x\to\infty \tag{G.5}$$

for all sufficiently small $q\in(0,1)$. (In fact, it will turn out that (G.5) holds for any $q\in(0,1)$.)

First, we consider the case $\beta_0\in(0,1)$. Let $0<\delta<(1-q)/m_n$. Then for sufficiently large $x\in\mathbb{R}_{++}$, we have $\lfloor\delta x\rfloor<\lfloor(1-q)x/m_n\rfloor$, and then

$$p(x,q)=\sum_{k=1}^{\lfloor\delta x\rfloor}\mathbb{P}\left(\sum_{i=1}^{k}\zeta_{i,0}^{(n)}>x\right)\mathbb{P}(X_0=k)+\sum_{k=\lfloor\delta x\rfloor+1}^{\lfloor(1-q)x/m_n\rfloor}\mathbb{P}\left(\sum_{i=1}^{k}\zeta_{i,0}^{(n)}>x\right)\mathbb{P}(X_0=k)$$
$$=: p_1(x,\delta)+p_2(x,\delta,q).$$

At first, we show that $p_2(x,\delta,q) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$ for all $0 < \delta < (1-q)/m_n$. Here, using that $\zeta_{i,0}^{(n)}$ is non-negative for each $i \in \mathbb{N}$, we obtain

$$\begin{aligned}
p_2(x,\delta,q) &\leqslant \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) \sum_{k=\lfloor \delta x \rfloor + 1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}(X_0 = k) \\
&\leqslant \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) \mathbb{P}(X_0 > \lfloor \delta x \rfloor) = \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) \mathbb{P}(X_0 > \delta x).
\end{aligned}$$

For sufficiently large $x \in \mathbb{R}_{++}$, we have $\frac{x}{\lfloor (1-q)x/m_n \rfloor} \geqslant \frac{m_n}{1-(q/2)}$, since $\frac{x}{\lfloor (1-q)x/m_n \rfloor} \to \frac{m_n}{1-q}$ as $x \to \infty$ and $\frac{m_n}{1-q} > \frac{m_n}{1-(q/2)}$. Hence, for sufficiently large $x \in \mathbb{R}_{++}$, we have

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) &= \mathbb{P}\left(\frac{1}{\lfloor (1-q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{x}{\lfloor (1-q)x/m_n \rfloor}\right) \\
&\leqslant \mathbb{P}\left(\frac{1}{\lfloor (1-q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{m_n}{1-(q/2)}\right).
\end{aligned}$$

Again by the strong law of large numbers (see (G.3)), $\frac{m_n}{1-(q/2)} > m_n$ yields

$$\mathbb{P}\left(\frac{1}{\lfloor (1-q)x/m_n \rfloor} \sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > \frac{m_n}{1-(q/2)}\right) \to 0 \qquad \text{as} \quad x \to \infty,$$

hence we obtain

$$\text{(G.6)} \qquad \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x\right) \to 0 \qquad \text{as} \quad x \to \infty.$$

Using that X_0 is regularly varying with index β_0, we have $\mathbb{P}(X_0 > \delta x) \sim \delta^{-\beta_0}\,\mathbb{P}(X_0 > x)$ as $x \to \infty$, hence $p_2(x,\delta,q) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$ for all $0 < \delta < (1-q)/m_n$ and $q \in (0,1)$. Now we turn to prove

$$\limsup_{\delta \downarrow 0} \limsup_{x \to \infty} \frac{p_1(x,\delta)}{\mathbb{P}(X_0 > x)} = 0.$$

