# 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 to this theorem pioneered by Lyons, Peres and Pemantle (1995) is extended to certain kinds of martingales defined for Galton-Watson processes with a general type space. Many examples satisfy stochastic domination conditions on the offspring distributions and suitable domination conditions combine nicely with general conditions for mean convergence 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 various branching processes is indicated. 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

In the one-type Galton-Watson process with mean family size $m \in (0, \infty)$ the Kesten-Stigum Theorem states that the non-negative martingale $W_n$, formed by the population size normed by its expectation, converges in mean exactly when $EN \log^+ N$ is finite, where $N$ is the family size. This result has been generalised to several more complicated branching processes and a number of different approaches to the proof are known. The original motivation for this study was the search for sharp results of this kind on the mean convergence of a signed martingale that arises in the homogeneous branching random walk. That signed martingale can be well approximated by other, non-negative, ones and the method initiated by Lyons, Peres and Pemantle (1995), which exploits a change of measure argument, seemed the most promising approach to the mean convergence of these approximating martingales. That initial study discusses the one-type Galton-Watson process but there have been several later papers dealing with other models. In

* *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

particular, Athreya (2000), considers a general multitype homogeneous process, which is fairly close to the framework adopted here. However, none of the existing results seemed applicable to the problem we faced. Rather than develop the theory in the context of the motivating example an attempt has been made to derive quite general results, which apply to many different examples, including the one we are particularly interested in. Hence, the most general part of this discussion has some of the character of a review and so we will show that the results obtained apply to a variety of models. Sometimes results that are technically new are obtained in this way, but that is not the main point.

Kesten-Stigum like results often contain the assertion that, when the martingale converges in mean the process dies out on the trajectories where the martingale limit is zero. No attempt is made here to seek general conditions for this to be so. MacPhee and Schuh (1983) and D'Souza and Biggins (1992) give examples of (varying environment) processes where this assertion fails.

The usual multitype branching process is considered, except that the type space, $\mathcal{S}$, is quite general. The process starts from a single individual of a specified type who produces a family whose members also have types in $\mathcal{S}$. These children produce families in a similar way, and so on. It is convenient for the formulation to insist that every person has a countably infinite number of children. More usual formulations can be embedded in this one by having one of the types in $\mathcal{S}$ as a 'ghost'-type, $\partial$, that always has all its children of type $\partial$ and is interpreted as the person and all their descendants being absent.

To be a bit more formal about the sample space, let $\mathcal{T}$ be the labelled nodes of the family tree in which every node has a countably infinite number of children. The basic random object constructed is $S$, which is a $\mathcal{S}$-valued function on the nodes of $\mathcal{T}$ and is drawn from the sample space $\mathfrak{B} = \mathcal{S}^{\mathcal{T}}$. Then $S(\nu)$ is the type of node $\nu \in \mathcal{T}$.

Reproduction depends on the parent's type; given that type, there is an associated distribution of the types of the children, called the family distribution. Let $\mathfrak{F} = \mathcal{S}^{\mathbb{N}}$ and for $f \in \mathfrak{F}$ write $f = (f_1, f_2, \ldots)$. The children of an individual are described by an element $f \in \mathfrak{F}$, $f_1$ is the type of the first child, $f_2$ the type of the second and so on. A kernel, $P_s(df)$, from $\mathcal{S}$ to $\mathfrak{F}$ specifies the reproduction mechanism. The family distribution for a parent of type $s$ is $P_s$ and the corresponding expectation is denoted by $E_s$.

The family tree is produced in the usual way from this specification of family distributions. Given the family history to generation $n$, individuals in that generation reproduce independently of each other with the family distribution for each parent's type. In this way, the law for the branching process $\mathbb{B}$ is constructed from the kernel $P_s(df)$ by using the theorem of Ionescu Tulcea. Here, $\mathbb{B}$ is the law given the type of the initial ancestor, but, for notational simplicity, the starting type is not explicitly recorded. The corresponding expectation is $E_{\mathbb{B}}$.

A finite non-negative function $H$ on $\mathcal{S}$ will be called mean-harmonic when $H(\tilde{s}) > 0$ for some $\tilde{s}$ and, with $f = (f_1, f_2, \ldots) \in \mathfrak{F}$,

$$E_s\left[\sum_i H(f_i)\right] = H(s) \quad \text{for all } s \in \mathcal{S}.$$

Thus mean-harmonic functions are 'conserved' on average under reproduction and so, as we will now see, produce martingales.

