Measure change in multitype branching*

J.D. Biggins[†]
The University of Sheffield

A.E. Kyprianou[‡]
The University of Utrecht

Running head: Multitype branching

Abstract

The Kesten-Stigum theorem for the one-type Galton-Watson process gives necessary and sufficient conditions for mean convergence of the martingale formed by the population size normed by its expectation. Here, the approach of Lyons, Peres and Pemantle (1995) to this theorem, which exploits a change of measure argument, is extended to martingales defined on Galton-Watson processes with a general type space through non-negative functions that are harmonic for the mean kernel. Many examples satisfy stochastic domination conditions on the offspring distributions that combine with the measure change argument to produce moment conditions, like the $X \log X$ condition of the Kesten-Stigum theorem; a general treatment of this phenomenon is given. The application of the approach to branching processes in varying environments and random environments is indicated; the results also apply to the general (Crump-Mode-Jagers) branching process once suitable results on what are called optional lines are obtained. However, the main reason for developing the theory was to obtain martingale convergence results in branching random walk that did not seem readily accessible with other techniques. These results, which are natural extensions of known results for martingales associated with binary branching Brownian motion, form the main application.

1 Introduction

Let $\mathcal{T}$ be the labelled nodes of the family tree in which every node has a countably infinite number of children; $\mathcal{T}$ is the (countable) index set for the process. Write $|\nu|$ for the generation of the node ν, $c(\nu)$ for the children of ν and 0 for the initial ancestor. Each node has a type (or mark), drawn from $\mathcal{S}$; more precisely, let S be a $\mathcal{S}$-valued function on the nodes, which therefore takes values in $\mathfrak{B} = S^{\mathcal{T}}$. The types of the first generation are given by an element of $\mathcal{S}^{\mathbb{N}} = \{(S_1, S_2, S_3, \ldots) : S_i \in \mathcal{S}\}$.

**Keywords and phrases.* branching, measure change, multitype, branching random walk, varying environment, random environment, martingales, harmonic functions, Crump-Mode-Jagers process, optional lines

MS2000 subject classification. Primary 60J80, secondary 60G42,

[†] *email:* J.Biggins@sheffield.ac.uk

[‡] *email:* kyprianou@math.uu.nl

The family size distributions, $\{P_s : s \in \mathcal{S}\}$, from which the branching process is built, are distributions on $\mathcal{S}^{\mathbb{N}}$, the possible types of the first generation; the corresponding expectation is E_s. Let $\mathfrak{B}_n$ be obtained by projecting $\mathfrak{B}$ onto the first n generations. In a slight abuse of notation, P_s is also used for the measure on $\mathfrak{B}_1$ that has $S(0) = s$ with probability one and then picks the first generation types according to P_s. A branching process is a Markov chain with a state in $\mathfrak{B}_n$ at time n and transition probabilities defined by the nth generation nodes producing independent families, with the distribution of the family of a node of type s being P_s. The resulting law for the chain is denoted by $\mathbb{B}$; for notational simplicity, the starting type, $S(0)$, is not explicitly recorded. This slightly unconventional definition produces more familiar branching processes when one of the types in $\mathcal{S}$ is a 'ghost'-type, ∂, that always has all its children of type ∂ and is interpreted as the node being absent. A realization of this chain can be identified, in the obvious way, with an element of $\mathfrak{B}$, and the measure describing the evolution can then be transferred to a measure on $\mathfrak{B}$. Let $\mathcal{G}_n$ be the σ-algebra generated by the types in the first n generations.

A finite non-negative function H on $\mathcal{S}$ will be called mean-harmonic when, abusing the notation E_s as promised,

$$E_s\left[\sum_i H(S_i)\right] = E_s\left[\sum_{|\nu|=1} H(S(\nu))\right] = H(s) \quad \text{for all } s \in \mathcal{S}$$

and $H(\tilde{s}) > 0$ for some $\tilde{s}$. Since we deal with a single such H, it will be convenient for many calculations to let h be the composition of H and S, a function from nodes to the non-negative reals. Throughout, the starting type is picked from $\mathcal{S}^H = \{s \in \mathcal{S} : H(s) > 0\}$ so that $H(S(0)) = h(0) > 0$.

The functions $\{W_n\}$ are defined by

$$W_n(S) = \sum_{|\sigma|=n} H(S(\sigma)) = \sum_{|\sigma|=n} h(\sigma) = \sum_{|\nu|=n-1} \sum_{\sigma \in c(\nu)} h(\sigma);$$

then

$$E_{\mathbb{B}}\left[W_n \middle| \mathcal{G}_{n-1}\right] = \sum_{|\nu|=n-1} E_{S(\nu)}\left[\sum_i H(S_i)\right] = \sum_{|\nu|=n-1} H(S(\nu)) = W_{n-1}$$

and so W_n forms a non-negative martingale with respect to $\mathcal{G}_n$ with $W_0 = h(0)$. Let $W = \limsup_n W_n$; of course W is actually $\lim_n W_n$ almost surely under $\mathbb{B}$, but it is convenient to have it defined everywhere. The main objective is to give conditions that determine when the martingale converges in mean; that is to obtain Kesten-Stigum like results for such martingales. The method used has been employed in various special cases of the framework adopted here. It is a natural extension and refinement of that employed by Lyons, Peres and Pemantle (1995), Lyons (1997) and Athreya (2000), and the connections between this treatment and those are not hard to see. The discussion in Waymire and Williams (1996) also has much in common with that here but it mostly confines branching to a b-ary tree and so is not directed towards classical Kesten-Stigum theorems; however, their framework is at first sight, rather different from here and so some points of contact are noted in Section 4. The key idea in all these papers is to exploit a change of measure to establish when the martingale converges in mean; the actual measure change has much longer history, as can be seen from the references in Lyons (1997).

In many branching models, the assertion that, when $E_{\mathbb{P}}W = 1$, the process dies out on $W = 0$ is included as part of a Kesten-Stigum theorem. Here, let Z_n be the number of nth generation nodes in S^H, then this assertion corresponds to the difference between the events $\{Z_n \to 0\}$ and $\{W = 0\}$ being $\mathbb{P}$-null, when $E_{\mathbb{P}}W = 1$. Examples mentioned in Section 4 show that this can fail. Clearly $\{Z_n \to 0\} \subset \{W = 0\}$, and, for simple models, showing that $\mathbb{P}(W = 0)$ satisfies the same equation as the extinction probability then settles the issue. No attempt is made here to tackle the generalization of this part of the Kesten-Stigum Theorem.

The generality of the type space brings many particular branching processes within the scope of the theorems. Some of these are discussed briefly later to illustrate this. However, the original motivation for this extension of earlier work was to study the convergence of a 'derivative' martingale — a name explained when it is defined properly in Section 8 — for a certain boundary case in the homogeneous branching random walk. The derivative martingale considered is the natural analogue of a martingale arising in binary branching Brownian motion that is associated with minimal-speed travelling wave of the KPP equation. Results on the convergence of that martingale, discussed in Neveu (1988) and Harris (1999), lead naturally to questions answered here for the branching random walk. An essential feature of this martingale is that it has the form of $\{W_n\}$ except that the corresponding function H takes negative values, so its convergence is not guaranteed. However, the derivative martingale turns out to be naturally connected to a non-negative one arising in the branching random walk with a barrier, which can be studied by the methods developed. The results on the derivative martingale also produce, as a by-product, information about an associated functional equation for the Laplace transform of the limit, corresponding to the smoothing transform of Durrett and Liggett (1983); this connection will be explained in Biggins and Kyprianou (2001).

To describe neatly the two measures that arise and their relationship, it turns out to be useful to augment the basic space by picking out a single line of descent. Formally, let $\xi = (\xi_0, \xi_1, \xi_2, \ldots)$ be a sequence drawn from $\mathcal{T}$ with $\xi_0 = 0$, and $\xi_{n+1} \in c(\xi_n)$. Thus ξ defines a line of descent starting from the initial ancestor. Let Ξ be the set of possible ξ. The new space is $\mathfrak{T} = \mathcal{S}^{\mathcal{T}} \times \Xi$ $(= \mathfrak{B} \times \Xi)$, its projection onto the first n-generations is $\mathfrak{T}_n$ and a branching process will now be a Markov chain with state in $\mathfrak{T}_n$ at time n. The line of descent ξ will be called the trunk — other names have also been used. (Informally, the "trunk" is what distinguishes the "bushes" which make up $\mathfrak{B}$, in which every branch is similar, from the "trees" which make up $\mathfrak{T}$, in which the "trunk" has special status.) Let $\mathcal{F}_n$ be the σ-algebra generated by the first n steps of the Markov chain, that is, the information on the development for the first n generations. Let $\mathcal{F}_n^*$ be the σ-algebra generated by $\mathcal{F}_{n-1}$ and the types of the nth generation, $\{S(\nu) : |\nu| = n\}$, but not the identification of the trunk at the nth generation, ξ_n; hence $\mathcal{F}_{n-1} \subset \mathcal{F}_n^* \subset \mathcal{F}_n$.

In producing the measure on this enlarged space, ξ is produced by an extra randomisation. Thus, the various types reproduce as before and then ξ_{n+1} is picked from the children of ξ_n, with probabilities proportional to the children's values of h when this makes sense. More precisely,

$$P(\xi_{n+1} = \nu | \mathcal{F}^*_{n+1}) = \frac{h(\nu) I(\nu \in c(\xi_n))}{\sum_{\sigma \in c(\xi_n)} h(\sigma)} \text{ when } \sum_{\sigma \in c(\xi_n)} h(\sigma) \in (0, \infty), \tag{1}$$

and is some arbitrary, but fixed, probability distribution on $c(\xi_n)$ otherwise. This defines a branching process with a trunk; call its probability law $\mathbb{P}$ and its expectation $E_{\mathbb{P}}$. There is no reason at the moment to use h to weight the possibilities in picking the trunk, but one will emerge. By construction, integrating out ξ maps $(\mathfrak{T}, \mathbb{P})$ to $(\mathfrak{B}, \mathbb{B})$.

Another approach to the construction starts by doubling the type space, working with $\mathcal{S} \times \{1, 2\}$. Types in $\mathcal{S}_1$ reproduce as before, producing only types in $\mathcal{S}_1$. For $s \in \mathcal{S}_2$, use P_s to generate a family from $\mathcal{S}^{\mathbb{N}}$; given the family, pick child j with probability $H(S_j)/(\sum_i H(S_i))$ when $0 < \sum_i H(S_i) < \infty$, and pick a child according to some fixed, arbitrary distribution otherwise; the chosen child is given its type (as generated in $\mathcal{S}^{\mathbb{N}}$) in $\mathcal{S}_2$, every other has its type in $\mathcal{S}_1$. Nodes in $\mathcal{S}_2$ give ξ.

An auxiliary branching process with a trunk, which will turn out to result from the change of measure, is described next. To define the development of this Markov chain, assume the state for the the first n generations is known. Then, reproduction from nth generation nodes not on the trunk, that is from $\nu \in \{\sigma : |\sigma| = n, \sigma \neq \xi_n\}$, is exactly the same as in $\mathbb{P}$ (or $\mathbb{B}$). When $S(\xi_n) = s$, the types of the children of ξ_n are given by generating a family from $\mathcal{S}^{\mathbb{N}}$ with the law having (Radon-Nikodym) derivative $\sum_i H(S_i)/H(s)$ with respect to P_s when $H(s) > 0$ and, for completeness, 1 when $H(s) = 0$. Finally, given the types, ξ_{n+1} is chosen exactly as previously, that is as in $\mathbb{P}$. Call the resulting measure $\mathbb{Q}$. To express the derivative more neatly, and for later developments, let $X(\nu)$ be defined by

$$\begin{aligned} X(\nu) &= I(H(S(\nu)) = 0) + I(H(S(\nu)) > 0) \frac{\sum_{\sigma \in c(\nu)} H(S(\sigma))}{H(S(\nu))} \\ &= I(h(\nu) = 0) + I(h(\nu) > 0) \frac{\sum_{\sigma \in c(\nu)} h(\sigma)}{h(\nu)}. \end{aligned}$$

Then, in constructing $\mathbb{Q}$, the types of the children of ξ_n are given by generating a family with the law having derivative $X(\xi_n)$ with respect to $P_{S(\xi_n)}$.

By assumption $h(\xi_0) = h(0) > 0$; under $\mathbb{Q}$, reproduction from ξ_0 produces, with probability one, a family with $0 < \sum_{\sigma \in c(0)} h(\sigma) < \infty$, which yields, using (1), ξ_1 with $h(\xi_1) > 0$, and so $h(\xi_n) > 0$ for all n. More formally, the following simple result says that under $\mathbb{Q}$ the types on the trunk develop as a Markov chain on $\mathcal{S}^H$ with a transition kernel arising naturally from the mean-harmonic function H.

Theorem 1 *Let $\zeta_n = S(\xi_n)$. Then $\zeta = \{\zeta_0, \zeta_1, \ldots\}$ forms a Markov chain on $\mathcal{S}^H$ under $\mathbb{Q}$ with the proper transition measure given by*

$$\frac{1}{H(s)} E_s \left[\sum_i H(S_i) I(S_i \in A) \right] \quad \textit{for } A \subset \mathcal{S}^H.$$

Proof. Compute $E_{\mathbb{Q}}[I(S(\xi_{n+1}) \in A)|\mathcal{F}_n]$ on $S(\xi_n) \in \mathcal{S}^H$. $\square$

For simplicity, write X for $X(0)$; which, under the prevailing assumption that $S(0) \in \mathcal{S}^H$, simplifies to $X = \sum_{\nu \in c(0)} h(\nu)/h(0)$. It turns out that the interplay between the development under $\mathbb{Q}$ of the Markov chain ζ and the distribution of X under P_{ζ_n} often determines when W_n converges in mean. The next theorem, which is a special case of Corollary 2 given later, illustrates this. In it, and the remainder of the paper, unadorned P and E will be used for probability and expectation on an (undefined) auxiliary probability space.