By Markov's inequality,

$$\mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} > x\right) \leqslant \frac{1}{x} \sum_{i=1}^{k} \mathbb{E}(\zeta_{i,0}^{(n)}) = \frac{m_n k}{x}$$

for all $k \in \mathbb{N}$ and $x \in \mathbb{R}_{++}$, and hence

$$\begin{aligned}
p_1(x,\delta) &\leqslant \frac{m_n}{x} \sum_{k=0}^{\lfloor \delta x \rfloor} k\,\mathbb{P}(X_0 = k) = \frac{m_n}{x}\,\mathbb{E}(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor \delta x \rfloor\}}) = \frac{m_n}{x} \int_0^\infty \mathbb{P}(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor \delta x \rfloor\}} > t)\,\mathrm{d}t \\
&= \frac{m_n}{x} \int_0^{\lfloor \delta x \rfloor} \mathbb{P}(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor \delta x \rfloor\}} > t)\,\mathrm{d}t \leqslant \frac{m_n}{x} \int_0^{\lfloor \delta x \rfloor} \mathbb{P}(X_0 > t)\,\mathrm{d}t.
\end{aligned}$$

Since $\mathbb{R}_+ \ni x \mapsto \mathbb{P}(X_0 > x)$ is locally integrable (due to the fact that it is bounded), it is integrable on intervals including 0 as well, and since it is regularly varying (at infinity) with index $-\beta_0$, by Karamata's theorem (see Theorem E.11),

$$\lim_{x\to\infty} \frac{x\,\mathbb{P}(X_0 > x)}{\int_0^x \mathbb{P}(X_0 > t)\,\mathrm{d}t} = 1 - \beta_0,$$

and hence

$$\int_0^{\lfloor \delta x \rfloor} \mathbb{P}(X_0 > t)\,\mathrm{d}t \sim \frac{1}{1-\beta_0} \lfloor \delta x \rfloor\, \mathbb{P}(X_0 > \lfloor \delta x \rfloor) = \frac{1}{1-\beta_0} \lfloor \delta x \rfloor\, \mathbb{P}(X_0 > \delta x)$$

as $x \to \infty$. Then using that $\mathbb{P}(X_0 > \delta x) \sim \delta^{-\beta_0}\, \mathbb{P}(X_0 > x)$ as $x \to \infty$, we have

$$\frac{p_1(x,\delta)}{\mathbb{P}(X_0 > x)} \leqslant \frac{m_n}{x} \frac{\int_0^{\lfloor \delta x \rfloor} \mathbb{P}(X_0 > t)\,\mathrm{d}t}{\mathbb{P}(X_0 > x)} \sim \frac{m_n}{1-\beta_0} \delta \frac{\mathbb{P}(X_0 > \delta x)}{\mathbb{P}(X_0 > x)} \sim \frac{m_n}{1-\beta_0} \delta^{1-\beta_0}$$

as $x \to \infty$. Consequently,

$$\limsup_{x\to\infty} \frac{p_1(x,\delta)}{\mathbb{P}(X_0 > x)} \leqslant \frac{m_n}{1-\beta_0} \delta^{1-\beta_0} \qquad \text{for all} \quad 0 < \delta < \frac{1-q}{m_n},$$

and hence $\limsup_{\delta\downarrow 0} \limsup_{x\to\infty} \frac{p_1(x,\delta)}{\mathbb{P}(X_0>x)} \leqslant \lim_{\delta\downarrow 0} \frac{m_n}{1-\beta_0} \delta^{1-\beta_0} = 0$. Combining the parts we get $p(x,q) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$ for any $q \in (0,1)$, as desired.

Next, we consider the case $\beta_0 \in (1,2)$. Using Lemma E.7, we check that there exists a non-negative random variable $\widetilde{\zeta}^{(n)}$ having the following properties:

- $\widetilde{\zeta}^{(n)}$ is regularly varying with index β_0,
- $\mathbb{P}(\zeta_{1,0}^{(n)} > x) \leqslant \mathbb{P}(\widetilde{\zeta}^{(n)} > x)$, $x \in \mathbb{R}_+$,
- $\mathbb{P}(\widetilde{\zeta}^{(n)} > x) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$,
- $\mathbb{E}(\zeta_{1,0}^{(n)}) \leqslant \mathbb{E}(\widetilde{\zeta}^{(n)}) < \infty$.