Write $|\nu|$ for the generation of the node $\nu \in \mathcal{T}$, $c(\nu)$ for the children of $\nu$ and 0 for the initial ancestor. Let $\mathcal{G}_n$ be the $\sigma$-algebra generated by the first $n$ generations. Also, though the notation will not be needed for some time, let $\{\nu_i : i = 0, 1, \ldots, |\nu|\}$ be the ancestry of $\nu$ ordered in the natural way, starting from $\nu_0 = 0$.

The functions $\{W_n\}$ are defined for $S \in \mathfrak{B}$ by

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

then

$$\begin{aligned} E_{\mathbb{B}}\left[W_n \middle| \mathcal{G}_{n-1}\right] &= \sum_{|\nu|=n-1} E_{\mathbb{B}}\left[\sum_{\sigma \in c(\nu)} H(S(\sigma)) \middle| \mathcal{G}_{n-1}\right] \\ &= \sum_{|\nu|=n-1} E_{S(\nu)}\left[\sum_i H(f_i)\right] = \sum_{|\nu|=n-1} H(S(\nu)) = W_{n-1} \end{aligned}$$

and so $W_n$ forms a non-negative martingale with respect to $\mathcal{G}_n$. 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. Let

$$\mathcal{S}^H = \{s \in \mathcal{S} : H(s) > 0\},$$

which are the types for the initial ancestor that give $W_0 > 0$. The main objective here is to give conditions that determine when the martingale $W_n$ converges in mean; that is to obtain Kesten-Stigum like results for such martingales. Clearly, this question is only interesting when the type of the initial ancestor lies in $\mathcal{S}^H$.

When $Y$ is defined on $\mathcal{S} \times \mathfrak{F}$, $E_s Y$ is defined, on $\mathcal{S}$, by $E_s Y = \int Y(s, f) P_s(df)$. Let $\mathfrak{B}_n$ be the projection of $\mathfrak{B}$ onto the first $n$ generations. Clearly, $\mathfrak{B}_1$ is equivalent to $\mathcal{S} \times \mathfrak{F}$, with the first component being the initial type and the second being the types of the first generation. The definition of $E_s Y$ therefore serves also for $Y$ defined on $\mathfrak{B}_1$.

Conditions are needed on the distribution of $W_1/W_0$ as the initial type, and hence $W_0$, varies. To describe these neatly, let

$$X(S) = \frac{W_1}{W_0} I(W_0 > 0) + I(W_0 = 0) \text{ for } S \in \mathfrak{B}.$$

The variables $W_1$ and $W_0$ and $X$ are all defined on the sample space $\mathfrak{B} = \mathcal{S}^{\mathcal{T}}$ but are actually determined on $\mathfrak{B}_1$. Hence, $X$ can be defined in a consistent way also on $\mathcal{S} \times \mathfrak{F}$. Specifically, with $f = (f_1, f_2, \ldots) \in \mathfrak{F}$, $X$ is given by

$$X(s, f) = \frac{\sum_i H(f_i)}{H(s)} I(H(s) > 0) + I(H(s) = 0). \tag{1-1}$$

With this definition, $P_s(X > x)$ makes sense.

One further ingredient is needed before a typical result can be stated. Let $\zeta = \{\zeta_0, \zeta_1, \ldots\}$ be the Markov chain on $\mathcal{S}^H$ with the (proper) transition measure given by

$$\frac{1}{H(s)} E_s\left[\sum_i H(f_i) I(f_i \in A)\right] \quad \text{for } A \subset \mathcal{S}^H. \tag{1-2}$$

The fact that $H$ is mean-harmonic ensures that this is a probability measure for any $s \in \mathcal{S}^H$. Whenever $\zeta$ occurs, it is assumed that the type of the initial ancestor of the branching process is in $\mathcal{S}^H$, so that $W_0 > 0$, and $\zeta_0$ is given by this type.

The interplay between the development of the Markov chain $\zeta$ and the distribution of $X$ under $P_{\zeta_n}$ often determines when $W_n$ converges in mean. The next theorem, which is a special case of Corollary 2.5 given later, gives the flavour. 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. The essence of the result is that estimates of the behaviour at infinity of a certain random function lead to moment conditions on suitable bounding variables for the family size distributions which imply either $E_{\mathbb{B}}W = W_0$ or $E_{\mathbb{B}}W = 0$.

