Size-biased branching population measures and the multi-type x log x condition

Peter Olofsson
Mathematics Department
Trinity University
San Antonio, Texas

June 14, 2008

Abstract

We investigate the $x \log x$ condition for a general (Crump–Mode–Jagers) multi-type branching process with arbitrary type space by constructing a size-biased population measure that relates to the ordinary population measure via the intrinsic martingale W_t. Sufficiency of the $x \log x$ condition for a non-degenerate limit of W_t is proved and conditions for necessity are investigated.

GENERAL BRANCHING PROCESS; XLOGX CONDITION; IMMIGRATION; SIZE-BIASED MEASURE

AMS 2000 SUBJECT CLASSIFICATION: PRIMARY 60J80

1 Introduction

The $x \log x$ condition is a fundamental concept in the theory of branching processes, being the necessary and sufficient condition for a supercritical branching process to grow as its mean. In a Galton-Watson process with offspring mean $m = E[X] > 1$, let Z_n be the number of individuals in the nth generation and let $W_n = Z_n/m^n$. Then W_n is a nonnegative martingale and hence $W_n \to W$ for some random variable W. The *Kesten-Stigum Theorem* is

Theorem 1.1 *If* $E[X\log^+ X] < \infty$ *then* $E[W] = 1$*; if* $E[X\log^+ X] = \infty$ *then* $W = 0$ *a.s.*

It can further be shown that $P(W = 0)$ must either be 0 or equal the extinction probability and hence $E[X\log^+ X] < \infty$ implies that $W > 0$ exactly on the set of nonextinction (see *e.g.* Athreya and Ney (1972)).

The analogue for general single-type branching processes appears in Jagers and Nerman (1984) and a partial result (establishing sufficiency) for general multi-type branching processes in Jagers (1989). Lyons, Pemantle and Peres (1995) give a slick proof of the Kesten-Stigum theorem based on comparisons between the Galton-Watson measure and another measure, the *size-biased* Galton-Watson measure, on the space of progeny trees. In Olofsson (1998), these ideas were further developed to analyze general single-type branching processes and the current paper proceeds to general multi-type branching processes with an arbitrary type space. In addition to providing a new proof of a known result, the ideas of size-biased processes now provide tools to further analyze conditions for necessity of the $x\log x$ condition.

A crucial concept for the Lyons-Pemantle-Peres (LPP) proof is that of *size bias*. If the offspring distribution is $\{p_0, p_1, ...\}$ and has $m = E[X]$, the size-biased offspring distribution is defined as

$$\widetilde{p}_k = \frac{kp_k}{m}$$

and a size-biased Galton-Watson tree is constructed in the following way: Let $\widetilde{X}$ denote a random variable that has the size-biased offspring distribution and let the ancestor v_0 have a number $\widetilde{X}_0$ of children. Pick one of these at random, call her v_1, give her a number $\widetilde{X}_1$ of children, and give her siblings ordinary Galton-Watson descendant trees. Pick one of v_1's children at random, give her a number $\widetilde{X}_2$ of children, give her sisters ordinary Galton-Watson descendant trees, and so on and so forth. With P_n denoting the ordinary Galton-Watson measure restricted to the n first generations, $\widetilde{P}_n$ denoting the measure that arises from the above construction, and $W_n = Z_n/m^n$, the relation

$$d\widetilde{P}_n = W_n dP_n, \tag{1.1}$$

holds. Hence, it is the fundamental martingale W_n that size-biases the Galton-Watson process. The construction of $\widetilde{P}$ can also be viewed as describing a Galton-Watson process with immigration where the immigrants are the siblings of the individuals on the path $(v_0, v_1, ...)$. Thus, the measure P is the ordinary Galton-Watson measure and the size-biased measure $\widetilde{P}$ is the measure of a Galton-Watson process with immigration where the i.i.d. immigration group sizes are distributed as $\widetilde{X} - 1$. The relation between P and $\widetilde{P}$ on the space of family trees can now be explored using results for processes with immigration which is the key to the proof.

The idea of using size-bias as such in branching processes appeared before LPP. One early example is Joffe and Waugh (1982), where size-biased Galton-Watson processes show up in the study of ancestral trees of randomly sampled individuals. This approach was further explored by Olofsson and Shaw (2002) with a view toward biological applications. The only approach that is anywhere near that of LPP, however, seems to be Waymire and Williams (1996), developed simultaneously with, and independently of LPP. Later applications and extensions of the powerful LPP method include Athreya (2000) and Biggins and Kyprianou (2004).

To make this paper self-contained, we give a short review of general multi-type branching processes and their $x \log x$ condition in the next section. As in the Galton-Watson case, branching processes with immigration are crucial in the proof; for that purpose we briefly discuss processes with immigration in Section 3 following Olofsson (1996). The size-biased measure on the space of population trees and its relation to the ordinary branching measure is investigated in Section 4 and in Section 5, sufficiency of the $x \log x$ condition is proved. Finally, in Section 6, we discuss various conditions for necessity.