By Lemma C.1, $\mathbb{E}((\zeta_{1,0}^{(n)})^r) < \infty$, and hence, by Lemma E.8, $\mathbb{P}(\zeta_{1,0}^{(n)} > x) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$. Thus, by Lemma E.7, there exists a monotone increasing, right-continuous, slowly varying (at infinity) function $L_{\widetilde{\zeta}^{(n)}}$ such that $L_{\widetilde{\zeta}^{(n)}}(x) \geqslant 1$, $x \in \mathbb{R}_+$, $\lim_{x\to\infty} L_{\widetilde{\zeta}^{(n)}}(x) = \infty$ and $\lim_{x\to\infty} L_{\widetilde{\zeta}^{(n)}}(x) \frac{\mathbb{P}(\zeta_{1,0}^{(n)}>x)}{\mathbb{P}(X_0>x)} = 0$. Hence, using also that $\mathbb{P}(X_0 \geqslant x) \leqslant 1$, $x \in \mathbb{R}_+$, there exists $x' \in \mathbb{R}_+$ such that $L_{\widetilde{\zeta}^{(n)}}(x) \frac{\mathbb{P}(\zeta_{1,0}^{(n)}>x)}{\mathbb{P}(X_0>x)} \leqslant 1$ and $\frac{\mathbb{P}(X_0>x)}{L_{\widetilde{\zeta}^{(n)}}(x)} \leqslant 1$ hold for all $x \geqslant x'$. Let $\widetilde{\zeta}^{(n)}$ be a random variable such that

$$\mathbb{P}(\widetilde{\zeta}^{(n)} > x) := \begin{cases} 1 & \text{if } x \leqslant x', \\ \frac{\mathbb{P}(X_0>x)}{L_{\widetilde{\zeta}^{(n)}}(x)} & \text{if } x > x'. \end{cases}$$

Such a non-negative random variable exists, since $\mathbb{R}_{++} \ni x \mapsto \frac{\mathbb{P}(X_0>x)}{L_{\widetilde{\zeta}^{(n)}}(x)}$ is monotone decreasing, converges to 0 as $x \to \infty$ and right-continuous. For all $q \in \mathbb{R}_{++}$,

$$\lim_{x\to\infty} \frac{\mathbb{P}(\widetilde{\zeta}^{(n)} > qx)}{\mathbb{P}(\widetilde{\zeta}^{(n)} > x)} = \lim_{x\to\infty} \frac{L_{\widetilde{\zeta}^{(n)}}(x)}{L_{\widetilde{\zeta}^{(n)}}(qx)} \frac{\mathbb{P}(X_0 > qx)}{\mathbb{P}(X_0 > x)} = 1 \cdot q^{-\beta_0} = q^{-\beta_0},$$

yielding that $\widetilde{\zeta}^{(n)}$ is regularly varying with index β_0. For $x \leqslant x'$, we have $\mathbb{P}(\zeta_{1,0}^{(n)} > x) \leqslant 1 = \mathbb{P}(\widetilde{\zeta}^{(n)} > x)$. For $x > x'$, we have

$$\mathbb{P}(\zeta_{1,0}^{(n)} > x) = L_{\widetilde{\zeta}^{(n)}}(x) \frac{\mathbb{P}(\zeta_{1,0}^{(n)} > x)}{\mathbb{P}(X_0 > x)} \mathbb{P}(\widetilde{\zeta}^{(n)} > x) \leqslant \mathbb{P}(\widetilde{\zeta}^{(n)} > x).$$

Further,

$$\lim_{x\to\infty} \frac{\mathbb{P}(\widetilde{\zeta}^{(n)} > x)}{\mathbb{P}(X_0 > x)} = \lim_{x\to\infty} \frac{\mathbb{P}(X_0 > x)}{L_{\widetilde{\zeta}^{(n)}}(x)\, \mathbb{P}(X_0 > x)} = 0,$$

since $\lim_{x\to\infty} L_{\widetilde{\zeta}^{(n)}}(x) = \infty$. Since $\mathbb{P}(\zeta_{1,0}^{(n)} > x) \leqslant \mathbb{P}(\widetilde{\zeta}^{(n)} > x)$, $x \in \mathbb{R}_+$, we have

$$\mathbb{E}(\zeta_{1,0}^{(n)}) = \int_0^\infty \mathbb{P}(\zeta_{1,0}^{(n)} > x)\,\mathrm{d}x \leqslant \int_0^\infty \mathbb{P}(\widetilde{\zeta}^{(n)} > x)\,\mathrm{d}x = \mathbb{E}(\widetilde{\zeta}^{(n)}),$$

and since $\widetilde{\zeta}^{(n)}$ is regularly varying with index $\beta_0 \in (1,2)$, we have $\mathbb{E}(\widetilde{\zeta}^{(n)}) < \infty$.