**Theorem 1.1** *For $x > 0$, let*

$$A(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x > 1),$$

*which is a random function of $\zeta$. Suppose $L$ is a positive increasing function that is slowly varying at infinity; $L$ may be different in (i) and (ii).*

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

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

*and that $\sup_{x>0}\{A(x)/L(x)\}$ is bounded above, almost surely. If $E[X^*L(X^*)] < \infty$ then $E_{\mathbb{B}}W = W_0$.*

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

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

*and that, for some $y$, $\inf_{x>y}\{A(x)/L(x)\}$ is bounded below by a positive constant, almost surely. If $E[X_*L(X_*)] = \infty$ then $E_{\mathbb{B}}W = 0$.*

For orientation, it is worth casting the simplest, much studied, case into the present framework, thereby illustrating that the type space often needs to be richer than in the traditional formulations. Consider the (one-type) Galton-Watson process with family size $N$ satisfying $EN = m \in (1, \infty)$. The appropriate type space is $\mathcal{S} = \{\partial, 0, 1, 2, \ldots\}$, where $\partial$ is a 'ghost' type and other nodes are typed by their generation. Hence, generically, a person of type $i$ gives birth to $N$ children of type $i + 1$, and the remaining children are of type $\partial$. 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, given by normalising the population size at generation $n$ by its expectation, $m^n$. Then $\zeta_i = i$ and so $A(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x > 1) = \sum_{i=1}^{\infty} I(m^i < x) \approx \log x / \log m$. Furthermore both $X^*$ and $X_*$ can be $N/m$. 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 familiar 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 signed martingale for a certain boundary case in the homogeneous branching random walk. We call that martingale a 'derivative' martingale for reasons explained in Section 5.

We now summarise how the treatment will develop. The next section describes the general results about the mean convergence of the martingale $W_n$. This is followed immediately by a section discussing several simple applications of these general results. Then, in Sections 4 and 5, the homogeneous branching random walk is introduced, the derivative martingale is described and discussed and the results obtained about it are stated. We need to consider the sum of $H(S(\nu))$ over collections of nodes other than the $n$th generation ones; specifically over what Jagers (1989) calls optional lines. Section 6 introduces the ideas and gives the results we need, which concern conditions for the limit over an increasing sequence of such lines to be $W$, that is, to be the same limit as when the lines are just formed by the generations. Sections 7 to 9 contains a discussion of various examples: the branching random walk in a random, ergodic environment; the multitype branching random walk; and the general branching (or CMJ) process. The approach to the derivative martingale is through a coupling to branching random walk with a barrier. That process is described and the relevant results about it obtained in Section 10. The last four sections give the proofs of the results claimed in Sections 2, 5 and 6. In particular, a full description of the measure change at the heart of the proof is deferred until Section 12, since the results described in Section 2 and, hence, the applications of them do not require knowledge of it.

## 2 Results on mean convergence of $W_n$

The various general results on mean convergence are now recorded. The proofs are in Sections 12 and 13. It is not necessary to read these proofs to follow the various applications of the results. All the results in this section involve conditions on the Markov chain $\zeta$ with transition kernel given at (1-2). It is not immediately clear that these conditions translate to useful conditions on the family distributions. In fact they often do and the main route for this is via Corollaries 2.5 and 2.7 given towards the end of the section.

The first theorem is the basic one with the subsequent ones being deductions from it that are designed to be easier to apply. Here, and throughout, $x \wedge y = \min\{x, y\}$.

**Theorem 2.1** *Let $X$ be given by (1-1) and $\zeta$ be the Markov chain with transition kernel (1-2).*

*(i) If*

$$\sum_{n=1}^{\infty} E_{\zeta_n}\left[X((H(\zeta_n)X) \wedge 1)\right] < \infty, \tag{2-1}$$

*almost surely, then $E_{\mathbb{B}}W = W_0$.*

*(ii) If, for all $y > 0$,*

$$\sum_{n=1}^{\infty} E_{\zeta_n}[XI(H(\zeta_n)X \geq y)] = \infty, \tag{2-2}$$

*almost surely, then $E_{\mathbb{B}}W = 0$.*

*(iii) If* $\limsup_n H(\zeta_n) = \infty$ *almost surely then* $E_{\mathbb{B}} W = 0$.

Writing $x(a \wedge 1) = xaI(a < 1) + xI(a \geq 1)$ splits (2-1) into two sums; the second of these is then the one in (2-2) when $y = 1$. This indicates that (2-1) and (2-2) are quite close; a necessary and sufficient condition for mean convergence of the martingale will be obtained when there are no intermediate cases. It is possible to refine the first part of this theorem slightly at some cost in elegance. The refinement is useful if $X$ becomes degenerate under $P_{\zeta_n}$ as $n$ increases, which is not usual. We will use this extension, given in the next result, only for one-type varying environment Galton-Watson processes.

**Proposition 2.2** *In Theorem 2.1 the condition (2-1) can be replaced by the weaker condition*

$$\liminf_n H(\zeta_n) < \infty \quad \textit{and} \quad \sum_{n=1}^{\infty} \frac{E_{\zeta_n}\left[\sum_i H(f_i)\left(\left(\sum_{j \neq i} H(f_j)\right) \wedge 1\right)\right]}{H(\zeta_n)} < \infty \tag{2-3}$$

*without changing the conclusion. When* $\sum_n H(\zeta_n) < \infty$ *almost surely (2-1) and (2-3) are equivalent.*

One other refinement, which will not be used, is also worth recording. Recall that, by assumption, the initial type is chosen with $W_0 > 0$.