2 The x log x Condition for General Branching Processes

In a general branching process, individuals are identified by descent. The ancestor is denoted by 0, the children of the ancestor by $1, 2, ...$ and so on, so that the individual $x = (x_1, ..., x_n)$ is the x_nth child of the x_{n-1}th child of...of the x_1th child of the ancestor. The set of all individuals can thus be described as

$$I = \bigcup_{n=0}^{\infty} N^n$$

At birth, each individual is assigned a *type* s chosen from the *type space* S, equipped with some appropriate σ-algebra $\mathcal{S}$. The type s determines a probability measure $P_s(\cdot)$, the *life law*, on the *life space* Ω, equipped with some appropriate σ-algebra $\mathcal{F}$. The information provided by a life $\omega \in \Omega$ may differ from one application to another but it must at least give the *reproduction process* ξ on $S \times R_+$. This process gives the sequence of birth times and types of the children of an individual. More precisely, let $(\tau(k), \sigma(k))$ be random variables on Ω denoting the birth time (age of the mother) and type of the kth child, respectively, and define

$$\xi(A \times [0,t]) = \#\{k : \sigma(k) \in A, \tau(k) \leq t\}$$

for $A \in \mathcal{S}$ and $t \geq 0$. We let $\tau(k) \equiv \infty$ if fewer than k children are born. The *population space* is defined as Ω^I, an outcome of which gives the lives of all individuals. The set of probability kernels $\{P_s(\cdot), s \in S\}$ defines a probability measure on Ω^I, the *population measure* P_s, where the ancestor's type is s. By using projections, we can view individual lives, types, etc, as random variables on Ω^I rather than Ω; if $\omega^I \in \Omega^I$, we for example define $\omega_x = U_x(\omega^I)$, the life of the individual x. With each individual $x \in I$, we associate its type σ_x, its birth time τ_x, and its life ω_x where σ_x is inherited from the mother (a function of the mother's life) and ω_x is chosen according to the probability distribution $P_{\sigma_x}(\cdot)$ on $(\Omega, \mathcal{F})$. The birth time τ_x is defined recursively by letting the ancestor be born at time $\tau_0 = 0$ and if x is the kth child of its mother y (i.e., $x = yk$), we let $\tau_x = \tau_y + \tau(k)$. Note that τ_x and τ_y refer to absolute time whereas $\tau(k)$ refers to the mother's age at x's birth.

An important entity is the *reproduction kernel*, defined by

$$\mu(s, dr \times dt) = E_s[\xi(dr \times dt)]$$

the expectation of $\xi(dr \times dt)$ when the mother is of type s. This kernel plays the role of $m = E[X]$ in the simple Galton-Watson process and determines the growth rate of the process as $e^{\alpha t}$ where α is called the Malthusian parameter. For the rest of this section, we leave out technical details and assumptions, instead focusing on the main definitions and results. The details can be found in Jagers (1989, 1992) and we simply refer to a process that satisfies all the conditions needed as a *Malthusian* process.

Given α, we define the Laplace transform of μ as

$$\widehat{\mu}(s,dr) = \int_0^\infty e^{-\alpha t}\mu(s,dr\times dt)$$

and under certain conditions this kernel has eigenmeasure π and eigenfunction h given by

$$\pi(dr) = \int_S \widehat{\mu}(s,dr)\pi(ds)$$

$$h(s) = \int_S h(r)\widehat{\mu}(s,dr) \tag{2.1}$$

where both π and $hd\pi$ can be normed to probability measures. The measure π is called the *stable type distribution* and $h(s)$ is called the *reproductive value* of an individual of type s. The interpretation of π and $hd\pi$ is that π is the distribution of the type of an individual chosen at random from an old population, and $hd\pi$ is the limiting type distribution backward in the family tree from this individual. The asymptotic age of child-bearing in this backward sense is denoted by β and satisfies

$$\beta = \int_{S\times S\times R_+} te^{-\alpha t}h(r)\mu(s,dr\times dt)\pi(ds) < \infty \tag{2.2}$$

To count, or measure, the population, *random characteristics* are used. A random characteristic is a real-valued process χ where $\chi(a)$ gives the contribution to the population of an individual of age a. Thus χ is a process defined on the life space and by letting χ_x be the characteristic pertaining to the individual x, the χ-counted population is defined as

$$Z_t^\chi = \sum_{x\in I}\chi_x(t-\tau_x)$$

which is the sum of the contributions of all individuals at time t (when the individual x is of age $t-\tau_x$). The simplest example of a random characteristic is $\chi(a) = I_{R_+}(a)$, the indicator for being born, in which case Z_t^χ is simply the total number of individuals born up to time t.