Let $(\widetilde{\zeta}_j^{(n)})_{j\in\mathbb{N}}$ be a sequence of independent identically distributed random variables with common distribution as that of $\widetilde{\zeta}^{(n)}$. By some properties of first order stochastic dominance (see, e.g., Shaked and Shanthikumar [25, part (b) of Theorem 1.A.3 and Theorem 1.A.4]), we have

$$\text{(G.7)} \qquad \mathbb{P}\left(\sum_{i=1}^k \zeta_{i,0}^{(n)} > x\right) \leqslant \mathbb{P}\left(\sum_{i=1}^k \widetilde{\zeta}_i^{(n)} > x\right)$$

for all $x \in \mathbb{R}_+$ and $k \in \mathbb{N}$. Put $\widetilde{m}_n := \mathbb{E}(\widetilde{\zeta}^{(n)})$. Let us consider the decomposition

$$\begin{aligned} p(x,q) = &\sum_{k=1}^{\lfloor (1-q)x/\widetilde{m}_n \rfloor} \mathbb{P}\left(\sum_{i=1}^k \zeta_{i,0}^{(n)} > x\right) \mathbb{P}(X_0 = k) \\ &+ \sum_{k=\lfloor (1-q)x/\widetilde{m}_n \rfloor + 1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}\left(\sum_{i=1}^k \zeta_{i,0}^{(n)} > x\right) \mathbb{P}(X_0 = k) =: p_1(x,q) + p_2(x,q), \qquad x \in \mathbb{R}_+. \end{aligned}$$

Here $m_n \leqslant \widetilde{m}_n$, and hence $\lfloor (1-q)x/\widetilde{m}_n \rfloor \leqslant \lfloor (1-q)x/m_n \rfloor$, $x \in \mathbb{R}_+$, $q \in (0,1)$. Applying Theorem F.1 with $\gamma := \frac{\widetilde{m}_n}{1-q} > \widetilde{m}_n$, we conclude the existence of a constant $C(q,n) \in \mathbb{R}_{++}$ (not depending on k and x, but on q and n) such that

$$\text{(G.8)} \qquad \mathbb{P}\left(\sum_{i=1}^k \widetilde{\zeta}_i^{(n)} > x\right) \leqslant C(q,n) k\, \mathbb{P}(\widetilde{\zeta}^{(n)} > x) \qquad \text{for all } x \geqslant \gamma k, \ k \in \mathbb{N}.$$

Using (G.7) and (G.8), we obtain

$$p_1(x,q) \leqslant \sum_{k=1}^{\lfloor (1-q)x/\widetilde{m}_n \rfloor} \mathbb{P}\left(\sum_{i=1}^{k} \widetilde{\zeta}_i^{(n)} > x \right) \mathbb{P}(X_0 = k)$$

$$\leqslant C(q,n) \sum_{k=1}^{\lfloor (1-q)x/\widetilde{m}_n \rfloor} k\, \mathbb{P}(\widetilde{\zeta}^{(n)} > x)\, \mathbb{P}(X_0 = k) \leqslant C(q,n)\, \mathbb{E}(X_0)\, \mathbb{P}(\widetilde{\zeta}^{(n)} > x), \qquad x \in \mathbb{R}_+.$$