**Proposition 2.3** *If either (2-1) or (2-3) holds with positive probability (rather than almost surely) then* $E_{\mathbb{B}} W > 0$*. If (2-2) holds with positive probability then* $E_{\mathbb{B}} W < W_0$*.*

The proofs of these three results are given in Section 12 and the basic measure change used in the proof is described there. The rest of the results described in this section arise from using stochastic domination conditions to simplify the series in Theorem 2.1. Their proofs, which require nothing from Section 12, are in Section 13.

In the first part of the next Theorem, note that $A$ is a function of $\zeta$, and so is random, and that the expectation in (2-4) is only over the auxiliary random variable $X^*$, not over $A$, which accounts for the qualification 'almost surely'. The same is true of (2-5) in the second part. For orientation, this 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 2.4**

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

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

*and that* $\zeta$ *is eventually in* $F$*, almost surely. Let the increasing function* $A$ *be defined by*

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

*If*

$$\int_1^{\infty} \frac{E\left[X^* A(X^* w)\right]}{w^2} dw < \infty \qquad (\textit{almost surely}) \tag{2-4}$$

*then* $E_{\mathbb{B}} W = W_0$.

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

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

*Let*

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

*If, for all* $w > 0$,

$$E[X_* A(X_* w)] = \infty \qquad (\textit{almost surely}) \tag{2-5}$$

*then* $E_{\mathbb{B}} W = 0$.

The next result shows that (2-4) and (2-5) become simple moment conditions when the appropriate $A$ can be bounded suitably.

**Corollary 2.5** *Let* $L$ *be a positive increasing function that is slowly varying at infinity and let* $\delta$ *be a constant satisfying* $\delta \in [0, 1)$; $L$ *and* $\delta$ *may be different in (i) and (ii).*

*(i) In Theorem 2.4(i) suppose, instead of (2-4), that*

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

*and* $E[(X^*)^{1+\delta} L(X^*)] < \infty$. *Then* $E_{\mathbb{B}} W = W_0$.

*(ii) In Theorem 2.4(ii) suppose, instead of (2-5) that, for some* $y$,

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

*and* $E[(X_*)^{1+\delta} L(X_*)] = \infty$. *Then* $E_{\mathbb{B}} W = 0$.

These results suffice for most purposes; indeed $g$ can often be taken to be the identity.

However, for our main example the natural upper bounds on the reproduction take a more complex form than that in Theorem 2.4(i) and so the next two results formulate a straightforward extension. Roughly, they say that if the upper bound involves the sum of several random variables it is enough to check moment conditions for them separately.

**Theorem 2.6** *Throughout,* $j \in \{1, 2, \ldots, J\}$. *Suppose that there are random variables* $X_j^*$, *positive functions* $g_j$ *on* $\mathcal{S}$ *and a subset* $F \subseteq \mathcal{S}$ *such that*

$$P_s(X > x) \leq P\left(\sum_j g_j(s) X_j^* > x\right) \quad \textit{for all } s \in F \subset \mathcal{S} \tag{2-6}$$

*and that* $\zeta$ *is eventually in* $F$, *almost surely. Let the increasing functions* $A_j$ *be defined by*

$$A_j(x) = \sum_i g_j(\zeta_i) I(x g_j(\zeta_i) H(\zeta_i) \geq 1).$$

*If*

$$\max_j \int_1^{\infty} \frac{E\left[X_j^* A_j(w X_j^*)\right]}{w^2} dw < \infty \qquad (\textit{almost surely}) \tag{2-7}$$

*then* $E_{\mathbb{B}} W = W_0$.

**Corollary 2.7** *Let $L_j$ be positive increasing functions that are slowly varying at infinity and let $\delta_j$ be constants satisfying $\delta_j \in [0,1)$. In Theorem 2.6 suppose, instead of (2-7), that*

$$\max_j \sup_{x>0} \left\{ A_j(x) / \left( x^{\delta_j} L_j(x) \right) \right\} < \infty \quad \textit{almost surely}$$

*and* $\max_j E[(X_j^*)^{1+\delta_j} L_j(X_j^*)]$ *is finite. Then* $E_{\mathbb{B}} W = W_0$.

# 3 Simple 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 $\partial$. 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 $n$th generation people. In this example, $\zeta_n = n$ and $H(\zeta_n) = H(n) = 1/EZ_n$, which are not random; hence, (2-1), (2-2) and (2-3) are all deterministic. A routine application of Proposition 2.2 gives the following result.