To capture the asymptotics of Z_t^χ, the crucial entity is the *intrinsic martingale* W_t, introduced by Nerman (1981) for single-type processes and generalized to multi-type processes in Jagers (1989). For its definition, denote x's mother my mx and let

$$\mathcal{I}_t = \{x : \tau_{mx} \leq t < \tau_x\} \tag{2.3}$$

the set of individuals whose mothers are born before time t but who themselves are not yet born at time t. Let

$$W_t = \frac{1}{h(\sigma_0)} \sum_{x \in \mathcal{I}_t} e^{-\alpha\tau_x} h(\sigma_x)$$

the individuals in $\mathcal{I}_t$ summed with time- and type-dependent weights, normed by the reproductive value of the ancestor. It can be shown that W_t is a martingale with respect to the σ-algebra $\mathcal{F}_t$ generated by the lives of all individuals born before t, and that $E_s[W_t] = 1$ for all $s \in S$. Hence, W_t plays the role that $W_n = Z_n/m^n$ does in the Galton-Watson process and the limit of Z_t^χ turns out to involve the martingale limit $W = \lim_{t\to\infty} W_t$. The main convergence result is of the form

$$e^{-\alpha t} Z_t^\chi \to \frac{E_\pi[\widehat{\chi}]}{\alpha\beta} h(s) W$$

where s is the type of the ancestor, $E_\pi = \int_S E_s \pi(ds)$, and $\widehat{\chi}$ is the Laplace transform of $\chi(t)$. As in the Galton-Watson case, the question is when the martingale limit W is non-degenerate. As $W_t \to W$ a.s. and $E_s[W_t] = 1$, L^1-convergence is equivalent to $E_s[W] = 1$ (Durrett (2005), p.258). Note that although it is the process Z_t^χ that is of interest and not W_t itself, the asymptotics are determined by W_t, one of many examples of the usefulness of finding an embedded martingale.

We are ready to formulate the general $x \log x$ condition and the main convergence result. For the reproduction process ξ, define the transform

$$\bar{\xi} = \int_{S\times R_+} e^{-\alpha t} h(r) \xi(dr \times dt)$$

which plays the role of X in the Galton-Watson process (in fact, in that case $\bar{\xi} = X/m$) and the $x \log x$ is defined in terms of $\bar{\xi}$. The $x \log x$ condition and convergence result are given in the following theorem from Jagers (1989).

Theorem 2.1 *Consider a general multi-type Malthusian branching process with*

$$E_\pi[\bar{\xi} \log^+ \bar{\xi}] < \infty$$

Then $E_s[W] = 1$ *for* $\pi-$*almost all* s *and*

$$e^{-\alpha t} Z_t^{\chi} \rightarrow \frac{E_\pi[\widehat{\chi}]}{\alpha\beta} h(s) W$$

in $L^1(P_s)$ *for* $\pi-$*almost all* s.

3 Processes with Immigration

Now consider a general branching process where new individuals $v_1, v_2, ...$ immigrate into the population according to some point process $\eta(dr \times dt)$ with points of occurrence and types $(\tau_1, \sigma_1), (\tau_2, \sigma_2), ...$ (for ease of notation, we write τ_k rather than τ_{v_k} etc). The kth immigrant initiates a branching process according to the population measure P_{σ_k}. The immigration process has the transform

$$\bar{\eta} = \int_0^\infty e^{-\alpha t} h(r) \eta(dr \times dt) = \sum_{k=1}^{\infty} e^{-\alpha \tau_k} h(\sigma_k)$$

and it can be shown that the process W_t is now a submartingale rather than a martingale (which is intuitively clear because immigrants are added to the set $\mathcal{I}_t$). The limit of W_t is therefore not automatically finite but needs a condition on the immigration process, established by the following lemma from Olofsson (1996).

Lemma 3.1 *Suppose that* $\bar{\eta} < \infty$. *Then* $W_t \rightarrow W$ *a.s. as* $t \rightarrow \infty$ *where* $W < \infty$ *a.s.*

The a.s. qualification refers to the joint probability measure of the immigration process and the population.

4 The size-biased population measure

Recall that Lyons-Pemantle-Peres's size-biased Galton-Watson measure was constructed from the size-biased offspring distribution. General branching processes require a more general concept of size-bias. In a general process, the offspring random variable X is replaced by the reproduction process ξ, the size of which is properly measured by the transform $\bar{\xi}$ which leads to the following definition.

Definition 4.1 *The size-biased life law $\widetilde{P}_s$ is defined as*

$$\widetilde{P}_s(d\omega) = \frac{\bar{\xi}(\omega)}{h(s)} P_s(d\omega)$$

The following lemma follows immediately from the definition of $\widetilde{P}_s$.