Theorem 2 *For $x > 0$, let $A(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x > 1)$. Suppose the positive increasing function L is slowly varying at infinity and that $\delta \in [0, 1)$; L and δ may be different in (i) and (ii).*

(i) Suppose that there is a random variable X^* with

$$P_s(X > x) \leq P(X^* > x) \quad \textit{for all } s \in \mathcal{S}$$

and that $\sup_{x>0}\{A(x)/(x^{\delta}L(x))\}$ is bounded above, $\mathbb{Q}$ almost surely (A is random!). If

$$E[(X^*)^{1+\delta}L(X^*)] < \infty$$

then $E_{\mathbb{P}}W = h(0)$.

(ii) Suppose that there is a random variable X_* with

$$P_s(X > x) \geq P(X_* > x) \quad \textit{for all } s \in \mathcal{S}$$

and that, for some y, $\inf_{x>y}\{A(x)/(x^{\delta}L(x))\}$ is bounded below by a positive constant, $\mathbb{Q}$ almost surely. If

$$E[(X_*)^{1+\delta}L(X_*)] = \infty$$

then $E_{\mathbb{P}}W = 0$.

The application of this result to the simplest (one-type) Galton-Watson process with generic family size N satisfying $EN = m \in (1, \infty)$ is now indicated briefly. First, type individuals by their generation, so the type space is $\mathcal{S} = \{\partial, 0, 1, 2, \ldots\}$; a person of type i gives birth to N' children of type $i+1$, where N' is just a copy of N, and the remaining children are of type ∂. Now the function H defined by $H(n) = m^{-n}$ and $H(\partial) = 0$ is mean harmonic and then W_n is the usual martingale. Both X^* and X_* can be N/m, $\zeta_i = i$ and so $A(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x > 1) = \sum_{i=1}^{\infty} I(m^i < x) \approx \log x / \log m$. Thus the two parts of the theorem combine to show the martingale converges in mean exactly when $EN \log N < \infty$ and the limit is zero when this fails, which is a Kesten-Stigum theorem.

We now summarise how the treatment will develop. The next section establishes general results about the mean convergence of the martingale W_n through the measure change argument. Often, particular models have, or are most easily understood under, stochastic bounds on the reproduction of the form employed in Theorem 2. The machinery to consider such cases is established in Section 3. Then, to make a break from general theory, several examples are discussed briefly to show how they fit into the framework;

these are the (one-type) Galton-Watson processes in a varying environment, the homogeneous Galton-Watson process with a finite irreducible type space, and the branching random walk in a random, ergodic environment. Some new results are obtained but the main point is to illustrate the applicability of the general results; it is clear that a similar discussion of other examples could be given.

Returning to general issues, there are natural reasons to want to consider the sum of $h(\nu)$ over collections of nodes other than the nth generation ones; specifically over what Jagers (1989) calls optional lines. Section 5 gives conditions for the limit over an increasing sequence of such lines to be W, so that the limit is the same as when the lines are just formed by the generations.

It was already mentioned that the original motivation for this work was to study the convergence of a certain derivative martingale in the homogeneous branching random walk through the branching random walk with a barrier. The latter process and the result obtained for it are described in Section 6 and the proof forms Section 7. Then, this result is used in Section 8 to prove convergence of the derivative martingale, and to give mild conditions for its limit to be non-trivial; also, the theory developed on optional lines is applied there.

2 Martingale measure change and mean convergence

The approach is based on a simplification of a result in Durrett (1996, Theorem 4.3.3) — see also Athreya (2000) — which is now stated and briefly discussed. The notation employed suggests how the result will be used.

Theorem 3 *Suppose $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures and $\mathcal{G}_n$ are increasing σ-algebras. Suppose further that, for all n, $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{G}_n$, with density W_n. Let $W = \limsup_n W_n$. Then*
(i) W_n is a $\mathbb{P}$-martingale and $1/W_n$ is a $\mathbb{Q}$-martingale;
(ii) $\int W d\mathbb{P}$ $(= E_{\mathbb{P}}W) = 1$ if and only if $\mathbb{Q}(W < \infty) = 1$;
(iii) $\int W d\mathbb{P}$ $(= E_{\mathbb{P}}W) = 0$ if and only if $\mathbb{Q}(W = \infty) = 1$.

Any non-negative, mean one, martingale defines a change of probability measure (from $\mathbb{P}$ to $\mathbb{Q}$ above); clearly, if the resulting $\mathbb{Q}$ is tractable it can be used to study the mean convergence of the original martingale through the last two parts of the lemma. When $\mathbb{P}$ is the law of a Markov chain with filtration $\{\mathcal{G}_n\}$, $\mathbb{Q}$ is too; then the measure change is often called Doob's h-transform. Note that this measure change only concerns $\mathbb{P}$ and $\mathbb{Q}$ on the σ-algebra generated by $\{\mathcal{G}_n\}$, leaving some freedom over the definition of $\mathbb{P}$ and hence of $\mathbb{Q}$. In the branching context, the introduction of the trunk exploits this freedom.

Returning to branching processes, recall that, $X(\nu) = \sum_{\sigma \in c(\nu)} h(\sigma)/h(\nu)$ when $h(\nu) > 0$ and is one when $h(\nu) = 0$. Then, when $H(s) > 0$,

$$E_{\mathbb{P}}\left[X(\nu)|S(\nu) = s\right] = \frac{1}{H(s)} E_s \left[\sum_i H(S_i)\right] = 1,$$

because H is mean harmonic, and, by definition, $E_{\mathbb{P}}\left[X(\nu)|S(\nu) = s\right] = 1$ when $H(s) = 0$. Exploiting the trunk, there is now a simpler martingale than W_n that can be constructed, by forming a product (down ξ) using these adapted positive terms with expectation one. In fact it is useful to define these products for any node. To do this, let $\{\nu_i : i = 0, 1, \ldots |\nu|\}$ be the ancestry of ν ordered in the natural way, starting from $\nu_0 = 0$. Now (with $0.\infty = 0$), let

$$\overline{W}(\nu) = \prod_{i=0}^{|\nu|-1} X(\nu_i).$$

It turns out that $\overline{W}(\xi_n)$ is a martingale linking, in the sense of Theorem 3, $\mathbb{P}$ and $\mathbb{Q}$. The probability laws $\mathbb{P}$ and $\mathbb{Q}$ are constructed from conditional probabilities (using the Theorem of Ionescu Tulcea) defined on $\mathfrak{T}_{n+1}$ given the state in $\mathfrak{T}_n$. The following straightforward lemma, on derivatives, from measure theory is the key to the relationship between these measures.

Lemma 1 *Let P be a probability measure on $\mathfrak{U}$, p a conditional probability from $\mathfrak{U}$ to $\mathfrak{V}$ and P^* the resulting joint probability measure. Let Q, q and Q^* be defined similarly, with Q absolutely continuous with respect to P and, for each $u \in \mathfrak{U}$, q absolutely continuous with respect to p. Then*

$$\frac{dQ^*}{dP^*} = \frac{dq}{dp}\frac{dQ}{dP}.$$

Lemma 2 *$\overline{W}(\xi_n)$ is the derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{F}_n$.*

Proof. The result is true for $n = 0$, assume it also holds for $n = r$. Let p_{r+1} and q_{r+1} be the conditional probability measures on $\mathfrak{T}_{r+1}$ given the state in $\mathfrak{T}_r$ that are used in the construction of $\mathbb{P}$ and $\mathbb{Q}$ respectively; both p_{r+1} and q_{r+1} are products of the family distributions appropriate to the types of the nodes. To generate the $(r+1)$th generation under $\mathbb{Q}$, all nodes except ξ_r use the same law as in $\mathbb{P}$ and ξ_r uses the law which has the derivative $X(\xi_r)$ with respect to $P_{S(\xi_r)}$. Thus, overall, the derivative dq_{r+1}/dp_{r+1} is $X(\xi_r)$. Let $\mathbb{P}_r$ and $\mathbb{Q}_r$ be $\mathbb{P}$ and $\mathbb{Q}$ restricted to $\mathcal{F}_r$; then, applying Lemma 1,

$$\frac{d\,\mathbb{Q}_{r+1}}{d\,\mathbb{P}_{r+1}} = \frac{dq_{r+1}}{dp_{r+1}}\frac{d\,\mathbb{Q}_r}{d\,\mathbb{P}_r} = X(\xi_r)\overline{W}(\xi_r) = \overline{W}(\xi_{r+1})$$

as required. □

Recall that $\mathcal{G}_n$ is the σ-algebra generated by the types, but not the trunk, in the first n generations. The idea now is to integrate out ξ to get the derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{G}_n$. For this to work, using h to choose the trunk in (1) turns out to be critical. The next lemma gives the essential formula for the integration; it computes $\mathbb{P}[\xi_n = \nu|\mathcal{G}_n]$ on the set $\{\overline{W}(\nu) > 0\}$.

Lemma 3 *For a fixed ν, let $n = |\nu|$. Then*

$$E_{\mathbb{P}}[\overline{W}(\nu)I(\xi_n = \nu)|\mathcal{G}_n] = \overline{W}(\nu)\mathbb{P}[\xi_n = \nu|\mathcal{G}_n] = \frac{h(\nu)}{h(0)}, \quad \mathbb{P} \textit{ almost surely.}$$

Proof. The first equality is obvious; for the second it is necessary to prove rather more than stated. Let $r \geq n$; then we prove that $\overline{W}(\nu)\mathbb{P}[\xi_n = \nu|\mathcal{G}_r] = h(\nu)/h(0)$. This is done by induction on n. The result is true for $n=0$. Suppose it is true for $(n-1)$. Let $\sigma\{\mathcal{F}_n^*, \mathcal{G}_r\}$ be the σ-algebra generated by the two components. Then, since the process is Markov,

$$\mathbb{P}(\xi_n = \nu|\sigma\{\mathcal{F}_n^*, \mathcal{G}_r\}) = \mathbb{P}(\xi_n = \nu|\mathcal{F}_n^*) = p(\nu)I(\xi_{n-1} = \nu_{n-1})$$

where $\{p(\nu) : \nu \in c(\nu_{n-1})\}$ is a proper probability distribution that is $\mathcal{G}_n$-measurable. Thus, taking expectations conditional on $\mathcal{G}_r$, multiplying by $\overline{W}(\nu) = X(\nu_{n-1})\overline{W}(\nu_{n-1})$ and using the result for $(n-1)$,

$$\begin{aligned}\overline{W}(\nu)\mathbb{P}(\xi_n = \nu|\mathcal{G}_r) &= p(\nu)X(\nu_{n-1})\overline{W}(\nu_{n-1})\mathbb{P}(\xi_{n-1} = \nu_{n-1}|\mathcal{G}_r) \\ &= p(\nu)X(\nu_{n-1})h(\nu_{n-1})/h(0).\end{aligned}$$

When $h(\nu_{n-1})X(\nu_{n-1}) \in (0, \infty)$, $h(\nu_{n-1})X(\nu_{n-1}) = \sum_{\sigma\in c(\nu_{n-1})} h(\sigma)$ and, from (1),

$$p(\nu) = \frac{h(\nu)}{\sum_{\sigma\in c(\nu_{n-1})} h(\sigma)} = \frac{h(\nu)}{h(\nu_{n-1})X(\nu_{n-1})};$$

substitution now shows that the formula holds. This leaves cases where $h(\nu_{n-1})X(\nu_{n-1}) \notin (0, \infty)$: if $X(\nu_{n-1}) = 0$ then $h(\nu) = 0$; if $h(\nu_{n-1}) = 0$ then, since

$$E\left[\left.\sum_{\sigma\in c(\nu_{n-1})} h(\sigma)\right| \mathcal{G}_{n-1}\right] = h(\nu_{n-1}) = 0,$$

$h(\nu) = 0$ almost surely; finally, $E_{\mathbb{P}}[h(\nu_{n-1})X(\nu_{n-1})] \leq E_{\mathbb{P}}[W_{n-1}] = h(0) < \infty$ so that $I(h(\nu_{n-1})X(\nu_{n-1}) = \infty)$ is $\mathbb{P}$-null. □

Proposition 1 *$W_n/h(0)$ is the derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{G}_n$. Hence Theorem 3 applies.*

Proof. The derivative sought equals $E_{\mathbb{P}}[\overline{W}(\xi_n)|\mathcal{G}_n]$. For a fixed ν with $|\nu| = n$,

$$E_{\mathbb{P}}[\overline{W}(\xi_n)I(\xi_n = \nu)|\mathcal{G}_n] = E_{\mathbb{P}}[\overline{W}(\nu)I(\xi_n = \nu)|\mathcal{G}_n] = \frac{h(\nu)}{h(0)},$$

by Lemma 3; thus

$$E_{\mathbb{P}}[\overline{W}(\xi_n)|\mathcal{G}_n] = E_{\mathbb{P}}\left[\left.\sum_{|\nu|=n} \overline{W}(\nu)I(\xi_n = \nu)\right| \mathcal{G}_n\right] = \frac{W_n}{h(0)}.$$

□

The next proposition is an intermediate result that it is convenient to isolate on the way to the main theorems on mean convergence of the martingale W_n to W.