**Corollary 3.1** *Let $N_i' = N_i - 1$. If* $\inf_n EZ_n > 0$ *and*

$$\sum_{i=1}^{\infty} E\left[ \frac{N_i}{m_i} \left( \left( \frac{N_i'}{EZ_{i+1}} \right) \wedge 1 \right) \right] < \infty$$

*then* $E_{\mathbb{B}} W = 1$.

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

$$y \wedge 1 \leq K \left( 1 - \frac{1 - e^{-y}}{y} \right),$$

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

To illustrate the use of bounding variables, Corollary 2.5 yields the following result for varying environment processes. 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 already indicated.

**Corollary 3.2**

*(i) Suppose that* $\liminf (EZ_n)^{1/n} > 1$ *and there is a random variable $X^*$ such that*

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

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

*(ii) Suppose that there are positive integers* $\{n(1), n(2), \ldots\}$ *such that* $\sup_j \left( EZ_{n(j)} \right)^{1/j}$ *is finite and there is a random variable $X_*$ with*

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

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

*Proof.* This is an application of Corollary 2.5. 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 $EZ_n$ imposed 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).

### 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 $\rho$ 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 2.5 translates to one part of the multitype Kesten-Stigum Theorem. To see this, note first that the second component of $\zeta_n$ 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.

Combining this example with the previous one 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.

It is worth looking briefly at a homogeneous Galton-Watson process on a general type space $\Sigma$, with $(\sigma_1, \sigma_2, \ldots)$ being the types in a family. Suppose there is a function $\widetilde{H}$ and $m \in (0, \infty)$ such that $E_\sigma \sum_i \widetilde{H}(\sigma_i) = m\widetilde{H}(\sigma)$. Augmenting the type space to include generation, $H(n, \sigma) = \widetilde{H}(\sigma)/m^n$ is mean-harmonic. Write $\zeta_n = (n, \tilde{\zeta}_n)$ to identify the two components of $\zeta$. Then $\tilde{\zeta}_n$ is a Markov chain on $\Sigma$. Suppose $\tilde{\zeta}$ has a stationary distribution given by $\pi(d\sigma) = \widetilde{H}(\sigma)\nu(d\sigma)$ for some measure $\nu$. Let $Y = m^{-1} \sum_i \widetilde{H}(\sigma_i)$.

A little calculation shows that when $\tilde{\zeta}$ has its stationary distribution (2-1) holds when

$$\int E_\sigma \left[ Y \log^+ Y \right] I(\widetilde{H}(\sigma) > 0)\nu(d\sigma) < \infty.$$

Hence, when this holds, $W_n$ converges in mean for a set of initial types with probability one under $\pi$.

### 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. Recall that $\{\nu_i : i = 0, 1, \ldots, |\nu|\}$ is the ancestry of $\nu$. Let

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

so that $S'(\nu)$ contains all the information on the families of the ancestors of $\nu$. 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 2.5(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.

## 4 The branching random walk

The branching random walk in various guises will provide the more substantial test cases for the general results. The basic notation for it is introduced in this section,

Let $Z = \sum_i \delta(z_i)$ be a point process on the reals, with points at $\{z_i\}$ and intensity measure $\mu$. Points may be multiple, since $Z$ is a discrete measure with integer masses. Also, it is worth saying explicitly that $Z$ may have an infinite number of points. 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 (non-ghost) children of a person at $s$ is distributed like $Z$.

For a fixed real $\theta > 0$, let $m(\theta) = \int e^{-\theta z}\mu(dz)$ and assume this is finite; assume also that $-m'(\theta)$, interpreted as $\int z e^{-\theta z}\mu(dz)$, exists. Augment the types in $\mathbb{R}$ to include the generation; then

$$H(n, s) = e^{-\theta s} \prod_{i=0}^{n-1} \frac{1}{m(\theta)} = \frac{e^{-\theta s}}{m(\theta)^n}, \quad H(\partial) = 0 \tag{4-1}$$

is mean-harmonic.

Here and in all other examples the 'ghost' state $\partial$ contributes only zeros to the sums defining $W_n$. Hence, sums over $|\nu| = n$ can and will be interpreted as being over the nodes that occur, that is those that do not have type $\partial$. With this convention, the martingale derived from this mean-harmonic function is

$$W_n = \sum_{|\nu|=n} \frac{e^{-\theta S(\nu)}}{m(\theta)^n}. \tag{4-2}$$

This class of martingales has been studied by several authors, see, for example Kingman (1975), Biggins (1977), Liu (1997) and Lyons (1997). In particular, Lyons (1997) studies mean convergence using the techniques employed here; indeed, that paper was the main inspiration for the general results in this one.