Lemma 4.2 *Let P_s and $\widetilde{P}_s$ be as above and denote the set of realizations of reproduction processes by $(\Gamma, \mathcal{G})$. Then*

(i) For $A \in \mathcal{F}$,

$$\widetilde{P}_s(A) = \frac{E_s[\bar{\xi}; A]}{h(s)}$$

(ii) For every $\mathcal{G}-$measurable function $g : \Gamma \to R$,

$$\widetilde{E}_s[g(\xi)] = \frac{E_s[\bar{\xi} g(\xi)]}{h(s)}$$

In particular, note that $\widetilde{P}_s$ is indeed a probability measure for all $s \in S$ because

$$\widetilde{P}_s(\Omega) = \frac{1}{h(s)} E_s[\bar{\xi}] = 1$$

where $E_s[\bar{\xi}] = h(s)$ follows from (2.1). Also note that a size-biased process always contain points because

$$\widetilde{P}_s(\xi(S \times R_+) = 0) = E_s[\bar{\xi}; \xi(S \times R_+) = 0] = 0$$

To construct the size-biased population measure, consider an outcome $\omega^I \in \Omega^I$ and let ω_t^I denote the set of family trees that coincide with ω^I up to time t. The goal is to have $d\widetilde{P}_s = W_t dP_s$, that is

$$\widetilde{P}_s(d\omega_t^I) = W_t(\omega^I) P_s(d\omega_t^I)$$

Let x be an individual in $\mathcal{I}_t$, defined in (2.3). By the definition of W_t, the desired relation between $\widetilde{P}_s$ and P_s holds if

$$\widetilde{P}_s(d\omega_t^I; x) = \frac{h(\sigma_x)e^{-\alpha\tau_x}}{h(\sigma_0)} P_s(d\omega_t^I)$$

so we define $\widetilde{P}_s$ by this last expression. Note that

$$\begin{aligned}\widetilde{P}_s(\Omega^I) &= \sum_{x\in\mathcal{I}_t} \widetilde{P}_s(\Omega^I; x) \\ &= E_s[W_t] P_s(\Omega^I) = 1\end{aligned}$$

so that $\widetilde{P}_s$ is indeed a probability measure for all $s \in S$.

To further investigate the relation between $\widetilde{P}_s$ and P_s, let $\bar{\xi}_0$ denote $\bar{\xi}(\omega_0)$. The individual $x \in \mathcal{I}_t$ stems from some individual i in the first generation, thus we have $x = iy$ for some y. Let us start by manipulating the regular population measure to obtain

$$\begin{aligned}P_s(d\omega_t^I) &= P_s(d\omega_0) \prod_{j=1}^{\xi_0(t)} P_{\sigma_j}(d\omega_{t-\tau_j}^{(j)}) \\ &= P_s(d\omega_0) P_{\sigma_i}(d\omega_{t-\tau_i}^{(i)}) \prod_{j\neq i} P_{\sigma_j}(d\omega_{t-\tau_j}^{(j)}) \\ &= \bar{\xi}_0 P_s(d\omega_0) \frac{h(\sigma_i)}{\bar{\xi}_0} \frac{1}{h(\sigma_i)} P_{\sigma_i}(d\omega_{t-\tau_i}^{(i)}) \prod_{j\neq i} P_{\sigma_j}(d\omega_{t-\tau_j}^{(j)})\end{aligned}$$

Now multiply by $\dfrac{e^{-\alpha\tau_x} h(\sigma_x)}{h(s)}$ and note that $\tau_x = \tau_i + \tau_y$ to obtain

$$\begin{aligned}\widetilde{P}_s(d\omega_t^I; x) &= \frac{h(\sigma_x)e^{-\alpha\tau_x}}{h(s)} P_s(d\omega_t^I) \\ &= \frac{\bar{\xi}_0}{h(s)} P_s(d\omega_0) \cdot \frac{h(\sigma_i)e^{-\alpha\tau_i}}{\bar{\xi}_0} \cdot \frac{h(\sigma_y)e^{-\alpha\tau_y}}{h(\sigma_i)} \cdot P_{\sigma_i}(d\omega_{t-\tau_i}^{(i)}) \cdot \prod_{j\neq i} P_{\sigma_j}(d\omega_{t-\tau_j}^{(j)})\end{aligned}$$

$$= \quad \frac{\bar{\xi}_0}{h(s)} P_s(d\omega_0) \cdot \frac{h(\sigma_i)e^{-\alpha\tau_i}}{\bar{\xi}_0} \cdot \widetilde{P}_{\sigma_i}(d\omega^{(i)}_{t-\tau_i}) \cdot \prod_{j\neq i} P_{\sigma_j}(d\omega^{(j)}_{t-\tau_j})$$

where we used the fact that the individual x is also the individual y when i is viewed as the ancestor; thus $h(\sigma_x) = h(\sigma_y)$ (formally, recall the projections mentioned in Section 1).

The expression for $\widetilde{P}_s$ suggests the following construction. Start with the ancestor, now called v_0, and choose her life ω_0 according to the size-biased distribution $\frac{\bar{\xi}_0}{h(s)} P_s(d\omega_0)$. Pick one of her children, born in the reproduction process ξ_0, such that the ith child is chosen with probability $\frac{h(\sigma_i)e^{-\alpha\tau_i}}{\bar{\xi}_0}$. Call this child v_1, let her start a population according to the measure $\widetilde{P}_{\sigma_i}$ and give her sisters independent descendant trees, such that sister j follows the law P_{σ_j}. Continue in this way and define the measure $\widetilde{P}_s$ to be the joint distribution of the random tree and the random path $(v_0, v_1, ...)$; then $\widetilde{P}_s$ satisfies the recursive expression above.