Let $c'(\xi_n)$ be the children of ξ_n excluding ξ_{n+1} and let, by analogy with X and $X(\xi_i)$,

$$X' = \frac{\sum_{\nu\in c'(0)} h(\nu)}{h(0)} \quad \text{and} \quad X'(\xi_i) = \frac{\sum_{\nu\in c'(\xi_i)} h(\nu)}{h(\xi_i)}.$$

so that $h(\xi_i)X(\xi_i) = h(\xi_i)X'(\xi_i) + h(\xi_{i+1})$. Then $h(\xi_i)$ tracks the value of H along the types in the trunk, while X' concerns the reproduction along the trunk.

Proposition 2

(i) If

$$\mathbb{Q}\left(\liminf_n h(\xi_n) < \infty,\ \sum_{i=1}^{\infty} h(\xi_i)X'(\xi_i) < \infty\right) > 0 \qquad (2)$$

or

$$\mathbb{Q}\left(\sum_{i=1}^{\infty} h(\xi_i)X(\xi_i) < \infty\right) > 0, \qquad (3)$$

which implies (2), then $E_{\mathbb{P}}W > 0$*. Furthermore* $E_{\mathbb{P}}W = h(0)$*, and so* $\{W_n\}$ *converges in* $\mathbb{P}$*-mean, when the probability in either (2) or (3) is one.*

(ii) If

$$\mathbb{Q}\left(\limsup_n h(\xi_n)X(\xi_n) = \infty\right) > 0 \qquad (4)$$

then $E_{\mathbb{P}}W < h(0)$ *and so* $\{W_n\}$ *does not converge in* $\mathbb{P}$*-mean. Furthermore,* $E_{\mathbb{P}}W = 0$ *when this probability is one.*

Proof. Let S^ν be the function S on the sub-tree rooted at ν. Then, by partitioning the sum using the sub-trees emanating from the siblings of $\xi_1, \xi_2, \ldots, \xi_{n-1}$

$$W_n(S) = \sum_{|\nu|=n} h(\nu) = h(\xi_n) + \sum_{i=1}^{n-1} \sum_{\nu\in c'(\xi_i)} h(\nu)\frac{W_{n-i}(S^\nu)}{h(\nu)}.$$

Let $\mathcal{H}$ be the σ-algebra generated by the reproduction of the members of the trunk. (Technically, in the language of Jagers (1989), with L the optional line formed by all non-trunk children of the nodes forming the trunk, $\mathcal{H}$ is the pre-L sigma algebra.) The construction of $\mathbb{Q}$ means that away from the trunk it looks just like $\mathbb{P}$, and so $E_{\mathbb{Q}}[W_{n-i}(S^\nu)|\mathcal{H}]$ is $h(\nu)$ when $\nu \in c'(\xi_i)$. Since $1/W_n(S)$ is a positive martingale under $\mathbb{Q}$, $W_n(S)$ converges to W, $\mathbb{Q}$ almost surely. Then, by Fatou,

$$\begin{aligned}
E_{\mathbb{Q}}[W|\mathcal{H}] &= E_{\mathbb{Q}}[\lim_n W_n(S)|\mathcal{H}] \\
&\le \liminf_n \left(h(\xi_n) + \sum_{i=1}^{n-1}\sum_{\nu\in c'(\xi_i)} h(\nu)\right) \\
&= \liminf_n h(\xi_n) + \sum_{i=1}^{\infty}\sum_{\nu\in c'(\xi_i)} h(\nu) \\
&= \liminf_n h(\xi_n) + \sum_{i=1}^{\infty} h(\xi_i)X'(\xi_i) \le \sum_{i=1}^{\infty} h(\xi_i)X(\xi_i),
\end{aligned}$$

$\mathbb{Q}$ almost surely. Hence (3) implies (2) and either implies that $\mathbb{Q}(W < \infty) > 0$; in addition, $\mathbb{Q}(W < \infty) = 1$ when either probability is one. Theorem 3 now gives the conclusion to the first part.

For the second half, note that

$$W_n(S) = \sum_{|\nu|=n} h(\nu) \ge \sum_{\nu\in c(\xi_{n-1})} h(\nu) = h(\xi_{n-1})X(\xi_{n-1})$$

and so

$$\mathbb{Q}(W=\infty)\geq \mathbb{Q}\left(\limsup_n h(\xi_n)X(\xi_n)=\infty\right).$$

A further application of Theorem 3 completes the proof. □

Corollary 1 *If* $\mathbb{Q}(\limsup_n H(\zeta_n)=\infty)=1$ *then* $E_{\mathbb{P}}W=0$.

Proof. Since $H(\zeta_n)=H(S(\xi_n))=h(\xi_n)$, the result follows from Proposition 2(ii) and the fact that $h(\xi_i)X(\xi_i)\geq h(\xi_{i+1})$. □

Observe that $h(\xi_n)$, $X'(\xi_{n-1})$ and $X(\xi_{n-1})$ are $\mathcal{F}_n$-measurable; therefore, the series $\sum h(\xi_i)X'(\xi_i)$ and $\sum h(\xi_i)X(\xi_i)$ are amenable to the following standard result, proved by truncation and conditional Borel-Cantelli. It and the lemma after it translate the conditions in Proposition 2(i) to ones involving $\{H(\zeta_n), P_{\zeta_n}\}$, the development of the types of the trunk and the associated family laws, to give the main theorem.

Lemma 4 *Suppose Y_i are non-negative variables that are adapted to the increasing σ-algebras $\mathcal{F}_i$. Let $t(x)=I(x\geq 1)+xI(x<1)$. Then*

$$I\left(\sum_i Y_i<\infty\right)=I\left(\sum_i E[t(Y_{i+1})|\mathcal{F}_i]<\infty\right)\quad \textit{almost surely.}$$

Lemma 5 *Recall that $X=X(0)$. Let Y be a non-negative function on $\mathfrak{B}_1$ and $Y(\nu)$ the corresponding function of the reproduction from node ν. Then*

$$E_{\mathbb{Q}}[Y(\xi_i)|\mathcal{F}_i]=E_{S(\xi_i)}[XY]=E_{\zeta_i}[XY].$$

In particular, for any non-negative f,

$$E_{\mathbb{Q}}[f(X(\xi_i))|\mathcal{F}_i]=E_{\zeta_i}[Xf(X)] \text{ and } E_{\mathbb{Q}}[f(X'(\xi_i))|\mathcal{F}_i]=E_{\zeta_i}[Xf(X')].$$

Proof. This is no more than the definitions. Firstly, $S(\xi_i)=\zeta_i$. Secondly, under $\mathbb{Q}$, ξ_i produces children typed according to the law that has derivative $X(\xi_i)$ with respect to $P_{S(\xi_i)}$. □

Theorem 4 *As in Lemma 4, $t(x)=I(x\geq 1)+xI(x<1)$.*

(i) If, with $\mathbb{Q}$-probability one,

$$\liminf_n H(\zeta_n)<\infty \quad \textit{and} \quad \sum_{i=1}^{\infty} E_{\zeta_i}\left[Xt(H(\zeta_i)X')\right]<\infty, \tag{5}$$

or

$$\sum_{i=1}^{\infty} E_{\zeta_i}\left[Xt(H(\zeta_i)X)\right]<\infty \tag{6}$$

then $E_{\mathbb{P}}W=h(0)$.

(ii) If for all $y>0$, with $\mathbb{Q}$-probability one,

$$\sum_{i=1}^{\infty} E_{\zeta_i}[XI(H(\zeta_i)X\geq y)]=\infty \tag{7}$$

then $E_{\mathbb{P}}W=0$.

Proof. Apply Proposition 4 and Lemma 5 to the series in (2) and (3) for the first part; for the second part, conditional Borel-Cantelli and Lemma 5 show that (7) implies that the probability in (4) is one. □

Splitting t into its two parts splits (6) into two; the first of the sums is the one in (7) when $y = 1$. This indicates that (6) and (7) are quite close; a necessary and sufficient condition for mean convergence of the martingale will be obtained when there are no intermediate cases. A version of Theorem 4 could be formulated for the cases where the probabilities in (2), (3) and (4) are positive, rather than one, translating those parts of Proposition 2.

The collection $\{\overline{W}(\nu) : \nu \in \mathcal{T}\}$ is, essentially, a positive $\mathcal{T}$-martingale; see Waymire and Williams (1996), and references therein. That discussion takes such martingales, also called multiplicative cascades, as the fundamental object, whereas here it is the multitype branching process. The connection is discussed further in Section 4.

3 Stochastically bounded reproduction

The conditions in Theorem 4 can be simplified to moment conditions in many examples where there are bounds on $P_s(X > x)$ that are uniform in the type s. The next two elementary lemmas establish the framework for this. The first is well-known and proved by integration by parts.

Lemma 6 *If $P(X > x) \leq P(X^* > x)$ then, for any increasing non-negative function f, $E[f(X)] \leq E[f(X^*)]$.*

Lemma 7 *Suppose η is a measure on $(0, \infty)$, $A(x) = \eta(0, x]$ and, as in Lemma 4, $t(x) = I(x \geq 1) + xI(x < 1)$. Then*

$$\int t(x/y)\eta(dy) = \int_1^\infty \frac{A(wx)}{w^2} dw.$$

Proof.

$$\begin{aligned}
\int t(x/y)\eta(dy) &= \int \left(I(x \geq y) + xy^{-1}I(x < y)\right)\eta(dy) \\
&= A(x) + x\int_x^\infty y^{-1}\eta(dy) \\
&= A(x) + x\int_x^\infty \left(\int_y^\infty z^{-2}dz\right)\eta(dy) \\
&= A(x) + x\int_x^\infty z^{-2}(A(z) - A(x))dz \\
&= \int_1^\infty w^{-2}A(wx)dw,
\end{aligned}$$

as required. □

Using this Lemma the conditions involving t in Theorem 4 will be replaced by conditions on a suitable A. In the next theorem, note that A (and the associated measure

η) is a function of the development of the chain giving types along the trunk, that is of ζ, and so is random, and that the expectation in (8) and (9) is only over the auxiliary random variable X^*, not over A, which accounts for the qualification '$\mathbb{Q}$ almost surely'. For orientation, the theorem can be read first assuming the function g always takes the value one and the stochastic bounds hold for all types, that is with $F = \mathcal{S}$.

Theorem 5

(i) Suppose that there is a random variable X^*, a non-negative function g on $\mathcal{S}$ and a subset $F \subseteq \mathcal{S}$ such that

$$P_s(X > x) \leq P(g(s)X^* > x) \quad \text{for all } s \in F \subset \mathcal{S}$$

and that ζ is eventually in F, $\mathbb{Q}$ almost surely. Let the increasing function A be defined by

$$A(x) = \sum_i g(\zeta_i) I(xg(\zeta_i)H(\zeta_i) \geq 1),$$

which corresponds to the measure η with mass $g(\zeta_i)$ at $(g(\zeta_i)H(\zeta_i))^{-1}$ for each i in the notation of Lemma 7. If

$$\int_1^\infty \frac{E\left[X^* A(wX^*)\right]}{w^2} dw < \infty \qquad (\mathbb{Q} \text{ almost surely}) \tag{8}$$

then $E_{\mathbb{P}}W = h(0)$.

(ii) Suppose that there is a random variable X_*, a non-negative function g on $\mathcal{S}$ and a subset $F \subseteq \mathcal{S}$ such that

$$P_s(X > x) \geq P(g(s)X_* > x) \quad \text{for all } s \in F \subset \mathcal{S}.$$

Let

$$A(x) = \sum_i g\left(\zeta_i\right) I\left(xg\left(\zeta_i\right) H\left(\zeta_i\right) \geq 1\right) I(\zeta_i \in F).$$

If, for all $w > 0$,

$$E[X_* A(X_* w)] = \infty \qquad (\mathbb{Q} \text{ almost surely}) \tag{9}$$

then $E_{\mathbb{P}}W = 0$.

Proof. Since $xt(hx) = x(I(hx \geq 1) + hxI(hx < 1))$ is an increasing function of x, applying Lemma 6 shows that, when $\zeta_i \in F$,

$$E_{\zeta_i}\left[Xt(H(\zeta_i)X)\right] \leq E\left[g(\zeta_i)X^* t(H(\zeta_i)g(\zeta_i)X^*)\right],$$

where the expectation on the right is only over X^*. By assumption, $\mathbb{Q}(\zeta_i \in F$ eventually) is one and so (6) in Theorem 4 holds when

$$E\left[\sum_i g(\zeta_i)X^* t(H(\zeta_i)g(\zeta_i)X^*)\right] < \infty.$$

Now let η be the measure with atoms $g(\zeta_i)$ at $(g(\zeta_i)H(\zeta_i))^{-1}$ and note that

$$E\left[\sum_i g(\zeta_i)X^*t(H(\zeta_i)g(\zeta_i)X^*)\right] = E\int X^*t(X^*/y)\eta(dy).$$

Applying Lemma 7 completes the proof of (i).