Using (1-2), a straightforward calculation shows 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 $\mu$, so that $E[Y_i] = -m'(\theta)/m(\theta)$. Hence

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

With these observations it is straightforward to apply Theorem 2.1. However, rather than doing this, we include the result as a special case of the branching random walk in a random environment. This is discussed in Section 7, which could be read now.

It is easy to reduce the general case to the one where $\theta = 1$ and $m(\theta) = 1$. Specifically, given $\theta > 0$ and $Z$, let

$$Z^* = \sum_i \delta\left(\theta z_i - \log m(\theta)\right)$$

with points $z_i^* = \theta z_i - \log m(\theta)$. Let the intensity measure of $Z^*$ be $\mu^*$. Then

$$\int e^{-x} \mu^*(dx) = 1 \quad \text{and} \quad \int x e^{-x} \mu^*(dx) = -\left(\theta \frac{m'(\theta)}{m(\theta)} + \log m(\theta)\right).$$

The transformation of the type space $(n, z) \to (n, \theta z - n \log m(\theta))$ takes a realization of the branching process based on $Z$ to one based on $Z^*$.

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. In the branching random walk, let $z(\nu)$ be the displacement of $\nu$ from its parent and let $A(\nu) = e^{-z(\nu)}$; otherwise, when the node $\nu$ has type $\partial$, let $A(\nu) = 0$. Then the first generation point process $Z$ corresponds to $A = (A_1, A_2, \ldots)$, which is usually called the generator of the cascade and $m(1) = E \int e^{-z} Z(dz) = E \sum_i A_i$. This correspondence is discussed at greater length in Liu (1998), for example.

## 5 The derivative martingale in branching random walk

The derivative martingale considered, which will be defined shortly, 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. In particular, the approach used by Harris (1999) is adapted to yield convergence of the derivative martingale here.

Let

$$\partial W_n = \sum_{|\nu|=n} \left(S(\nu) + n \frac{m'(\theta)}{m(\theta}\right) \frac{e^{-\theta S(\nu)}}{m(\theta)^n}.$$

Differentiating (4-2) with respect to $\theta$ gives $-\partial W_n$ and indicates that $\partial W_n$ ought to be a martingale, a fact easily verified by direct calculation. For this reason, we call $\partial W_n$ a derivative martingale, even if $m(\xi)$ is not actually finite anywhere other than at $\xi = \theta$ so that the derivative is fictional.

The case when $\log m(\theta) = -\theta m'(\theta)/m(\theta)$ is particularly interesting and is the one we focus on. Known results for the martingale $W_n$ and its analogue for branching Brownian motion indicate that this case is a boundary one. 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.

Making the reduction to $\theta = 1$ and $m(\theta) = 1$ described in Section 4 and then focusing on the boundary case simplifies the assumptions about $\mu$ to

$$\int e^{-x}\mu(dx) = 1, \quad \int xe^{-x}\mu(dx) = 0 \tag{5-1}$$

and $\partial W_n$ becomes

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

Thus $\partial W_n$ corresponds to the 'harmonic' function $H(s) = se^{-s}$, which takes negative values. Since $\partial W_n$ is a signed martingale 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 using the results in Section 2. Kyprianou (2003) gives an extensive discussion of the use of change of measure ideas for branching Brownian motion and of the use of a barrier to discuss derivative martingales in that context. Here, the first result is that $\partial W_n$ does converge when a mild extra condition holds.

**Theorem 5.1** *Suppose (5-1) holds and $\int x^2 I(x<0)e^{-x}\mu(dx) < \infty$.*

*The martingale $\partial W_n$ converges to a finite non-negative limit, $\Delta$, almost surely. Furthermore,*

$$\Delta = \sum_j e^{-z_j}\Delta_j, \tag{5-2}$$

*where $\Delta_j$ are copies of $\Delta$ independent of each other and $Z$, and $\mathbb{B}(\Delta = 0)$ is equal to either the extinction probability or one.*

Moment conditions are needed to describe when the limit $\Delta$ is non-trivial. 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 $s \uparrow \infty$ and if $Z$ is concentrated on $(-s, \infty)$ then $\tilde{X}_3(s) = \tilde{X}_2$.