The individuals off the path $(v_0, v_1, ...)$ constitute a general branching process with immigration (the immigrants being the children of $v_0, v_1, ...$). To describe the immigration process, let $I_{j,k}$ be the indicator of the event that v_{j-1}'s kth child is *not* chosen to be v_j and denote the kth point in ξ_j by $\tau_k(j)$. The immigration process η is

$$\eta(ds \times dt) = \sum_{j,k} \delta_{\sigma_k(j)}(ds)\delta_{\tau_k(j)}(dt - \tau_j)I_{j,k}$$

which has

$$\bar{\eta} = \sum_{j,k} h(\sigma_k(j))e^{-\alpha\tau_j}e^{-\alpha\tau_k(j)}I_{j,k} \tag{4.1}$$

The sequence of types $\sigma_0, \sigma_1, ...$ of the immigrants $v_1, v_2, ...$ in the construction of the size-biased population measure is of particular interest to our analysis. This sequence is a Markov chain with transition probabilities given by

$$\widetilde{P}(\sigma_1 \in dr|\sigma_0 = s) = \sum_i \widetilde{P}(\sigma_1 \in dr, v_1 = i|\sigma_0 = s)$$

$$= \sum_i \widetilde{E}_s\left[\frac{e^{-\alpha\tau_i}h(\sigma_i)}{\bar{\xi}_0}\delta_{\sigma_i}(dr)\right]$$

$$= \frac{1}{h(s)}\sum_i E_s\left[e^{-\alpha\tau_i}h(\sigma_i)\delta_{\sigma_i}(dr)\right]$$

$$= \frac{1}{h(s)}\int_0^\infty E_s\left[e^{-\alpha t}h(r)\xi(dr \times dt)\right]$$

$$= \frac{h(r)}{h(s)}\widehat{\mu}(s, dr)$$

Now let $\nu(ds) = h(s)\pi(ds)$. As

$$\int_S \widehat{\mu}(s, dr)\pi(ds) = \pi(dr)$$

we get

$$\int_{s\in S} \frac{h(r)}{h(s)}\widehat{\mu}(s, dr)\nu(ds) = \nu(dr)$$

and thus the Markov chain of types has stationary distribution $\nu = hd\pi$.

In a similar fashion, the sequence of types and times, $(\sigma_0, T_0), (\sigma_1, T_1), ...$ is a Markov renewal process with transition kernel given by

$$\widetilde{P}(T_1 \in dt, \sigma_1 \in dr|\sigma_0 = s)$$

$$= \sum_i \widetilde{P}(T_1 \in dt, \sigma_1 \in dr, v_1 = i|\sigma_0 = s)$$

$$= \frac{h(r)}{h(s)}e^{-\alpha t}\mu(s, dr \times dt)$$

and the expected value under the stationary distribution $\nu = hd\pi$ is

$$\widetilde{E}_\nu[T_1] = \int_{S\times[0,\infty)} t\frac{h(r)}{h(s)}e^{-\alpha t}h(s)\mu(s, dr\times dt)\pi(ds)$$
$$= \int_{S\times[0,\infty)} te^{-\alpha t}h(r)\mu(s, dr\times dt)\pi(ds)$$
$$= \beta < \infty$$

by (2.2).

There is an interesting connection between the size-biased measure and the *stable population measure* from Jagers (1992). The latter is an asymptotic probability measure that is centered around a randomly sampled individual as $t \to \infty$. In such a stable population, the randomly sampled individual is born in a point process that has the size-biased distribution, and the asymptotic type distribution as time goes backwards through the individual's line of descent is $hd\pi$. The transition probabilities in this backward chain also involve $\widehat{\mu}(s, dr)$ but have weights that are expressed in terms of π rather than h as we have in the size-biased measure where time goes forward. This relation becomes clearer in a finite-type Galton Watson process where π and h are simply the left and right eigenvectors of the mean reproduction matrix.

5 Sufficiency of the $x \log x$ condition

We are soon ready to prove the general $x \log x$ theorem, the key to which is the relation between $\widetilde{P}_s$ and P_s. Recall that the two are related through W_t which is a martingale under P_s and a submartingale under $\widetilde{P}_s$. The following lemma relates the limiting behavior of W_t under $P_\pi = \int_S P_s \pi(ds)$ to its limiting behavior under $\widetilde{P}_\nu = \int_S \widetilde{P}_s h(s)\pi(ds)$.

Lemma 5.1 *Let* $W = \limsup_t W_t$. *Then*

(i) $\widetilde{P}_\nu(W = \infty) = 0 \quad \Rightarrow \quad E_\pi[W] = 1$

(ii) $\widetilde{P}_\nu(W = \infty) = 1 \quad \Rightarrow \quad E_\pi[W] = 0$

Proof. By Durrett (2005), p.239

$$\widetilde{P}_\nu(A) = E_\nu[W;A] + \widetilde{P}_\nu(A \cap \{W = \infty\})$$

If $\widetilde{P}_\nu(W = \infty) = 0$, we have $\widetilde{P}_\nu(A) = E_\nu[W;A]$ and get

$$\widetilde{P}_\nu(W = 0) = E_\nu[W; W = 0] = 0$$

and hence

$$E_\nu[W] = E_\nu[W : W > 0] = \widetilde{P}_\nu(W > 0) = 1$$

Moreover, as

$$E_\nu[W] = \int_S E_s[W]\nu(ds)$$

and as Fatou's lemma implies that $E_s[W] \leq 1$ for all s, we must have $E_s[W] = 1$ for $\nu-$almost all $s \in S$. As $\pi \ll \nu$, we also get $E_\pi[W] = 1$.