In a similar way, considering the series in (7),

$$\begin{aligned}\sum_{i=1}^{\infty} E_{\zeta_i}[XI(H(\zeta_i)X \geq y)] &\geq \sum_{i=1}^{\infty} E_{\zeta_i}[XI(H(\zeta_i)X \geq y)]I(\zeta_i \in F)\\ &\geq \sum_{i=1}^{\infty} E[g(\zeta_i)X_*I(H(\zeta_i)g(\zeta_i)X_* \geq y)]I(\zeta_i \in F)\\ &= E[X_*A(X_*/y)],\end{aligned}$$

giving (ii). □

Corollary 2 *The positive increasing function L is slowly varying at infinity and $\delta \in [0,1)$; L and δ may be different in (i) and (ii).*

(i) Suppose, in addition to the conditions of Theorem 5(i), that,

$$\sup_{x>0}\{A(x)/(x^\delta L(x))\} < \infty, \quad \mathbb{Q} \text{ almost surely.}$$

If $E[(X^*)^{1+\delta}L(X^*)] < \infty$ then $E_{\mathbb{P}}W = h(0)$.

(ii) Suppose, in addition to conditions of Theorem 5(ii), that, for some y,

$$\inf_{x>y}\{A(x)/(x^\delta L(x))\} > 0, \quad \mathbb{Q} \text{ almost surely.}$$

If $E[(X_*)^{1+\delta}L(X_*)] = \infty$ then $E_{\mathbb{P}}W = 0$.

Proof. Suppose $A(x)/(x^\delta L(x))$ is bounded above by C. Then

$$\begin{aligned}\int_1^\infty \frac{A(wx)}{w^2}dw &\leq C\int_1^\infty \frac{(wx)^\delta L(wx)}{w^2}dw\\ &= Cx^\delta L(x)\int_1^\infty \frac{w^\delta}{w^2}\frac{L(wx)}{L(x)}dw,\end{aligned}$$

and, using the representation theorem for slowly varying functions, for suitably small ϵ and then sufficiently large x

$$\int_1^\infty \frac{w^\delta}{w^2}\frac{L(wx)}{L(x)}dw \leq \int_1^\infty \frac{w^\delta}{w^2}(1+\epsilon)w^\epsilon dw = (1+\epsilon)(1-\delta-\epsilon)^{-1}.$$

Applying these bounds in (8) proves (i). For (ii), note first that, since $L(xw)/L(x) \to 1$ as $x \to \infty$, $E[(X_*)^{1+\delta}L(X_*w)]$ is infinite when $E[(X_*)^{1+\delta}L(X_*)]$ is. Suppose $A(x)/(x^\delta L(x))$ is bounded below by $C > 0$ for $x \geq y$; then

$$E[X_*A(X_*w)] \geq Cw^\delta E[(X_*)^{1+\delta}L(X_*w)I(X_*w \geq y)],$$

which is infinite when $E[(X_*)^{1+\delta}L(X_*w)]$ is. Now apply (9). $\square$

The results with g the identity suffice for most purposes. However, for our main example the natural bounds on the reproduction take a more complicated form. In it, there are random variables X_1^* and X_2^* such that, for some positive function g on $\mathcal{S}$,

$$P_s(X > x) \leq P(X_1^* + g(s)X_2^* > x). \tag{10}$$

The next lemma implies that, when this happens, moment conditions for $E_{\mathbb{P}}W = h(0)$ can be derived separately for X_1^* and X_2^*. The fact that g multiplies X_2^* but not X_1^* will lead to the moment conditions on the two being different.

Lemma 8 *Suppose (10) holds. Then, for any $h > 0$,*

$$E_s[Xt(hX)] \leq 2E[X_1^* t(hX_1^*)] + 2E[g(s)X_2^* t(hg(s)X_2^*)].$$

Proof. This follows directly from the inequality $(x+y)t(h(x+y)) \leq 2xt(hx) + 2yt(hy)$. $\square$

4 Examples

Varying environment

Let $\mathcal{S} = \{\partial, 0, 1, 2, \ldots\}$ and let a person of type i give birth only to children of type $i+1$ and ∂. Assume the initial ancestor is of type 0. Let N_i be the (generic) family size for a person of type i, that is the number of children of type $i+1$, and let $EN_i = m_i$. Then

$$H(n) = \prod_{i=0}^{n-1} \frac{1}{m_i}, \quad H(0) = 1, \quad H(\partial) = 0$$

is mean-harmonic and the corresponding martingale is Z_n/EZ_n, where Z_n is the number of nth generation people. In this example $H(\zeta_n) = H(n)$, which is not random, and the distributions of X' and X depend only on n. Hence (5), (6) and the sum in (7) are all deterministic.

A routine application of Theorem 4(i) gives the following result.

Corollary 3 *If*

$$\liminf_n \sum_{i=0}^{n-1} (-\log m_i) < \infty$$

and, with $N_i' = N_i - 1$,

$$\sum_{i=1}^{\infty} E\left[\frac{N_i}{m_i} t\left(H(i)\frac{N_i'}{m_i}\right)\right] < \infty$$

then $E_{\mathbb{P}}W = 1$.

It is worth noting that, since $N_i' < N_i$ and, for a suitable $K > 0$,

$$t(y) = I(y \geq 1) + yI(y < 1) \leq K\left(1 - \frac{1-e^{-y}}{y}\right),$$

this result contains the main result, Theorem 5, of Goettge (1975).

To illustrate the use of bounding variables, Corollary 2 yields the following result for varying environment process. The special case when X^* and X_* are the same and $\{n(1), n(2), \ldots\} = \mathbb{N}$ leads to a classical Kesten-Stigum result, as was indicated in the Introduction.

Corollary 4

(i) Suppose that there is a random variable X^* such that

$$P(N_n/m_n > x) \leq P(X^* > x)$$

and that

$$\liminf \frac{\sum_{i=0}^{n-1} \log m_i}{n} > 0.$$

If $E[X^* \log^+(X^*)]$ is finite then $E_{\mathbb{P}} W = 1$.

(ii) Suppose that there is a random variable X_* and positive integers $\{n(1), n(2), \ldots\}$ with

$$P(N_n/m_n > x) \geq P(X_* > x) \quad \text{for } n \in \{n(1), n(2), \ldots\}$$

and

$$\limsup_j \frac{\sum_{i=0}^{n(j)-1} \log m_i}{j} < \infty.$$

If $E[X_* \log^+(X_*)]$ is infinite then $E_{\mathbb{P}} W = 0$.

Proof. This is an application of Corollary 2. For the first part, $A(x) = \sum_n I(H(n)x \geq 1)$ and the condition on the means translates to the restriction that for some $C > 0$ and $a > 0$, $H(n) < Ce^{-an}$ for all n; hence, $A(x)/\log(x+2)$ is bounded above. For the second part, taking $F = \{n(1), n(2), \ldots\}$ gives $A(x) = \sum_j I(H(n(j))x \geq 1)$; now the condition on the means yields, for some $C > 0$ and $a > 0$, $H(n(j)) > Ce^{-aj}$ for all j, which implies that $\inf_{x>y} A(x)/\log x$ is positive for y large enough. □

The growth conditions on the means here produce exponential decay rates for $H(n)$, leading to the $X \log X$ conditions. Other growth assumptions on the means will yield alternative results; an observation already made, in his notation, by Goettge (1975).

Before leaving this example, it is worth recalling that there are examples of varying environment processes where $E_{\mathbb{P}} W = 1$ but $\mathbb{P}(Z_n \to \infty, W = 0) > 0$, so that the survival set is larger than the set where W is positive. Relevant examples can be seen in MacPhee and Schuh (1983) and D'Souza and Biggins (1992).

Homogeneous, finite type space

Consider an irreducible, homogeneous, multitype, supercritical Galton-Watson process with the finite type space $\{1, 2, \ldots, p\}$, which, with the minor extra assumption of positive regularity, is the framework in the original Kesten-Stigum theorem (1966). Let $\mathcal{S} = \mathbb{Z}^+ \times \{\partial, 1, 2, \ldots, p\}$, where the first component of $\mathcal{S}$ tracks the generation. If ρ is the

Perron-Frobenius eigenvalue of the mean matrix and v the corresponding strictly positive right eigenvector then

$$H(n,j) = \rho^{-n} v_j, \quad H(n,\partial) = 0$$

is mean-harmonic and W_n is a weighted sum of the numbers in the n generation, with type j having weight $\rho^{-n} v_j$. Corollary 2 translates to one part of the multitype Kesten-Stigum theorem. To see this, note first that the second component of $\zeta_n = S(\xi_n)$ forms an irreducible Markov chain on $\{1, 2, \ldots, p\}$. The sum of all the offspring variables can be used for X^* and any component of any one of them for X_*. In either case the associated $A(x)$ looks like $\log x$; in the lower bound, this is a consequence of the chosen type having a finite mean recurrence time under the chain on $\{1, 2, \ldots, p\}$.

The full multitype Kesten-Stigum theorem considers the convergence of the vector formed by the numbers of each type, not just a particular weighted sum of the components. The best way to get from one to the other, in this model and more complex ones, is by establishing (by law of large number arguments) the stabilization of the proportion of each type; see, for example, Section V.6 of Athreya and Ney (1972). Kurtz et al. (1997) also discuss the multitype Galton-Watson process through the change of measure argument.

Reproduction depending on family history

In the language adopted here, Waymire and Williams (1996) allow reproduction at a node to depend on the reproduction of the node's ancestors. This can be accommodated easily by augmenting the type suitably. Let

$$S'(\nu) = \{S(\sigma) : \sigma = 0 \text{ or } \sigma \in c(\nu_i),\ i = 0, 1, \ldots, |\nu| - 1\},$$

so that $S'(\nu)$ contains all the information on the families of the ancestors of ν. Using S' as the new type allows reproduction to depend on the family history. There is an obvious consistency condition — the relevant part of a child's type must agree with the parent's. A very simple illustration is a one-type 'Galton-Watson' process in which a person's family size has a fixed mean m but a distribution that varies with the number of siblings that person has. Then Corollary 2(i) implies that the martingale Z_n/m^n will converge in mean when the various family size distributions are dominated by a distribution with a finite $x \log x$ moment.

Branching random walk in a random environment

In the original formulation, the basic data is a function from the type space $\mathcal{S}$ into probability laws on $\mathcal{S}^{\mathbb{N}}$, giving the family size distributions $\{P_s : s \in \mathcal{S}\}$. Denote the set of such functions by $\mathcal{L}$. In a sense, the collection of family size distributions, that is the element of $\mathcal{L}$ used, defines the external environment. Thus, a natural generalization is to allow some choice from $\mathcal{L}$; the varying environment process, already described, can be viewed in this way. Usually in a random environment the elements from $\mathcal{L}$ used in successive generations form a stationary sequence; here a branching random walk with a stationary environment sequence is considered.

First, the homogeneous branching random walk is described briefly. Let Z be a point process on the reals, with intensity measure μ. Branching random walk is a branching process with types in $\mathbb{R} \cup \partial$, corresponding to position. The point process describing the relative positions of the children of a person at s is distributed like Z. For a real θ, let $m(\theta) = \int e^{-\theta z}\mu(dz)$ and assume this is finite; assume also that $-m'(\theta)$, interpreted as $\int ze^{-\theta z}\mu(dz)$, exists. Augment the type space to include the generation; then

$$H(n,s) = e^{-\theta s}\prod_{i=0}^{n-1}\frac{1}{m(\theta)} = \frac{e^{-\theta s}}{m(\theta)^n}, \quad H(n,\partial) = 0$$

is mean-harmonic. Using Theorem 1, it is straightforward to check that $\zeta_n = (n, \sum_{i=0}^{n-1} Y_i)$, where $\{Y_i\}$ are independent identically distributed, with the law that has derivative $e^{-\theta x}/m(\theta)$ with respect to μ, so that $E[Y_i] = -m'(\theta)/m(\theta)$. Hence

$$-\log H(\zeta_n) = \sum_{i=0}^{n-1}(\theta Y_i + \log m(\theta)).$$

In the process with a random environment, the law for the point process Z varies; when that law is η let the corresponding expectation be E^η and let $m_\eta(\theta) = E^\eta\left[\int e^{-\theta x}Z(dx)\right]$. Let the law used in generation n be $\lambda(n)$, where $\lambda = \{\lambda(n)\}$ forms a stationary sequence with the marginal law P^*. Assume that $m_\eta(\theta)$ is finite and $m'_\eta(\theta)$ exists, P^* almost surely. Finally, denote the conditional branching law given λ by $\mathbb{P}$. It should really be something like $\mathbb{P}^\lambda$, but precision is sacrificed to simplicity.

Suppose the environment λ is given. Then, again augmenting the type space by the generation,

$$H(n,s) = e^{-\theta s}\prod_{i=0}^{n-1}\frac{1}{m_{\lambda(i)}(\theta)}, \quad H(n,\partial) = 0$$

is mean-harmonic for the branching process. Suppose the parent has reproduction law η, then the variable X becomes $X = (m_\eta(\theta))^{-1}\int e^{-\theta x}Z(dx)$, where Z has law η.

The next lemma is a straightforward application of definitions.

Lemma 9 *Given λ, $E_{(n,s)}[f(X)] = E^{\lambda(n)}[f(X)]$. Let $\overline{E}$ be the expectation over λ, then, by stationarity, $\overline{E}[E_{(n,s)}[f(X)]] = \int E^\eta[f(X)]P^*(d\eta)$.*

The following theorem extends some of the results in Biggins (1977) and Lyons (1997). When $\theta = 0$ it covers the Galton-Watson case and when the environment is fixed it covers the homogeneous branching random walk. It is worth stressing that, since $\mathbb{P}$ is a conditional law, the conclusions are conditional ones, holding almost surely as λ varies over realisations.