Hence for each $q \in (0,1)$,

$$\limsup_{x\to\infty} \frac{p_1(x,q)}{\mathbb{P}(X_0 > x)} \leqslant C(q,n)\, \mathbb{E}(X_0) \limsup_{x\to\infty} \frac{\mathbb{P}(\widetilde{\zeta}^{(n)} > x)}{\mathbb{P}(X_0 > x)} = 0,$$

where the last step follows by the corresponding property of $\widetilde{\zeta}^{(n)}$. Moreover,

$$p_2(x,q) \leqslant \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x \right) \sum_{k=\lfloor (1-q)x/\widetilde{m}_n \rfloor + 1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}(X_0 = k)$$

$$\leqslant \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x \right) \mathbb{P}(X_0 > \lfloor (1-q)x/\widetilde{m}_n \rfloor)$$

$$= \mathbb{P}\left(\sum_{i=1}^{\lfloor (1-q)x/m_n \rfloor} \zeta_{i,0}^{(n)} > x \right) \mathbb{P}\left(X_0 > \frac{(1-q)x}{\widetilde{m}_n} \right).$$

Since X_0 is regularly varying with index β_0, we have

$$\lim_{x\to\infty} \frac{\mathbb{P}\left(X_0 > \frac{(1-q)x}{\widetilde{m}_n} \right)}{\mathbb{P}(X_0 > x)} = \left(\frac{\widetilde{m}_n}{1-q} \right)^{\beta_0},$$

hence, for each $q \in (0,1)$, applying (G.6), we conclude

$$\limsup_{x\to\infty} \frac{p_2(x,q)}{\mathbb{P}(X_0 > x)} \leqslant 0 \cdot \left(\frac{\widetilde{m}_n}{1-q} \right)^{\beta_0} = 0.$$

Finally, we turn to the case $\beta_0 = 1$. For each $q \in (0,1)$, we have

$$p(x,q) = \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} > x \right) \mathbb{P}(X_0 = k)$$

$$= \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} - k m_n > x - k m_n \right) \mathbb{P}(X_0 = k).$$

Let $r' \in (1,2]$. According to Lemma 2.1 in Robert and Segers [24] with $\gamma = \frac{m_n q}{1-q}$, there exist positive numbers v and $C = C(v,q,n)$ such that for all $x \in \mathbb{R}_+$ and $k \in \mathbb{N}$ with

$k \leqslant \lfloor (1-q)x/m_n \rfloor$,

$$\mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} - km_n > x - km_n\right) \leqslant k\,\mathbb{P}\big(\zeta_{1,0}^{(n)} - m_n > v(x-km_n)\big) + \frac{C}{(x-km_n)^{r'}},$$

since $x - km_n \geqslant \gamma k$ for all $x \in \mathbb{R}_+$ and $k \in \mathbb{N}$ with $k \leqslant \lfloor (1-q)x/m_n \rfloor$. Consequently,

$$\begin{aligned}
p(x,q) &= \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}\left(\sum_{i=1}^{k} \zeta_{i,0}^{(n)} - km_n > x - km_n\right) \mathbb{P}(X_0 = k) \\
&\leqslant \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}(X_0 = k) \left(k\,\mathbb{P}\big(\zeta_{1,0}^{(n)} - m_n > v(x-km_n)\big) + \frac{C}{(x-km_n)^{r'}}\right) \\
&\leqslant \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} \mathbb{P}(X_0 = k) \left(k\,\mathbb{P}\big(\zeta_{1,0}^{(n)} > v(x-km_n)\big) + \frac{C}{(x-km_n)^{r'}}\right) \\
&\leqslant \mathbb{P}(\zeta_{1,0}^{(n)} > qvx) \sum_{k=1}^{\lfloor (1-q)x/m_n \rfloor} k\,\mathbb{P}(X_0 = k) + \frac{C}{(qx)^{r'}} \\
&\leqslant \mathbb{P}(\zeta_{1,0}^{(n)} > qvx)\, \mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right) + \frac{C}{(qx)^{r'}},
\end{aligned}$$