Next, suppose that $\widetilde{P}_\nu(W = \infty) = 1$. As W is an a.s. finite martingale limit under P_s, we have $P_\nu(W = \infty) = \int_S P_s(W = \infty)h(s)\pi(ds) = 0$. Hence, the measures $\widetilde{P}_\nu(\cdot)$ and $E_\nu[W;\ \cdot\ ]$ are mutually singular and as $\widetilde{P}_\nu(W > 0) > 0$, we get

$$E_\nu[W] = E_\nu[W; W > 0] = 0$$

which implies that $E_\pi[W] = 0$ as well, and the proof of the lemma is complete. ■

We now aim to prove that finite $x \log x$ moment under P_π implies that $\bar{\eta} < \infty$, landing us at part (i) of Lemma 5.1. To that end, choose $g(x) = \log^+(x)$ in Lemma 4.2 to obtain

$$\begin{aligned} E_\pi[\bar{\xi}\log^+\bar{\xi}] &= \int_S E_s[\bar{\xi}\log^+\bar{\xi}]\pi(ds) \\ &= \int_S h(s)\widetilde{E}_s[\log^+\bar{\xi}]\pi(ds) \\ &= \widetilde{E}_\nu[\log^+\bar{\xi}] \end{aligned}$$

where, we recall, $\nu = hd\pi$ is the stationary distribution for the type sequence $\sigma_0, \sigma_1, ...$ The following lemma brings us one step closer to the proof.

Lemma 5.2 *Consider a general branching process with immigration process* η *as above. If* $\widetilde{E}_\nu[\log^+ \bar{\xi}] < \infty$*, then* $W = \lim_{t\to\infty} W_t$ *exists and is finite* $\widetilde{P}_\nu-$*a.s.*

Proof. First note that

$$\widetilde{E}_\nu[\log^+ \bar{\xi}] < \infty \Rightarrow \sum_n \widetilde{P}_\nu(\log^+ \bar{\xi} > cn) < \infty \text{ for all } c > 0$$

Now consider the sequence $\bar{\xi}_0, \bar{\xi}_1, ...$, which is stationary under $\widetilde{P}_\nu$ which gives

$$\sum_n \widetilde{P}_\nu(\log^+ \bar{\xi}_n > cn) < \infty$$

which implies that, for all $c > 0$,

$$\widetilde{P}_\nu(\log^+ \bar{\xi}_n > cn \text{ i.o.}) = 0$$

which, by (4.1) gives

$$\begin{aligned} \bar{\eta} &\leq \sum_{j,k} h(\sigma_k(j)) e^{-\alpha\tau_j} e^{-\alpha\tau_k(j)} \\ &= \sum_{j=1}^{\infty} e^{-\alpha\tau_j} \bar{\xi}_j < \infty \ \widetilde{P}_\nu - \text{a.s.} \end{aligned}$$

as the τ_j are sums of the T_k which, being the regeneration times in a Markov renewal process, obey the law of large numbers (Alsmeyer (1994)), so that $\tau_j \sim \beta j$ as $j \to \infty$. By Lemma 3.1 we conclude that $\lim_t W_t$ exists and is finite $\widetilde{P}_\nu-$a.s. ∎

Proof of Theorem 2.1. $E_\pi[\bar{\xi}\log^+ \bar{\xi}] = \widetilde{E}_\nu[\log^+ \bar{\xi}] < \infty$ implies $\widetilde{P}_\nu(W = \infty) = 0$ and Lemma 5.1 gives $E_\nu[W] = 1$. Moreover, as

$$E_\nu[W] = \int_S E_s[W]\nu(ds)$$

and as Fatou's lemma implies that $E_s[W] \leq 1$ for all s, we must have $E_s[W] = 1$ for $\nu-$almost all $s \in S$. As $\pi \ll \nu$, the proof is complete. ∎

6 Necessity of the $x \log x$ condition

For single-type processes, the condition of having finite $x \log x$ moment is both sufficient and necessary. Using the size-bias method, this can be established by using the first and second Borel–Cantelli lemmas, respectively. However, in the multi-type setting, the main result in Jagers (1989) establishes only sufficiency and as pointed out in Athreya (2000), additional conditions are typically needed in the general multi-type setting. The method of size-biased branching processes provides a way of investigating conditions for necessity, and although the second Borel–Cantelli lemma can not be used due to dependence, more general versions can be employed. This section is explorative in nature and does not provide any definite solutions but outlines two different approaches through the conditional Borel–Cantelli lemma and the Kochen–Stone lemma, respectively. Below, by "the $x \log x$ condition," we mean $E_\pi[\bar{\xi} \log^+ \bar{\xi}] < \infty$.