Theorem 6 *Assume that the environment λ is ergodic and that*

$$\kappa = \int\left(-\theta\frac{m'_\eta(\theta)}{m_\eta(\theta)} + \log m_\eta(\theta)\right)P^*(d\eta)$$

exists.

(i) If $\kappa < 0$ then $E_{\mathbb{P}} W = 0$.

(ii) If $\kappa > 0$ and $\int E^{\eta}[X \log X] P^*(d\eta) < \infty$ then $E_{\mathbb{P}} W = 1$.

(iii) If λ is a collection of independent identically distributed variables, then $E_{\mathbb{P}} W = 0$ when (a) $\kappa = 0$ or when (b) $0 < \kappa < \infty$ and $\int E^{\eta}[X \log X] P^*(d\eta) = \infty$.

Proof. Let the real random variable Y_η have the law with density $e^{-\theta x}/m_\eta(\theta)$ with respect to μ_η. Then, as before,

$$-\log H(\zeta_n) = \sum_{i=0}^{n-1} \left(\theta Y_{\lambda(i)} + \log m_{\lambda(i)}(\theta)\right),$$

where, given the $\lambda(i)$, the Y's are independent variables. Then $\{(\lambda(n), Y_{\lambda(n)})\}$ is stationary and, by careful use of the pointwise ergodic theorem,

$$\frac{-\log H(\zeta_n)}{n} \to \int \left(-\theta \frac{m'_\eta(\theta)}{m_\eta(\theta)} + \log m_\eta(\theta)\right) P^*(d\eta) = \kappa.$$

When κ is less than 0, Corollary 1 applies to show that W is zero, proving (i).

When κ is greater than zero, $H(\zeta_n)$ is eventually contained in an interval of the form $(0, d^n)$, with $d < 1$. Then (6) in Theorem 4 is finite when

$$\sum_{n=1}^{\infty} E_{\zeta_n}[X t(d^n X)] < \infty$$

and, by Lemma 9,

$$\overline{E}\left[\sum_{n=1}^{\infty} E_{\zeta_n}[X t(d^n X)]\right] = \int E^{\eta}\left[\sum_{n=1}^{\infty} X t(d^n X)\right] P^*(d\eta),$$

which is finite when $\int E^{\eta}[X \log X] P^*(d\eta) < \infty$. This proves (ii).

When $\kappa = 0$ and the $\{\lambda(i)\}$ are independent, $-\log H(\zeta_n)$ is a zero-mean random walk and so has its lim sup at infinity; thus, Corollary 1 again shows that W is zero.

When $0 < \kappa < \infty$, $H(\zeta_n)$ is ultimately contained in an interval of the form (d^n, ∞), with $d < 1$. Then, using Lemma 9, the series in (7) in Theorem 4 is infinite when

$$\sum_{n=1}^{\infty} E^{\lambda(n)}[X I(d^n X \geq y)] = \infty$$

and the terms here are bounded by one and are independent. Conditional Borel-Cantelli now shows this holds exactly when $\int E^{\eta}[X \log X] P^*(d\eta) = \infty$. $\square$

The example can be taken further, allowing reproduction to depend on the node, not just its generation. Each node ν has a law $\lambda(\nu)$ attached to it, with λ forming an ergodic sequence down every line of descent. Now, augmenting the type by the node,

$$H(\nu, s) = e^{-\theta s} \prod_{i=0}^{|\nu|-1} \frac{1}{m_{\lambda(\nu_i)}(\theta)}, \quad H(\nu, \partial) = 0$$

is mean harmonic, given λ, and the arguments leading to Theorem 6 continue to apply, but the martingale W_n may now be too complicated to be interesting.

This example is fairly simple because the mean-harmonic function factorises, with one factor depending on the original type space and the other depending only on the environment. Most multitype random environment branching processes do not have this property.

Multiplicative cascades

There is a direct correspondence between the branching random walk and what are called multiplicative cascades. To make this correspondence, the type space $\mathbb{R} \cup \partial$, used in the branching random walk, becomes $[0, \infty)$ by taking s to e^s, with the convention that $e^\partial = 0$. This transforms the addition of displacements along lines of descent, which define the branching random walk, into multiplications.

When the node ν has type ∂ let $A(\nu) = 0$; otherwise, let $z(\nu)$ be the displacement of ν from its parent and let $A(\nu) = e^{z(\nu)}$. Then $A = (A_1, A_2, \ldots)$ is usually called the generator of the cascade and $m(1) = E \int e^{-z} Z(dz) = E \sum_i A_i$.

Waymire and Williams (1996) consider such cascades on the b-ary tree with the b non-zero terms in A being conditionally independent given the family history of the parent and each having mean one. Then $m(1) = b^{-1}$. Augmenting the type space with the generation, so that it becomes $\mathbb{Z}^+ \times [0, \infty)$, the function $H(n, s) = s/b^n$ is mean harmonic. Now Proposition 2 here can be seen to be very closely related to Corollary 2.3 in Waymire and Williams (1996), with the basic measure change being their Theorem 2.3.

Other examples

There are several other examples that might also have been considered in detail. Combining the first two examples leads to the multitype Galton-Watson process in a varying environment, which was considered using other methods in Biggins, Cohn and Nerman (1999). The parts of that discussion which consider martingales arising from mean-harmonic functions, which are there just called harmonic, can certainly be tackled using the ideas developed here. Another example is the multitype branching random walk with a finite type space; Kyprianou and Rahimzadeh Sani (2001) discuss martingale convergence for this through the measure change argument. Olofsson (1998) uses the change of measure argument in the context of the general branching process; the next section, on optional lines, provides the theory to connect this model with the branching random walk already discussed. The connection is explained at the end of that section.

5 Optional lines

There are natural reasons to want to consider the sum of $h(\nu)$ over collections of nodes other than the nth generation ones; see, for example, Jagers (1989), Chauvin (1991), Biggins and Kyprianou (1997) and Kyprianou (2000). In particular, Jagers (1989) establishes the basic framework.

A function $\mathcal{L}$ from the nodes $\mathcal{T}$ to $\{0,1\}$ codes for membership of the set of nodes $\{\nu : \mathcal{L}(\nu) = 1\}$. This set, and the corresponding function $\mathcal{L}$, is called a line if no member of it is the ancestor of any other, so that $\mathcal{L}(\nu) = 1$ implies that $\mathcal{L}(\nu_i) = 0$ for all $i < |\nu|$. Lines in this sense cut across the tree, in (complete) contrast to lines of descent; however, a line does not have to include a node from every line of descent, so it does not have to cut all branches from the root. Although, formally, a line $\mathcal{L}$ is a function on the nodes it will often be convenient to identify $\mathcal{L}$ with set of nodes where the function takes the value one.

Informally, an optional line, $\mathcal{L}$, is a random line with the property that its position is determined by the history of the process up to the line. Its σ-algebra, $\mathcal{G}_{\mathcal{L}}$, is the information on the reproduction of all individuals that are neither on the line nor a descendent of any member of the line. Jagers (1989) shows that the branching property, which is that, given $\mathcal{G}_n$, different individuals in generation n give rise to independent copies of the original tree, extends to any optional line.

Unfortunately, optional lines seem to be too general for some of the results sought here, necessitating some restriction. Specifically, the optional line $\mathcal{L}$ will be called simple when, for all ν, the function $\mathcal{L}(\nu)$ is measurable with respect to $\mathcal{G}_{|\nu|}$ (not $\mathcal{F}_{|\nu|}$); thus, whether ν is on the line or not is determined by looking at the process up to generation $|\nu|$, ignoring the trunk.

For the line $\mathcal{L}$ let

$$W_{\mathcal{L}} = \sum_{\nu \in \mathcal{T}} \mathcal{L}(\nu) H(S(\nu)) = \sum_{\nu \in \mathcal{T}} \mathcal{L}(\nu) h(\nu).$$

Clearly W_n is $W_{\mathcal{L}}$ when $\mathcal{L}$ is the (non-random) line formed by all nth generation nodes. Lines can be ordered through the natural ordering of the nodes; so, $\mathcal{L}^{(1)} \leq \mathcal{L}^{(2)}$ if every node on the line $\mathcal{L}^{(2)}$ is on, or has an ancestor on, the line $\mathcal{L}^{(1)}$. An important question is when $W_{\mathcal{L}}$ defines a martingale as $\mathcal{L}$ varies through some increasing collection of optional lines. The next lemma provides the key to this property, which relies on expectation being preserved. Define N by $\mathcal{L}(\xi_N) = 1$, with $N = \infty$ when there is no such N; thus, N is the generation in which ξ hits the line. Often it will be easy to see when N is finite under $\mathbb{Q}$.

Lemma 10 *When $\mathcal{L}$ is a simple optional line, $\mathbb{Q}(N < \infty) = E_{\mathbb{P}} W_{\mathcal{L}} / h(0)$, and so $E_{\mathbb{P}}[W_{\mathcal{L}}] = h(0)$ if and only if $\mathbb{Q}(N < \infty) = 1$.*

Proof. The steps in the next calculation are justified by: conditioning on $\mathcal{F}_n$ and using Lemma 2 to move from $E_{\mathbb{Q}}$ to $E_{\mathbb{P}}$; conditioning on $\mathcal{G}_n$ and using that $\mathcal{L}$ is a simple optional line; and, finally, using Lemma 3.

$$\mathbb{Q}(N = n) \quad = \quad E_{\mathbb{Q}} \left[\sum_{|\nu| = n} \mathcal{L}(\nu) I(\xi_n = \nu) \right]$$

$$
\begin{aligned}
&= \sum_{|\nu|=n} E_{\mathbb{P}}\left[\mathcal{L}(\nu)\overline{W}(\nu)I(\xi_n=\nu)\right] \\
&= \sum_{|\nu|=n} E_{\mathbb{P}}\left[\mathcal{L}(\nu)\overline{W}(\nu)\mathbb{P}(\xi_n=\nu|\mathcal{G}_n)\right] \\
&= \sum_{|\nu|=n} E_{\mathbb{P}}\left[\mathcal{L}(\nu)\frac{h(\nu)}{h(0)}\right].
\end{aligned}
$$

Summing over n now gives the result. □

Let $\mathcal{A}_\nu$ be the σ-algebra generated by $\{S(\nu_i) : i = 0, 1, \ldots, |\nu|\}$. Then it is reasonable to call an optional line very simple when, for all ν, the function $\mathcal{L}(\nu)$ is measurable with respect to $\mathcal{A}_\nu$; then, whether ν is on the line or not is determined by looking at the types in its ancestry. For very simple optional lines, N is a stopping time for the Markov chain $\zeta = \{S(\xi_0), S(\xi_1), \ldots\}$. The main application is of this form, but there is no advantage to the extra restriction in developing the theory. It is worth noting that the definition of an optional line used in Kyprianou (2000) to prove a particular case of Lemma 10 is between simple and very simple.

Once $W_{\mathcal{L}}$ defines a martingale as $\mathcal{L}$ varies, it is natural to wonder whether its limit is the same as that of the martingale W_n. Let $\mathcal{L}_n$ be the (function corresponding to the) line formed by members of $\mathcal{L}$ in the first n generations and the nth generation nodes with no ancestor in $\mathcal{L}$.

Lemma 11 *For any (not necessarily simple) optional line* $\mathcal{L}$, $E_{\mathbb{P}}[W_n|\mathcal{G}_{\mathcal{L}}] = W_{\mathcal{L}_n}$.

Proof. Recall that $W_r(S^\nu)$ is W_r defined on the sub-tree rooted at ν. Then

$$W_n = \sum_{|\nu|\le n} \mathcal{L}_n(\nu) W_{n-|\nu|}(S^\nu)$$

and, when $\mathcal{L}_n(\nu) = 1$, $E_{\mathbb{P}}\left[W_{n-|\nu|}(S^\nu)\middle|\, \mathcal{G}_{\mathcal{L}}\right] = h(\nu)$. Hence

$$
\begin{aligned}
E_{\mathbb{P}}[W_n|\mathcal{G}_{\mathcal{L}}] &= E_{\mathbb{P}}\left[\sum_{|\nu|\le n} \mathcal{L}_n(\nu) W_{n-|\nu|}(S^\nu)\middle|\, \mathcal{G}_{\mathcal{L}}\right] \\
&= \sum_{|\nu|\le n} \mathcal{L}_n(\nu) E_{\mathbb{P}}\left[W_{n-|\nu|}(S^\nu)\middle|\, \mathcal{G}_{\mathcal{L}}\right] \\
&= \sum_{|\nu|\le n} \mathcal{L}_n(\nu) h(\nu) = W_{\mathcal{L}_n}.
\end{aligned}
$$

□

In general $\mathcal{L}_n$ need not be optional when $\mathcal{L}$ is and so $W_{\mathcal{L}_n}$ need not be $\mathcal{G}_n$-measurable. However, for simple optional lines it is, and then, as the next two lemmas show, much more can be said.

Lemma 12 *Let* $\mathcal{L}$ *be a simple optional line. Then* $\mathcal{L}_n$ *is a simple optional line and* $(W_{\mathcal{L}_n}, \mathcal{G}_n)$ *is a positive martingale with a limit at least* $W_{\mathcal{L}}$. *When* $E_{\mathbb{P}}[W_{\mathcal{L}}] = h(0)$, *(i)* $W_{\mathcal{L}_n} = E_{\mathbb{P}}[W_{\mathcal{L}}|\mathcal{G}_n]$, *(ii) the martingale converges in mean to* $W_{\mathcal{L}}$, *and (iii)* $E_{\mathbb{P}}[W_n|\mathcal{G}_{\mathcal{L}}] = E_{\mathbb{P}}[W_{\mathcal{L}}|\mathcal{G}_n]$.