where for the last but one step, we used that $x - km_n \geqslant qx$ for $k \in \{1, \ldots, \lfloor (1-q)x/m_n \rfloor\}$. Since $r' \in (1,2]$, by Lemma E.4, we have $C/(qx)^{r'} = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$, so we only have to work with the first term. If $\mathbb{E}(X_0) < \infty$, then $\mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right) \leqslant \mathbb{E}(X_0) < \infty$ also holds, and

$$\frac{\mathbb{P}(\zeta_{1,0}^{(n)} > qvx)}{\mathbb{P}(X_0 > x)} = \frac{\mathbb{P}(X_0 > qvx)}{\mathbb{P}(X_0 > x)} \cdot \frac{\mathbb{P}(\zeta_{1,0}^{(n)} > qvx)}{\mathbb{P}(X_0 > qvx)} \to 0 \qquad \text{as } x \to \infty,$$

where we used that X_0 is regularly varying with index 1, and that $\mathbb{P}(\zeta_{1,0}^{(n)} > x) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$ also holds (as it was already proved earlier). Now we consider the case $\mathbb{E}(X_0) = \infty$. By Markov's inequality, $\mathbb{P}(\zeta_{1,0}^{(n)} > qvx) \leqslant \mathbb{E}((\zeta_{1,0}^{(n)})^r)/(qvx)^r$ (note that in this case, $\mathbb{E}((\zeta_{1,0}^{(n)})^r)$ exists, see Lemma C.1), and using the fact that $\limsup_{x\to\infty} \frac{\mathbb{E}(X_0 \mathbb{1}_{\{X_0 \leqslant x\}})}{x^s\,\mathbb{P}(X_0 > x)} = 0$ for some $1 < s < r$ (see the remark after Theorem 3.2 in Robert and Segers [24]), we have

$$\begin{aligned}
&\frac{\mathbb{P}(\zeta_{1,0}^{(n)} > qvx)\, \mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right)}{\mathbb{P}(X_0 > x)} \leqslant \frac{\mathbb{E}((\zeta_{1,0}^{(n)})^r)\, \mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right)}{(qvx)^r\, \mathbb{P}(X_0 > x)} \\
&= \frac{\mathbb{E}((\zeta_{1,0}^{(n)})^r)}{(qv)^r} \cdot \frac{\mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right)}{x^s\, \mathbb{P}(X_0 > x)} \cdot \frac{1}{x^{r-s}} \\
&= \frac{\mathbb{E}((\zeta_{1,0}^{(n)})^r)}{(qv)^r} \cdot \frac{\lfloor (1-q)x/m_n \rfloor^s}{x^s} \cdot \frac{\mathbb{P}\left(X_0 > \lfloor (1-q)x/m_n \rfloor\right)}{\mathbb{P}(X_0 > x)} \\
&\quad \times \frac{\mathbb{E}\left(X_0 \mathbb{1}_{\{X_0 \leqslant \lfloor (1-q)x/m_n \rfloor\}}\right)}{\lfloor (1-q)x/m_n \rfloor^s\, \mathbb{P}(X_0 > \lfloor (1-q)x/m_n \rfloor)} \cdot \frac{1}{x^{r-s}} \to 0 \qquad \text{as } x \to \infty.
\end{aligned}$$

Putting parts together, we have $p(x,q) = \mathrm{o}(\mathbb{P}(X_0 > x))$ as $x \to \infty$, as desired. $\square$

G.2 Remark. For a corresponding result for (first-order) Galton–Watson processes (without immigration), see Barczy et al. [2, Proposition 2.2]. A formal application of Proposition G.1 also gives this result, namely, for each $n \in \mathbb{N}$, we have $\mathbb{P}(X_n > x) \sim m_\xi^{n\beta_0}\,\mathbb{P}(X_0 > x)$ as $x \to \infty$. In case of $m_\xi = 0$ and $m_\eta \in \mathbb{R}_{++}$, Proposition 2.2 in Barczy et al. [2] gives that $\mathbb{P}(X_n > x) \sim m_\eta^{\frac{n}{2}\beta_0}\,\mathbb{P}(X_0 > x)$ as $x \to \infty$ if $n \in \mathbb{N}$ is even, and $\mathbb{P}(X_n > x) \sim m_\eta^{\frac{n+1}{2}\beta_{-1}}\,\mathbb{P}(X_{-1} > x)$ as $x \to \infty$ if $n \in \mathbb{N}$ is odd. $\square$

Acknowledgements

We would like to thank the referee and Prof. Yuliya Mishura, Co-editor-in-chief, for their comments that helped us to improve the paper.