The conditional Borel–Cantelli lemma states that if $\mathcal{F}_n$ is a filtration and A_n a sequence of events with $A_n \in \mathcal{F}_n$, then

$$\{A_n \text{ i.o.}\} = \left\{ \sum_{n=1}^{\infty} P(A_n | \mathcal{F}_{n-1}) = \infty \right\}$$

see Durrett (2005). For us, $A_n = \{\log^+ \bar{\xi}_n > cn\}$, the σ-field $\mathcal{F}_{n-1}$ gives the type of the nth individual, and we get

$$P(\log^+ \bar{\xi}_n > n | \mathcal{F}_{n-1}) = P_{\sigma_n}(\log^+ \bar{\xi} > cn)$$

and the question becomes under which conditions

$$\sum_{n=1}^{\infty} \widetilde{P}_{\sigma_n}(\log^+ \bar{\xi} > cn) = \infty \quad \widetilde{P}_\nu - \text{a.s.}$$

given that

$$\sum_{n=1}^{\infty} \widetilde{P}_\nu(\log^+ \bar{\xi} > cn) = \infty$$

One more step is necessary in order to invoke part (ii) of Lemma 5.1, namely to argue that $\log^+ \bar{\xi}_n > cn$ i.o implies that $W = \infty$ $\widetilde{P}_\nu$–a.s. Consider W_{τ_n}, the value of W_t at the time of the arrival of the nth immigrant. As all the children of this immigrant belongs to $\mathcal{I}_{\tau_n}$ and the kth child is born at time $\tau_n + \tau(k)$ and has type $\sigma(k)$ where $\tau(k)$ and $\sigma(k)$ are the points in the reproduction process ξ_n of the nth immigrant, we get

$$\begin{aligned} W_{\tau_n} &= \frac{1}{h(\sigma_0)} \sum_{x \in \mathcal{I}_{\tau_n}} h(\sigma_x) e^{-\alpha \tau_x} \\ &\geq \frac{e^{-\tau_n}}{h(\sigma_0)} \sum_{k=1}^{\infty} h(\sigma(k)) e^{-\alpha \tau(k)} I_{j,k} \\ &\geq \frac{e^{-\tau_n}}{h(\sigma_0)} (\bar{\xi}_n - C) \end{aligned}$$

where $C = \sup_s h(s) < \infty$ (Jagers (1989)). Hence, if $\log^+ \bar{\xi}_n > cn$ i.o, we have $W = \limsup_t W_t = \infty$ $\widetilde{P}_\nu-$a.s.

A potential problem is that a random variable can have infinite expectation under $\widetilde{E}_\nu$ but finite expectations under $\widetilde{E}_s$ for all $s \in S$. The following condition for $x \log x$ necessity precludes this possibility.

Proposition 6.1 *If the Markov chain of types has one positive recurrent state s such that $E_s[\bar{\xi} \log^+ \xi] = \infty$ and $\sigma_k = s$ for some k, then the $x \log x$ condition is necessary.*

Proof. The result follows from the following observation regarding infinite series. Let $a_n \geq 0$ be a decreasing sequence of real numbers such that $\sum_n a_n = \infty$, let $X_1, X_2, ...$ be i.i.d. nonnegative random variables with finite mean μ, and let $T_n = X_1 + X_2 + ... + X_n$. Then

$$\sum_{n=1}^{\infty} a_{T_n} = \infty \quad \text{a.s.}$$

This holds because if $k \geq \mu$ is an integer, then $a_{T_n} \geq a_{nk}$ a.s. for large n by the strong law of large numbers and obviously $\sum_n a_{nk} = \infty$ for all fixed k. Finally, apply this result to $a_n = P_s(\log^+ \bar{\xi} > cn)$. ■

An obvious special case of Proposition 6.1 is if the type space is finite. We leave it as an open problem whether it is possible to construct a branching process that has $E_\pi[\bar{\xi}\log^+\bar{\xi}] = \infty$, $E_s[\bar{\xi}\log^+\bar{\xi}] < \infty$ for all $s \in S$, and $\sum_{n=1}^{\infty} P_{\sigma_n}(\log^+\bar{\xi} > cn) < \infty \quad P_\nu -$ a.s., thus demonstrating a case when the $x\log x$ condition is not necessary.

Another approach is to consider the rate of convergence toward the stationary distribution ν; if this convergence is fast enough, necessity of the $x\log x$ condition follows. To simplify the analysis, let $Y = \sum_n \log^+\bar{\xi} \cdot I_{\{n-1\le\bar{\xi}\le n\}}$ which is nonnegative integer-valued and has finite mean if and only if $\bar{\xi}$ does. The result is

Proposition 6.2 *Suppose that*

$$\sum_{k\ge 1} nE_\nu \left| \frac{1}{n}\sum_{k=1}^{n} P_{\sigma_k}(Y=n) - P_\nu(Y=n)\right| < \infty$$

Then the $x\log x$ condition is necessary.