Proof. It is immediate from the definitions that $\mathfrak{L}_n$ is a simple optional line. Hence $W_{\mathfrak{L}_n}$ is $\mathcal{G}_n$-measurable. Let $A_{\mathfrak{L},n}$ be the line formed by members of the nth generation that are neither in $\mathfrak{L}$ nor have an ancestor in $\mathfrak{L}$, so that (with the ancestry of ν being $\{\nu_i : i = 0, 1, \ldots, |\nu|\}$)

$$A_{\mathfrak{L},n}(\nu) = I(|\nu| = n) \prod_{i=0}^{n} (1 - \mathfrak{L}(\nu_i)).$$

Then $A_{\mathfrak{L},n}$ is a simple optional line when $\mathfrak{L}$ is a simple optional line. By definition,

$$W_{\mathfrak{L}_{n+1}} = \sum_{|\nu| \leq n} \left(\mathfrak{L}(\nu) h(\nu) + A_{\mathfrak{L},n}(\nu) \sum_{\sigma \in c(\nu)} h(\sigma) \right)$$

Now, when $|\nu| = n$,

$$E_{\mathbb{P}} \left[\sum_{\sigma \in c(\nu)} h(\sigma) \middle| \mathcal{G}_n \right] = h(\nu)$$

and, because $\mathfrak{L}$ is simple, everything else in the expression for $W_{\mathfrak{L}_{n+1}}$ is $\mathcal{G}_n$-measurable. Thus,

$$E_{\mathbb{P}} \left[W_{\mathfrak{L}_{n+1}} | \mathcal{G}_n \right] = \sum_{|\nu| \leq n} \left(\mathfrak{L}(\nu) + A_{\mathfrak{L},n}(\nu) \right) h(\nu) = W_{\mathfrak{L}_n},$$

and so is a martingale, and $\lim_n W_{\mathfrak{L}_n} \geq W_{\mathfrak{L}}$. Hence $E_{\mathbb{P}}[W_{\mathfrak{L}_n}] = E_{\mathbb{P}}[W_{\mathfrak{L}_0}] = h(0)$ and

$$W_{\mathfrak{L}_n} = \lim_{m \to \infty} E_{\mathbb{P}} [W_{\mathfrak{L}_m} | \mathcal{G}_n] \geq E_{\mathbb{P}} [W_{\mathfrak{L}} | \mathcal{G}_n];$$

$E_{\mathbb{P}}[W_{\mathfrak{L}}] = h(0)$ forces equality here, which in turn implies that $W_{\mathfrak{L}_n}$ converges to $W_{\mathfrak{L}}$. Hence, $W_{\mathfrak{L}_n} = E_{\mathbb{P}}[W_{\mathfrak{L}} | \mathcal{G}_n]$ and, by Lemma 11, $W_{\mathfrak{L}_n} = E_{\mathbb{P}}[W_n | \mathcal{G}_{\mathfrak{L}}]$, proving (iii). ☐

Theorem 6.7 of Jagers (1989) gives similar conclusions to the next lemma, but for general optional lines.

Lemma 13 *Let $\mathfrak{L}'$ and $\mathfrak{L}$ be simple optional lines with $\mathfrak{L}' \leq \mathfrak{L}$ and $E_{\mathbb{P}}[W_{\mathfrak{L}}] = h(0)$. Then $E_{\mathbb{P}}[W_{\mathfrak{L}} | \mathcal{G}_{\mathfrak{L}'}] = W_{\mathfrak{L}'}$.*

Proof. Let N' and N be the generations where ξ hits $\mathfrak{L}'$ and $\mathfrak{L}$ respectively. Then $N' \leq N$ and so, by Lemma 10, $E_{\mathbb{P}}[W_{\mathfrak{L}}] = h(0)$ implies that $E_{\mathbb{P}}[W_{\mathfrak{L}'}] = h(0)$. Since $\mathfrak{L}' \leq \mathfrak{L}$, $\mathcal{G}_{\mathfrak{L}'} \subset \mathcal{G}_{\mathfrak{L}}$ and so, by Lemma 11,

$$W_{\mathfrak{L}'_n} = E_{\mathbb{P}} [W_n | \mathcal{G}_{\mathfrak{L}'}] = E_{\mathbb{P}} [E_{\mathbb{P}} [W_n | \mathcal{G}_{\mathfrak{L}}] | \mathcal{G}_{\mathfrak{L}'}] = E_{\mathbb{P}} [W_{\mathfrak{L}_n} | \mathcal{G}_{\mathfrak{L}'}].$$

Letting n go to infinity and applying Lemma 12(ii) completes the proof. ☐

The final lemma in this sequence gives a fairly simple necessary condition for one of the hypotheses of the main theorem to hold.

Lemma 14 *Let $\{\mathfrak{L}[t] : t \geq 0\}$ be (not necessarily simple) optional lines that are increasing with t. If $(h(\nu)\mathfrak{L}[t](\nu)) \to 0$ as $t \to \infty$ for each ν with $|\nu| \leq (n-1)$ then $W_{\mathfrak{L}[t]_n} \to W_n$.*

Proof. Let $\mathcal{G}_*$ be the σ-algebra generated by $\{\mathcal{G}_{\mathcal{L}[t]} : t \geq 0\}$. Lemma 11 implies that $(W_{\mathcal{L}[t]_n}, \mathcal{G}_{\mathcal{L}[t]})$ is a positive martingale, and so $W_{\mathcal{L}[t]_n}$ converges as $t \to \infty$, to $E_{\mathbb{P}}[W_n|\mathcal{G}_*]$. Now

$$W_{\mathcal{L}[t]_n} \leq \left(W_n + \sum_{|\nu| \leq n-1} \mathcal{L}[t](\nu)h(\nu) \right)$$

and $\sum_{|\nu|\leq n-1} h(\nu) = \sum_{i=0}^{n-1} W_i$ which is finite. Hence, letting $t \to \infty$ and using dominated convergence,

$$E_{\mathbb{P}}[W_n|\mathcal{G}_*] = \lim_{t\to\infty} W_{\mathcal{L}[t]_n} \leq W_n$$

which implies that $E_{\mathbb{P}}[W_n|\mathcal{G}_*] = W_n$, as required. □

Recall that W is the almost sure limit of the martingale $(W_n, \mathcal{G}_n)$. Obviously, in the next theorem, Lemma 10 is useful for checking that $E_{\mathbb{P}}\left[W_{\mathcal{L}[t]}\right] = h(0)$.

Theorem 7 *Let $\{\mathcal{L}[t] : t \geq 0\}$ be simple optional lines that are increasing with t and satisfy $E_{\mathbb{P}}\left[W_{\mathcal{L}[t]}\right] = h(0)$ for every t. Then $(W_{\mathcal{L}[t]}, \mathcal{G}_{\mathcal{L}[t]})$ is a positive martingale. If, for each n, $W_{\mathcal{L}[t]_n}$ tends to W_n, almost surely, as $t \to \infty$ then $W_{\mathcal{L}[t]}$ converges to W almost surely.*

Proof. The martingale property follows immediately from Lemma 13. Let W' be the limit of $W_{\mathcal{L}[t]}$. By Lemma 12(i), $E_{\mathbb{P}}\left[W_{\mathcal{L}[t]}|\mathcal{G}_n\right] = W_{\mathcal{L}[t]_n}$; letting $t \to \infty$, Fatou gives $E_{\mathbb{P}}[W'|\mathcal{G}_n] \leq W_n$ and then letting $n \to \infty$ gives $W' \leq W$. Again, let $\mathcal{G}_*$ be the σ-algebra generated by $\{\mathcal{G}_{\mathcal{L}[t]} : t \geq 0\}$. By Lemma 12(iii), $E_{\mathbb{P}}\left[W_{\mathcal{L}[t]}|\mathcal{G}_n\right] = E_{\mathbb{P}}\left[W_n|\mathcal{G}_{\mathcal{L}[t]}\right]$; letting n and then t go to infinity shows that $W' \geq E_{\mathbb{P}}[W|\mathcal{G}_*]$. Hence $E_{\mathbb{P}}[W' - W] \geq 0$, but $W' \leq W$. Hence $W' = W$, completing the proof. □

The results are easier to establish when the original martingale $\{W_n\}$ converges in mean. This can be seen from the next theorem, which we will not need.

Theorem 8 *Suppose W_n converges in mean to W. Let $\{\mathcal{L}[t] : t \geq 0\}$ be optional lines that are increasing with t and satisfy $E_{\mathbb{P}}\left[W_{\mathcal{L}[t]}\right] = h(0)$ for every t. Then $(W_{\mathcal{L}[t]}, \mathcal{G}_{\mathcal{L}[t]})$ is a positive martingale and $W_{\mathcal{L}[t]}$ converges in mean to W.*

Proof. Let $n \to \infty$ in Lemma 11 to show that $E_{\mathbb{P}}[W|\mathcal{G}_{\mathcal{L}[t]}] \geq W_{\mathcal{L}[t]}$; both sides have expectation $h(0)$, forcing equality. Standard martingale theory now gives the results. □

An example: first crossings in the branching random walk

The remainder of this section concerns a homogeneous branching random walk, as described in Section 4, with $E\int e^{-\alpha t} Z(dt) = 1$ for some $\alpha > 0$ and $\beta = E\int te^{-\alpha t}Z(dt) > 0$; then $e^{-\alpha s}$ is mean-harmonic, giving the martingale $W_n = \sum_\nu e^{-\alpha S(\nu)} I(|\nu| = n, S(\nu) \neq \partial)$ with associated martingale limit W. Let

$$\mathcal{C}[t](\nu) = I(S(\nu) > t \text{ but } S(\nu_i) \leq t \text{ for } i < |\nu|),$$

which is the line formed by picking out individuals born to the right of t but with all their antecedents born to the left t. Now let $W_{\mathcal{C}[t]} = \sum_\nu e^{-\alpha S(\nu)} I(\nu \in \mathcal{C}[t])$.

Proposition 3 *$W_{\mathcal{C}[t]}$ is a martingale converging to W.*

Proof. The $\mathcal{C}[t]$ are simple optional lines that are increasing in t. Under $\mathbb{Q}$, ζ is a random walk with mean $\beta > 0$; hence, ζ is certain to cross $\mathcal{C}[t]$ and so Lemma 10 applies. Furthermore, $e^{-\alpha S(\nu)} I(S(\nu) > t) \to 0$ as $t \to \infty$ and so Lemma 14 applies. Hence Theorem 7 applies to show that $W_{\mathcal{C}[t]}$ is a martingale converging to W. □

The martingale $W_{\mathcal{C}[t]}$ occurs in the general branching process. Let $X = \int e^{-\alpha t} Z(dt)$. Theorem 6 gives $E_{\mathbb{P}} W = 1$ if $E[X \log X] < \infty$ and $E_{\mathbb{P}} W = 0$ if both $\beta < \infty$ and $E[X \log X] = \infty$. When combined with Proposition 3, this includes the conclusion of Theorem 2.1 in Olofsson (1998), which deals with the case where Z is concentrated on $(0, \infty)$ and $\beta < \infty$; that paper should be consulted for references to earlier treatments of this result and more context on the general branching process.

6 Branching random walk with a barrier

Let $Z = \sum_i \delta(z_i)$ be a point process on the reals, with points at $\{z_i\}$. As indicated in Section 4, a homogeneous branching random walk based on Z has types in $(-\infty, \infty)$ generated in the obvious way. The homogeneous process is now modified by the removal of lines of descent from the point where they cross into $(-\infty, 0]$, to give a process with an absorbing barrier. This construction couples the process with a barrier, which is the topic of this section, to the homogeneous one. This kind of process has been considered before; see Kesten (1978) and Biggins et al. (1991).

Formally, the branching process considered in this section has the type space, corresponding to position, $\mathcal{S} = [0, \infty)$ and has the point process describing the positions of the children of a person at s distributed like

$$\sum_i \delta(s + z_i) I(s + z_i > 0);$$

thus, the relative positions are distributed like Z, except that children with positions in $(-\infty, 0]$ do not appear. The 'ghost' state ∂ and the associated details are omitted from the discussion; hence, sums over $|\nu| = n$ should now be interpreted as being over the nodes that occur, that is those that do not have type ∂.

Let the intensity measure of Z be μ and assume that

$$\int e^{-x} \mu(dx) = 1 \quad \text{and} \quad \int x e^{-x} \mu(dx) = 0. \tag{11}$$

It turns out that this provides the most interesting case; a claim not justified here. The results can be transferred easily to any Z that takes the required form after its points have been scaled by a non-zero θ.

Let Y_n be independent identically distributed variables with their law having density e^{-x} with respect to μ and let S_n be the random walk with increments $\{Y_n\}$. For $x > 0$, let $V(x)$ be the expected number of visits S_n makes to $(-x, 0]$ before first hitting $(0, \infty)$, and let $V(0) = 1$. Some results from random walk theory are important for the motivation and the formulation; these are recorded in the following lemma. The first two parts are consequences of V being, essentially, the renewal function for the weak descending ladder height process of $\{S_n\}$; the final part is Lemma 1 of Tanaka (1989). Similar results can be found in Bertoin and Doney (1994). The relevant material is reviewed at the start of Biggins (2001), which develops the random walk results needed to prove Theorem 10 stated at the end of this section.