**Theorem 5.2** *Suppose (5-1) holds and that $\int x^2 e^{-x}\mu(dx) < \infty$.
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 $L_4(x) = (\log x)^2/\phi(x)$.*

*The limit $\Delta$ in Theorem 5.1 has infinite mean, and so is not identically zero, when both $E[\tilde{X}_1 L_1(\tilde{X}_1)]$ and $E[\tilde{X}_2 L_2(\tilde{X}_2)]$ are finite. The limit is identically zero when either $E[\tilde{X}_1 L_3(\tilde{X}_1)]$ is infinite or, for some $s$, $E[\tilde{X}_3(s) L_4(\tilde{X}_3(s))]$ is infinite.*

There is a (small) gap between the slowly varying functions used in the two sets of moment conditions in Theorem 5.2. In the proof, this gap arises from oscillations in the $\zeta$. 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.

The equation (5-2) is an example of a smoothing transform, in the sense of Durrett and Liggett (1983) and Liu (1998); the Laplace transform of $\Delta$ satisfies an associated functional equation. It was this functional equation that was important in the study of the convergence of $\partial W_n$ in Kyprianou (1998) and Liu (2000). In contrast, the results here yield results about the functional equation as a by-product, as explained in Biggins and Kyprianou (2001). For that study, it turns out to be important to consider the analogue of $\partial W_n$ in which the sum is formed over sets of nodes other than the generations. Specifically, the set of nodes $\mathcal{C}[t]$ is defined by individuals born to the right of $t$ but with all their antecedents born to the left of $t$. Thus $\mathcal{C}[t]$ focuses attention on nodes near $t$, regardless of generation, thereby making sums over $\mathcal{C}[t]$ comparatively well behaved. The set of nodes $\mathcal{C}[t]$ is what is called an optional line; the next section describes and develops the necessary theory about such lines.

Define $\partial W_{\mathcal{C}[t]}$ by

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

Results on optional lines and ideas used in the proof of Theorem 5.1 yield the following theorem.

**Theorem 5.3** *Suppose that (5-1) holds and that $\int x^2 I(x<0) e^{-x}\mu(dx) < \infty$. Let $\Delta$ be the limit of $\partial W_n$ introduced in Theorem 5.1. Then $\partial W_{\mathcal{C}[t]}$ converges to $\Delta$ almost surely.*

# 6 Optional lines

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

A function $\ell$ from the nodes $\mathcal{T}$ to $\{0,1\}$ identifies the set of nodes $\{\nu : \ell(\nu) = 1\}$. This set, and the corresponding function $\ell$, is called a *line* if no member of it is the ancestor of any other, so that (with the ancestry of $\nu$ being $\{\nu_i : i = 0, 1, \ldots, |\nu|\}$) $\ell(\nu) = 1$ implies that $\ell(\nu_i) = 0$ for all $i < |\nu|$. Lines in this sense cut across the family 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 $\ell$ is a function on the nodes it will often be convenient to identify $\ell$ with set of nodes where the function takes the value one. The $\sigma$-algebra $\mathcal{G}_\ell$ contains full information on the life-histories of all individuals that are neither in $\ell$ nor a descendent of any member of $\ell$. For the line $\ell$ let

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

Clearly $W_n$ is $W_\ell$ when $\ell$ is the line formed by all $n$th generation nodes. The partial ordering of $\mathcal{T}$ by 'is an ancestor of' ($<$) induces a partial order on lines, with $\ell_1 \leq \ell_2$ when every member of $\ell_2$ is a descendent (not necessarily strict) of some member of $\ell_1$. It therefore makes sense to speak of an increasing collection lines.

Informally, an *optional line*, $\mathfrak{L}$, is a random line with the property that its position is determined by the history of the process up to the position of the line. More precisely, an optional line $\mathfrak{L}$ is a random line with the property that, for any fixed line $\ell$, $\{\mathfrak{L} \leq \ell\} \in \mathcal{G}_\ell$. Then the $\sigma$-algebra associated with the optional line $\mathfrak{L}$, $\mathcal{G}_\mathfrak{L}$, 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. Important questions are when $W_\mathfrak{L}$ defines a martingale as $\mathfrak{L}$ varies through some increasing collection of optional lines and, when it does, whether its limit is the same as that of the martingale $W_n$.

Unfortunately, optional lines seem to be too general for some of the results sought here, necessitating some restriction. The optional line $\mathfrak{L}$ will be called *simple* when, for all $\nu$, the function $\mathfrak{L}(\nu)$ is measurable with respect to $\mathcal{G}_{|\nu|}$; thus, whether $\nu$ is on the line or not is determined by looking at the process up to generation $|\nu|$. Let $\mathcal{A}_\nu$ be the $\sigma$-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 $\nu$, the function $\mathfrak{L}(\nu)$ is measurable with respect to $\mathcal{A}_\nu$; then, whether $\nu$ is on the line or not is determined by looking at the types in its ancestry.