Proof. The condition in the proposition implies that

$$\sum_{n\ge 1} n \left| \frac{1}{n}\sum_{k=1}^{n} P_{\sigma_k}(Y=n) - P_\nu(Y=n)\right| < \infty \quad P_\nu\text{-a.s.}$$

which yields

$$\begin{aligned}
\sum_{k=1}^{\infty} P_{\sigma_k}(Y>k) &= \sum_{k=1}^{\infty}\sum_{n>k} P_{\sigma_k}(Y=n) \\
&= \sum_{n>1} n\left(\frac{1}{n}\sum_{k=1}^{n} P_{\sigma_k}(Y=n)\right) \\
&\ge \sum_{n>1} nP_\nu(Y=n) - \sum_{n\ge 1} n\left|\frac{1}{n}\sum_{k=1}^{n} P_{\sigma_k}(Y=n) - P_\nu(Y=n)\right| = \infty
\end{aligned}$$

if the first term is infinite and the second is finite. ■

Note that if $\{\sigma_n\}$ is positive Harris recurrent with stationary distribution ν, the ergodic theorem yields

$$\frac{1}{n}\sum_{k=1}^{n} P_{\sigma_k}(Y=j) \to P_\nu(Y=j) \quad P_\nu\text{-a.s.}$$

for all j so our condition means that this convergence is, in some sense, "fast enough."

Another generalization of Borel–Cantelli is the Kochen-Stone lemma that states that if $\sum_n P(A_n) = \infty$, then

$$P(A_n \text{ i.o.}) \geq \limsup_n \frac{\{\sum_{k=1}^n P(A_k)\}^2}{\sum_{1\leq j,k\leq n} P(A_j \cap A_k)}$$

We can apply it to prove

Proposition 6.3 *Let $A_n = \{\log^+ \bar{\xi}_n > cn\}$. If the (indicators of the) A_n are pairwise negatively correlated, the $x \log x$ condition is necessary.*

Proof. Becuase $\widetilde{P}_\nu(A_j \cap A_k) \leq \widetilde{P}_\nu(A_j)\widetilde{P}_\nu(A_k)$, we get

$$\widetilde{P}_\nu(A_n \text{ i.o.}) \geq \limsup_n \frac{\{\sum_{k=1}^n \widetilde{P}_\nu(A_k)\}^2}{\sum_{1\leq j,k\leq n} \widetilde{P}_\nu(A_j \cap A_k)} \geq 1$$

■

7 Acknowledgements

The author would like to thank Gerold Alsmeyer and Olle Häggström for fruitful discussions, the former of which took place more than 10 years ago.

8 References

ALSMEYER, G. (1994) On the Markov renewal theorem. *Stoch. Proc. Appl.* **50**, 37–56.

ATHREYA, K.B. (2000) Change of measures for Markov chains and the LlogL theorem for branching processes. *Bernoulli* **6(2)**, 323–338.

ATHREYA, K.B. and NEY, P. (1972) Branching Processes. Springer, Berlin.

BIGGINS, J.D. and KYPRIANOU, A. E. (2004) Measure change in multitype branching. *Adv. in Appl. Prob.* **36(2)**, 544–581

DURRETT, R. (2005) Probability: Theory and Examples, third edition. Duxbury Press.

JAGERS, P. (1989) General branching processes as Markov fields. *Stoch. Proc. Appl.* **32**, 183–212.

JAGERS, P. (1992) Stabilities and instabilities in population dynamics. *J. Appl. Prob.* **29**, 770–780.

JAGERS, P. and NERMAN, O. (1984) The growth and composition of branching populations. *Adv. Appl. Prob.* **16**, 221–259.

JOFFE, A. and WAUGH, W. A. O'N. (1982) Exact distributions of kin numbers in a Galton-Watson process. *J. Appl. Prob.* **19(4)**, 767–775.

LYONS, R., PEMANTLE, R. and PERES, Y. (1995) Conceptual proofs of L log L criteria for mean behaviour of branching processes. *Annals of Probability* **3**, 1125–1138.

NERMAN, O. (1981) On the convergence of supercritical general (C–M–J) branching processes. *Z. Wahrscheinlichkeitsth.* **57**, 365–395.

OLOFSSON, P. (1996) General branching processes with immigration. *J. Appl. Prob.*, **33**, 940–948.

OLOFSSON, P. (1998) The $x \log x$ condition for general branching processes. *J. Appl. Prob.*, **35**, 537–544.

OLOFSSON, P. and SHAW (2002). Exact sampling formulas for multi-type Galton-Watson processes. *J. Math. Biol*, **45**, 279–293

WAYMIRE, E. and WILLIAMS, S.C. (1996) A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* **348(2)**, 585–632.

Peter Olofsson
Mathematics Department
Trinity University
One Trinity Place
San Antonio, TX 78212

polofsso@trinity.edu

www.trinity.edu/polofsso