References

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

[2] Barczy, M., Bősze, Zs. and Pap, G. (2018). Regularly varying non-stationary Galton–Watson processes with immigration. *Statistics & Probability Letters* **140** 106–114.

[3] Basrak, B., Kulik, R. and Palmowski, Z. (2013). Heavy-tailed branching process with immigration. *Stochastic Models* **29(4)** 413–434.

[4] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular variation.* Cambridge University Press, Cambridge.

[5] Bősze, Zs. and Pap, G. (2019). Regularly varying nonstationary second-order Galton–Watson processes with immigration. *Stochastic Models* **35(2)** 132–147.

[6] Buraczewski, D., Damek, E. and Mikosch, T. (2016). *Stochastic models with power-law tails. The equation $X = AX + B$.* Springer, Cham.

[7] Dénes, A., Kevei, P., Nishiura, H. and Röst, G. (2013). Risk of infectious disease outbreaks by imported cases with application to the European Football Championship 2012. *International Journal of Stochastic Analysis* **Art. ID 576381**, 9 pp.

[8] Du, J. G. and Li, Y. (1991). The integer valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis* **12(2)** 129–142.

[9] Faÿ, G., González-Arévalo, B., Mikosch, T. and Samorodnitsky, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Systems* **54** 121–140.

[10] Heyer, H. (2010). *Structural aspects in the theory of probability, 2nd ed.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

[11] Horn, R. A. and Johnson, Ch. R. (2013). *Matrix Analysis, 2nd ed.* Cambridge University Press, Cambridge.

[12] Hult, H. and Samorodnitsky, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli* **14(3)** 838–864.

[13] Ispány, M. and Pap, G. (2014). Asymptotic behavior of critical primitive multi-type branching processes with immigration. *Stochastic Analysis and Applications* **32(5)** 727–741.

[14] Kaplan, N. (1974). The supercritical p-dimensional Galton–Watson process with immigration. *Mathematical Biosciences* **22** 1–18.

[15] Kashikar, A. S. (2017). Estimation of growth rate in second order branching process. *Journal of Statistical Planning and Inference* **191** 1–12.

[16] Kashikar, A. S. and Deshmukh, S. R. (2015). Probabilistic properties of second order branching process. *Annals of the Institute of Statistical Mathematics* **67** 557–572.

[17] Kashikar, A. S. and Deshmukh, S. R. (2016). Estimation in second order branching processes with application to swine flu data. *Communications in Statistics – Theory and Methods* **45(4)** 1031–1046.

[18] Kimmel, M. and Axelrod, D. E. (2015). *Branching Processes in Biology, 2nd ed.* Springer, New York.

[19] Latour, A. (1997). The multivariate GINAR(p) process. *Advances in Applied Probability* **29(1)** 228–248.

[20] Pénisson, S. (2014). Estimation of the infection parameter of an epidemic modeled by a branching process. *Electronic Journal of Statistics* **8(2)** 2158–2187.

[21] Pénisson, S. and Jacob, Ch. (2012). Stochastic methodology for the study of an epidemic decay phase, based on a branching model. *International Journal of Stochastic Analysis* **Art. ID 598701**, 32 pp.

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

[23] Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer.

[24] Robert, C. Y. and Segers, J. (2008) Tails of random sums of a heavy-tailed number of light-tailed terms. *Insurance: Mathematics and Economics* **43** 85–92.

[25] Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders.* Springer Series in Statistics. Springer, New York.

[26] Tang, Q. and Yan, J. (2002). A sharp inequality for the tail probabilities of sums of i.i.d. r.v.'s with dominatedly varying tails. *Science in China (Series A)* **45(8)** 1006-1011.