Lemma 15 *(i) As $x \to \infty$, $V(x)/x$ converges to a positive constant, which is finite if $\int x^2 e^{-x}\mu(dx) < \infty$. (ii) When $V(x)/x$ has a finite limit, $a(x+1) \leq V(x) \leq b(x+1)$for suitable $a > 0$ and $b < \infty$. (iii) $E[V(Y_1+s)I(Y_1+s>0)] = V(s)$.*

Lemma 16 *When (11) holds, $H(s) = V(s)e^{-s}$ is mean-harmonic for the branching random walk with a barrier.*

Proof. For any non-negative g,

$$\begin{aligned} E_s\left[\sum_{|\nu|=1} g(S(\nu))\right] &= E\left[\sum_i g(s+z_i)I(s+z_i>0)\right] \\ &= \int g(z+s)I(s+z>0)\mu(dz). \end{aligned}$$

Hence, using Lemma 15(iii) for the final equality,

$$\begin{aligned} E_s\left[\sum_{|\nu|=1} V(S(\nu))e^{-S(\nu)}\right] &= \int V(z+s)e^{-z-s}I(s+z>0)\mu(dz) \\ &= e^{-s}E[V(Y_1+s)I(Y_1+s>0)] = e^{-s}V(s), \end{aligned}$$

as required. □

The martingale now being studied is

$$W_n = \sum_{|\nu|=n} V(S(\nu))e^{-S(\nu)},$$

with its limit being W; the behaviour of the trunk under the measure change this martingale induces is described next.

Lemma 17 *Under $\mathbb{Q}$, the Markov development of the types on the trunk, ζ, has the transitions from s given by the law*

$$\frac{V(z+s)}{V(z)}I(z+s>0)e^{-z}\mu(dz).$$

Proof. Substitute for H and the reproduction process in Theorem 1. □

This transition mechanism, which has arisen previously, in, for example, Tanaka (1989) and Bertoin and Doney (1994), can reasonably be called a random walk conditioned to stay positive for reasons explained in Bertoin (1993). In particular, Tanaka (1989) gives a sample path construction of the process that can be used to give rather precise information on the long term behaviour of ζ, which will be described in the next section, but first the following simple consequence of his results is recorded.

Lemma 18 *Under* $\mathbb{Q}$, $\zeta_n \to \infty$ *as* $n \to \infty$.

This is enough for the application of the ideas on optional lines. In the same way as at the end of Section 5, $\mathcal{C}[t]$ is the line formed by picking out individuals born to the right of t but with all their antecedents born to the left of t. These lines are particularly useful since they focus attention on nodes near t, regardless of generation, thereby making sums over them comparatively well behaved. As before, $\mathcal{C}[t]$ are increasing very simple optional lines, and so in Lemma 10, N is the first time ζ exceeds t; Lemma 18 now shows that $\mathbb{Q}(N < \infty)$. Furthermore, $V(S(\nu))e^{-S(\nu)}I(S(\nu) > t) \to 0$ as $t \to \infty$ and so Lemma 14 applies. Hence Theorem 7 applies to give the following result.

Theorem 9 *The martingale* $(W_{\mathcal{C}[t]}, \mathcal{G}_{\mathcal{C}[t]})$ *converges to* W*, which is the limit of the martingale* $(W_n, \mathcal{G}_n)$.

Obviously mean convergence of W_n will require moment conditions. To state these, let

$$\tilde{X}_1 = \sum_j z_j e^{-z_j} I(z_j > 0),$$

$$\tilde{X}_2 = \sum_j e^{-z_j} \quad \text{and} \quad \tilde{X}_3(s) = \sum_j e^{-z_j} I(z_j > -s).$$

Note that $\tilde{X}_3(s) \uparrow \tilde{X}_2$ as $s \uparrow \infty$. The next section is devoted to the proof of the following result.

Theorem 10 *Assume* $\int x^2 e^{-x} \mu(dx) < \infty$.
Let $\phi(x) = \log\log\log x$, $L_1(x) = (\log x)\phi(x)$, $L_2(x) = (\log x)^2\phi(x)$, $L_3(x) = (\log x)/\phi(x)$ *and* $L_4(x) = (\log x)^2/\phi(x)$.

(i) If both $E[\tilde{X}_1 L_1(\tilde{X}_1)]$ *and* $E[\tilde{X}_2 L_2(\tilde{X}_2)]$ *are finite then* W_n *converges in mean.*

(ii) If $E[\tilde{X}_1 L_3(\tilde{X}_1)]$ *is infinite or, for some* s*, or* $E[\tilde{X}_3(s) L_4(\tilde{X}_3(s))]$ *is infinite then* $W_n \to 0$ *almost surely.*

There is a (small) gap between the slowly varying functions used in the two parts. There are some grounds for thinking this gap is inevitable, reflecting oscillations in the trunk. The gap in the random variables, between using $\tilde{X}_2$ in the first part and $\tilde{X}_3(s)$ in the second, arises from the upper and lower bounds on the reproduction having slightly different forms.

7 Proof of Theorem 10

The proof is an application of Corollary 2, Lemma 8 and the following results. First, the simple Lemma 18 needs to be supplemented by information on how fast ζ goes to infinity; the following result, taken from Biggins (2001), provides relevant estimates.

Theorem 11 *Let* $D(x) = \sum_n I(\zeta_n \le x)$ *and let* $\phi(x) = \log\log x$ *for* $x > 3$*. For suitable (non-random)* L *and* U

$$\limsup_{x\to\infty} \frac{D(x)}{x^2\phi(x)} \le U < \infty \quad \textit{and} \quad \liminf_{x\to\infty} \frac{D(x)}{x^2/\phi(x)} \ge L > 0,$$

almost surely.

One consequence of this, or Lemma 18, is that in applying the second parts of Theorem 5 and Corollary 2 it will be enough to consider the reproduction far above the barrier, that is, $F \subseteq \mathcal{S}$ in those results can be taken as $[s, \infty)$ for any large s. The next lemma is also a simple application of Theorem 11, providing another relevant estimate. It would be easy to prove more, replacing V by a more general function, but the result will only be needed for this case.

Lemma 19 *Let* $\tilde{D}(x) = \sum_n V(\zeta_n)^{-1} I(\zeta_n \le x) = \int_0^x V(z)^{-1} D(dz)$*. For suitable (non-random)* $\tilde{L}$ *and* $\tilde{U}$

$$\limsup_{x\to\infty} \frac{\tilde{D}(x)}{x\phi(x)} \le \tilde{U} < \infty \quad \textit{and} \quad \liminf_{x\to\infty} \frac{\tilde{D}(x)}{x/\phi(x)} \ge \tilde{L} > 0.$$

Proof. Let $D_*(x) = (L-\epsilon)x^2/\phi(x)$ for $x > x_0 > e^e$ and $D_* = 0$ otherwise, with x_0, which is random, large enough that $D(x) \ge D_*(x)$ for all $x \ge 0$. Then, using Lemma 15(i) and Fubini,

$$\begin{aligned}
\tilde{D}(x) &= \int_0^x V(z)^{-1} D(dz) \\
&\ge \int_0^x \frac{1}{b(z+1)} D(dz) \\
&= \frac{1}{b}\int_0^x \left(\int_z^\infty \frac{dy}{(y+1)^2}\right) D(dz) \\
&= \frac{1}{b}\int_0^\infty D(\min\{x,y\}) \frac{dy}{(y+1)^2} \\
&\ge \frac{1}{b}\int_0^\infty D_*(\min\{x,y\}) \frac{dy}{(y+1)^2} \\
&= \frac{1}{b}\int_0^x \frac{1}{z+1} D_*(dz) \\
&= \frac{L-\epsilon}{b}\int_{x_0}^x \frac{1}{z+1}\left(\frac{2z}{\log\log z} - \frac{z}{\log z(\log\log z)^2}\right) dz.
\end{aligned}$$

Estimating the integral for large x gives the lower bound; the derivation of the upper bound is similar. □

The next result derives suitable bounding variables for use in Corollary 2. It is here that the gap, mentioned already, between using $\tilde{X}_2$ in the upper bound but $\tilde{X}_3(s_0)$ as a lower bound arises.

Lemma 20 *Under* P_s *(that is, when the parent is at* s*), for suitable* $0 < a < b < \infty$*,*

$$X \le b\frac{\tilde{X}_1}{V(s)} + \frac{b}{a}\tilde{X}_2, \qquad X \ge a\frac{\tilde{X}_1}{V(s)}$$

and, for $s > s_0$,

$$X \geq \frac{a}{b}\tilde{X}_3(s_0).$$

Proof. When the parent is at s, applying Lemma 15(i),

$$\begin{aligned}
X &= \frac{\sum_j V(z_j+s)e^{-(z_j+s)}I(z_j+s>0)}{V(s)e^{-s}} \\
&= \frac{\sum_j V(z_j+s)e^{-z_j}I(z_j>-s)}{V(s)} \\
&\leq \frac{\sum_j b(z_j+s+1)e^{-z_j}I(z_j>-s)}{V(s)} \\
&\leq \frac{b}{V(s)}\sum_j z_j e^{-z_j}I(z_j>0) + \frac{b}{a}\sum_j e^{-z_j} \\
&= b\frac{\tilde{X}_1}{V(s)} + \frac{b}{a}\tilde{X}_2,
\end{aligned}$$

as required. Similarly

$$X \geq \frac{\sum_j a(z_j+s+1)e^{-z_j}I(z_j>-s)}{V(s)} \geq \frac{a\sum_j z_j e^{-z_j}I(z_j>0)}{V(s)} = a\frac{\tilde{X}_1}{V(s)}$$

and, for $s > s_0$

$$X \geq \frac{\sum_j a(z_j+s+1)e^{-z_j}I(z_j>-s)}{b(s+1)} \geq \frac{a}{b}\sum_j e^{-z_j}I(z_j>-s_0) = \frac{a}{b}\tilde{X}_3(s_0).$$

□

In applying the first of these bounds in Theorem 5(i) and Corollary 2(i), Lemma 8 shows that the variables $\tilde{X}_1/V(s)$ and $\tilde{X}_2$ can actually be treated separately; clearly, $\tilde{X}_1/V(s)$ and $\tilde{X}_3(s_0)$ can be considered separately in part (ii) of these results. In dealing with $\tilde{X}_1$, $V(s)^{-1}$ provides the function $g(s)$. Hence there are two functions A of the form defined in the main theorems that have to be considered. These are introduced and estimated in the next Lemma. Notice how accounting for the function $V(s)^{-1}$ that multiplies $\tilde{X}_1$ alters the associated growth rate.

Lemma 21 *Let $A_1(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x \geq 1)$. Then*

$$\limsup_{x\to\infty}\frac{A_1(x)}{(\log x)^2\log\log\log x} < \infty \quad \textit{and} \quad \liminf_{x\to\infty}\frac{A_1(x)}{(\log x)^2/\log\log\log x} > 0.$$

Let $A_2(x) = \sum_{i=1}^{\infty} V(\zeta_i)^{-1}I(H(\zeta_i)x \geq V(\zeta_i))$. Then

$$\limsup_{x\to\infty}\frac{A_2(x)}{(\log x)\log\log\log x} < \infty \quad \textit{and} \quad \liminf_{x\to\infty}\frac{A_2(x)}{(\log x)/\log\log\log x} > 0;$$

Proof.

$$A_1(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x \geq 1) \;=\; \sum_{i=1}^{\infty} I(V(\zeta_i)e^{-\zeta_i}x \geq 1)$$
$$= \sum_{i=1}^{\infty} I(\zeta_i \leq \log x + \log V(\zeta_i)).$$

Since, $0 \leq \log V(x) \leq \log b(x+1)$ and for any $\epsilon > 0$ there is an $\gamma > 0$ such that $\log b(x+1) \leq \gamma + \epsilon x$ it follows that

$$\sum_{i=1}^{\infty} I(\zeta_i \leq \log x) \leq A_1(x) \leq \sum_{i=1}^{\infty} I\left((1-\epsilon)\zeta_i \leq \log x + \eta\right),$$

that is

$$D(\log x) \leq A_1(x) \leq D((1-\epsilon)^{-1}(\log x + \eta)).$$

The results in Theorem 11 complete the required estimation of A_1.

Similarly

$$A_2(x) \;=\; \sum_{i=1}^{\infty} V(\zeta_i)^{-1} I(H(\zeta_i)x \geq V(\zeta_i))$$
$$= \sum_{i=1}^{\infty} V(\zeta_i)^{-1} I(\zeta_i \leq \log x)$$

and the estimates in Lemma 19 finish the proof. □

These estimates of the growth in terms of slowly varying functions translate into moment conditions through Corollary 2 to give Theorem 10. The estimates for A_1 gives the moment conditions on $\tilde{X}_2$ and $\tilde{X}_3(s)$, while A_2 gives those on $\tilde{X}_1$.

8 The derivative martingale in branching random walk

Much as at the start of Section 6, let Z be a point process on the reals with intensity measure μ and assume that

$$\int e^{-x}\mu(dx) = 1, \;\; \int xe^{-x}\mu(dx) = 0 \;\text{ and }\; \int x^2 e^{-x}\mu(dx) < \infty.$$

The spatially homogeneous branching random walk based on Z is considered; this has types in $(-\infty, \infty)$ generated in the obvious way. The first assumption on μ implies that $H(s) = e^{-s}$ is mean-harmonic. The corresponding martingale,

$$W_n = \sum_{|\nu|=n} e^{-S(\nu)},$$

belongs to a class of martingales studied by several authors, see, for example Kingman (1975), Biggins (1977), Liu (1997) and Lyons (1997).

Lemma 22 *(i) $E_{\mathbb{P}}W = 0$, (ii) $\mathbb{P}(W = 0) = 1$ and (iii) $\inf\{S(\nu) : |\nu| = n\} \to \infty$, almost surely as $n \to \infty$.*

This result is contained in Biggins (1977); the second and third assertions are immediate consequences of $E_{\mathbb{P}}W = 0$. Informally, the last part says that every line of descent goes to infinity. A proof that $E_{\mathbb{P}}W = 0$ without the side condition that $\int x^2 e^{-x}\mu(dx)$ is finite was given by Lyons (1997), can also be obtained fairly directly from Corollary 1 and is contained in Theorem 6(iii)(a); however, this strengthening is not needed here.

If $m(\phi) = \int e^{-\phi x}\mu(dx)$ is finite then

$$W_n(\phi) = \sum_{|\nu|=n} \frac{e^{-\phi S(\nu)}}{m(\phi)^n}$$

is a positive martingale, with W_n being $W_n(1)$. Differentiating $W_n(\phi)$ with respect to ϕ and setting $\phi = 1$ shows that

$$\partial W_n = \sum_{|\nu|=n} S(\nu)e^{-S(\nu)}$$

ought to be a martingale, a fact easily verified by direct calculation. This is called the derivative martingale, even if $m(\phi)$ is not actually finite anywhere other than at $\phi = 1$ so that the derivative is fictional. The derivative martingale does not fall directly into the general framework developed earlier because the function $H(s) = se^{-s}$ producing it takes negative values. The basic idea for dealing with this problem, which was used by Harris (1999) in the context of branching Brownian motion, is to approximate the spatially homogeneous process by one with an absorbing barrier. The last part of Lemma 22 suggests that, for large b, a barrier at $-b$ should make little difference. Then, for the approximating process, the counterpart of ∂W_n will be a non-negative martingale amenable to the theory developed.

Known results for the martingales $W_n(\phi)$ and analogous martingales for branching Brownian motion, which we do not attempt to describe, indicate that the case under consideration ($\phi = 1$) is a boundary one. In the branching Brownian motion context the boundary case corresponds to the travelling wave solution to the associated reaction-diffusion equation with slowest speed. This was the context of the study of Harris (1999) whose approach is adapted here to yield convergence of the derivative martingale. Kyprianou (2001) gives an extensive discussion of the use of change of measure ideas for branching Brownian motion, and other branching diffusions, and of the use of a barrier to discuss derivative martingales in that context.

The convergence of derivative martingales and related questions have been considered before; see, for example, Biggins (1991, 1992) and Barral (2000) for non-boundary cases. More relevantly, the convergence in the boundary case has been considered by Kyprianou (1998) and Liu (2000), drawing on results from a related functional equation; the approach here, which is more direct, gives convergence under weaker conditions.

Theorem 12 *The martingale ∂W_n converges to a finite non-negative limit, Δ, almost surely. Furthermore, $\mathbb{P}(\Delta = 0)$ is equal to either the extinction probability or one.*

The limit Δ has infinite mean, and so is not identically zero, when the conditions of Theorem 10(i) hold. The limit is identically zero when the conditions of Theorem 10(ii) hold.

Proof. Let t_b be the event that no node in the branching random walk has a position to the left of $-b$; then, by Lemma 22(iii), t_b increases to an event with probability one as $b \to \infty$. Use the homogeneous branching random walk to construct a branching random walk with a barrier at $-b$; on t_b the processes with and without a barrier agree. To make the coupling precise, let $I_b(\nu)$ be one if the node ν is retained in the process with a barrier at $-b$ and zero otherwise. Now, by Lemma 15(ii), $V(b)^{-1}\sum_{|\nu|=n} V(b+S(\nu))e^{-S(\nu)}I_b(\nu)$ is a positive martingale, which must converge to a finite limit, denoted by B_b. Hence, using Lemma 15(i) and Lemma 22,

$$\begin{aligned}
B_b &= \lim_{n\to\infty}\sum_{|\nu|=n}\frac{V(b+S(\nu))}{V(b)}e^{-S(\nu)}I_b(\nu)\\
&= \frac{1}{V(b)}\lim_{n\to\infty}\sum_{|\nu|=n}V(b+S(\nu))e^{-S(\nu)}I_b(\nu)\\
&= \frac{C}{V(b)}\lim_{n\to\infty}\sum_{|\nu|=n}(b+S(\nu))e^{-S(\nu)}I_b(\nu)\\
&\le \frac{C}{V(b)}\lim_{n\to\infty}(W_n b+\partial W_n)\\
&= \frac{C}{V(b)}\lim_{n\to\infty}\partial W_n;
\end{aligned}$$

furthermore, equality holds on t_b. Thus ∂W_n converges to $\Delta = C^{-1}V(b)B_b$ on t_b. Letting $b \to \infty$ completes the proof that ∂W_n has a finite, non-negative limit.

Splitting on the first generation and letting n go infinity shows, drawing on Lemma 22(i), that

$$\Delta(S) = \sum_{|\nu|=1} e^{-S(\nu)}\Delta(S^\nu). \tag{12}$$

Hence, $\mathbb{P}(\Delta = 0)$ is a fixed point of the generating function of the family size and so must have the stated property.

It has already been shown that $V(b)B_b \le C\Delta$ with equality on t_b. When Theorem 10(i) holds, $E_{\mathbb{P}}B_b = 1$ and then $V(b) \le CE_{\mathbb{P}}\Delta$ for any b; thus $E_{\mathbb{P}}\Delta = \infty$. Similarly, when the conditions of Theorem 10(ii) hold B_b, and hence Δ, is zero on t_b for every b. □

The equation (12) is an example of a smoothing transform, in the sense of Durrett and Liggett (1983) and Liu (1998); the Laplace transform of Δ satisfies an associated functional equation. It was this functional equation that was important in the study of the convergence of ∂W_n in Kyprianou (1998) and Liu (2000). In contrast, the results here yield results about the functional equation as a by-product, as will be explained in Biggins and Kyprianou (2001). For that study, it turns out to be important to consider the analogue of ∂W_n over certain optional lines.

Specifically, the optional lines $\mathcal{C}[t]$ have the same definition as in Section 6 and

$$\partial W_{\mathcal{C}[t]} = \sum_{\nu \in \mathcal{C}[t]} S(\nu) e^{-S(\nu)}.$$

It is worth noting explicitly that, unlike $(W_{\mathcal{C}[t]}, \mathcal{G}_{\mathcal{C}[t]})$, $(\partial W_{\mathcal{C}[t]}, \mathcal{G}_{\mathcal{C}[t]})$ is not a martingale (though it is a submartingale). Nonetheless, the ideas in the proof of Theorem 12 yield the following result.

Corollary 5 *$\partial W_{\mathcal{C}[t]}$ converges to Δ almost surely.*

Proof. Note first that Proposition 3 shows that $W_{\mathcal{C}[t]}$ and W_n have the same limit; by Lemma 22, this limit is zero and so $\inf\{S(\nu) : \nu \in \mathcal{C}[t]\} \to \infty$ as $t \to \infty$. Now, applying Theorem 9 shows that on t_b

$$\begin{aligned} \Delta &= \frac{V(b)}{C} \lim_{n\to\infty} \sum_{|\nu|=n} \frac{V(b+S(\nu))}{V(b)} e^{-S(\nu)} \\ &= \frac{V(b)}{C} \lim_{t\to\infty} \sum_{\nu\in\mathcal{C}[t]} \frac{V(b+S(\nu))}{V(b)} e^{-S(\nu)} \\ &= \lim_{t\to\infty} (W_{\mathcal{C}[t]} b + \partial W_{\mathcal{C}[t]}) = \lim_{t\to\infty} \partial W_{\mathcal{C}[t]}. \end{aligned}$$

Letting $b \to \infty$ completes the proof. □

References

[1] Athreya, K. (2000) Change of measures for Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli.* **6**, 323–338

[2] Athreya, K.B. and Ney, P.E. (1972) *Branching Process.* Springer-Verlag, Berlin.

[3] Barral, J. (2000) Differentiability of multiplicative processes related to branching random walk. *Ann. Inst. H. Poincaré* **36**, 407–417.

[4] Bertoin, J. (1993) Splitting at the infimum and excursions in half lines for random walks and Lévy processes. *Stoc. Proc. Appl.* **47**, 17–35.

[5] Bertoin, J. and Doney, R.A. (1994) On conditioning a random walk to stay positive. *Ann. Probab.* **22**, 2152–2167.

[6] Biggins, J.D. Uniform convergence of martingales in the one-dimensional branching random walk. *IMS lectures notes — Monograph Series. Selected proceedings of the Sheffield Symposium on Applied Probability, 1989,* Eds. I.V. Basawa and R.L. Taylor. (1991), **18**, 159–173.

[7] Biggins, J.D. (1992). Uniform convergence of martingales in the branching random walk. *Ann. Probab.* **20**, 137–151.

[8] Biggins, J.D. (2001). Random walk conditioned to stay positive. Submitted to *LMS*; preprint available at
`http://www.shef.ac.uk/~st1jdb/rwctsp.html`

[9] Biggins, J.D. and Kyprianou, A.E. (2001). The smoothing transform; the boundary case. Preliminary version available at: `http://www.shef.ac.uk/~st1jdb/tsttbc.html`

[10] Biggins, J.D., Lubachevsky, B.D., Shwartz, A. and Weiss, A. (1991). A branching random walk with a barrier. *Ann. Appl. Probab.* **1**, 573–581.

[11] D'Souza, J.C. and Biggins, J.D. (1992). The Supercritical Galton-Watson process in varying environments. *Stoc. Proc. Appl.* **42**, 39–47.

[12] Durrett, R. (1996) *Probability: Theory and Examples. 2nd Edition* Duxbury, Belmont, CA.

[13] Durrett, R. and Liggett, M. (1983). Fixed points of the smoothing transform. *Z. Wahrsch. verw. Gebeite*, **64**, 275–301.

[14] Goettge, R.T. (1976). Limit Theorems for the Supercritical Galton-Watson Process in Varying Environments, *Math. Biosci.* **28** 171–190.

[15] Harris, S.C. (1999) Travelling waves for the FKPP equation via probabilistic arguments, *Proc. Roy. Soc. Edin.* **129A** 503–517.

[16] Jagers, P. (1989). General branching processes as Markov fields. *Stoc. Proc. Appl.* **32**, 183–212.

[17] Kesten, H. (1978) Branching Brownian motion with absorption. *Stoc. Proc. Appl* **7** 9–47.

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

[19] Kyprianou, A.E. (1998) Slow variation and uniqueness of solutions to the functional equation in the branching random walk. *J. Appl. Probab.* **35** 795–802.

[20] Kyprianou, A.E. (2001) Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris; probabilistic analysis. Submitted to: *Ann. Inst. H. Poincaré*

[21] Kyprianou, A., and Rahimzadeh Sani A. (2001) Martingale convergence and the functional equation in the multi-type branching random walk. *Bernoulli.* **7**, 593–604.

[22] Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y. (1997). A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes* (K.B. Athreya, P. Jagers, eds.). *IMA Volumes in Mathematics and its Applications* **84**, 181–185. Springer-Verlag, New York.

[23] Liu, Q. (1997). Sur une equation fonctionelle et ses applications: une extension du theoreme de Kesten-Stigum concernant des processus de branchement *Adv. Appl. Probab.* **29**, 353–373

[24] Liu, Q. (1998) Fixed points of a generalized smoothing transform and applications to the branching processes. *Adv. Appl. Probab.* ***30**, 85–112.*

[25] Liu, Q. (2000) On generalized multiplicative cascades. *Stoc. Proc. Appl.* **86**, 263–286

[26] Lyons, R. (1997). A simple path to Biggins' martingale convergence. In *Classical and Modern Branching Processes* (K.B. Athreya, P. Jagers, eds.). *IMA Volumes in Mathematics and its Applications* **84**, 217–222. Springer-Verlag, New York.

[27] Lyons, R., Pemantle R. and Peres, Y. (1995) Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes. *Ann. Probab.* **23**, 1125–1138.

[28] MacPhee, I.M. and Schuh, H.J. (1983) A Galton-Watson branching process in varying environments with essentially constant means and two rates of growth, *Austral. J. Statist.* **25** 329-338.

[29] Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*, Holden-Day, London.

[30] Neveu, J. (1988). Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987,* eds: E. Çinlar, K.L. Chung, R.K. Getoor. Progress in Probability and Statistics, **15**, 223-241. Birkhäuser, Boston.

[31] Olofsson, P. (1998) The $x \log x$ condition for general branching process. *J. Appl. Probab.* **35**, 537–544

[32] Tanaka, H. (1989) Time reversal of random walks in one-dimension. *Tokyo J. Math.* **12**, 159–174

[33] Waymire, E.C. and Williams, S.C. (1996). A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* **348**, 585–632.

J.D. Biggins
Department of Probability and Statistics,
Hicks Building, The University of Sheffield,
Sheffield, S3 7RH,
U.K.

A.E. Kyprianou
The University of Utrecht
Department of Mathematics
Buadapestlaan 6, 3584CD
The Netherlands