For a very simple optional line the rule applied to a line of descent to determine membership of the line can be applied to any trajectory of the Markov chain $\zeta$ introduced at (1-2). Let $N(\zeta)$ be the, possibly infinite, stopping time obtained in this way. The next lemma provides the key to the martingale property, which relies on expectation being preserved. Later, in Lemma 14.1, a more general result is given which applies to simple optional lines, rather than very simple ones, but which depends more heavily on the measure change argument.

**Lemma 6.1** *When $\mathfrak{L}$ is a very simple optional line $E_\mathbb{B}[W_\mathfrak{L}] = W_0$ if and only if $N(\zeta) < \infty$ almost surely.*

Lemma 6.1 (or its generalization Lemma 14.1) is useful for checking the hypothesis that expectations are preserved in the next theorem. Recall that $W$ is the almost sure limit of the martingale $(W_n, \mathcal{G}_n)$. Let $\mathfrak{L}_n$ be the (function corresponding to the) line formed by members of $\mathfrak{L}$ in the first $n$ generations and the $n$th generation nodes with no ancestor in $\mathfrak{L}$.

**Theorem 6.2** *Let $\{\mathfrak{L}[t] : t \geq 0\}$ be simple optional lines that are increasing with $t$ and satisfy $E_\mathbb{B}\left[W_{\mathfrak{L}[t]}\right] = W_0$ for every $t$. Then $(W_{\mathfrak{L}[t]}, \mathcal{G}_{\mathfrak{L}[t]})$ is a positive martingale. If, for each $n$, $W_{\mathfrak{L}[t]_n}$ tends to $W_n$, almost surely, as $t \to \infty$ then $W_{\mathfrak{L}[t]}$ converges to $W$ almost surely.*

The final lemma in this section gives a fairly simple necessary condition for $W_{\mathcal{L}[t]_n} \to W_n$, which is one of the hypotheses of Theorem 6.2.

**Lemma 6.3** *Let $\{\mathcal{L}[t] : t \geq 0\}$ be (not necessarily simple) optional lines that are increasing with $t$. If $H(S(\nu))\mathcal{L}[t](\nu) \to 0$ as $t \to \infty$ for each $\nu$ with $|\nu| \leq (n-1)$ then $W_{\mathcal{L}[t]_n} \to W_n$.*

The results in Theorem 6.2 are easier to establish when the original martingale $\{W_n\}$ converges in mean. However, in our main applications mean convergence does not necessarily hold.

There is an important particular case of increasing lines that sets the scene for a discussion of the general branching process. Let

$$\mathcal{I}[t](\nu) = I(H(S(\nu)) < e^{-t} \text{ but } H(S(\nu_i)) \geq e^{-t} \text{ for } i < |\nu|),$$

which is the (very simple) optional line formed by picking out individuals whose value of $H$ is below $e^{-t}$ but whose antecedents' values are not. Obviously these lines increase with $t$. The following proposition is a straightforward application of previous three results.

**Proposition 6.4** *If $\liminf_n H(\zeta_n) = 0$ then $(W_{\mathcal{I}[t]}, \mathcal{G}_{\mathcal{I}[t]})$ is a martingale converging to $W$.*

# 7 Branching random walk in a random environment

In the original formulation, the basic data are contained in a function from the type space $\mathcal{S}$ into probability laws on $\mathcal{S}^{\mathbb{N}}$, giving the family distributions $\{P_s : s \in \mathcal{S}\}$. Denote the set of such functions by $\mathfrak{L}$. In a sense, the collection of family distributions, that is the element of $\mathfrak{L}$ used, defines the external environment. Thus, a natural generalization is to allow some choice from $\mathfrak{L}$; the varying environment process, already described, can be viewed in this way. In a random environment the elements from $\mathfrak{L}$ used in successive generations usually form a stationary sequence; here a branching random walk with a stationary environment sequence is considered. In this process, the law for the point process $Z$ varies; when that law is $\eta$, let the corresponding expectation be $E^\eta$ and let $m_\eta(\theta) = E^\eta\left[\int e^{-\theta x} Z(dx)\right]$. Let the law used in generation $n$ be $\lambda(n) \in \mathfrak{L}$, where $\lambda = \{\lambda(n)\}$ forms a stationary sequence with the marginal law $P^*$; thus, $\lambda$ is a realization of the random environment. Assume that $m_\eta(\theta)$ is finite and $m'_\eta(\theta)$ exists, $P^*$ almost surely. Finally, denote the conditional branching law given $\lambda$ by $\mathbb{B}$. It should really be something like $\mathbb{B}^\lambda$, but precision is sacrificed to simplicity.

Suppose the environment $\lambda$ is given. Then, again augmenting the type space by the generation,

$$H(n, s) = e^{-\theta s} \prod_{i=0}^{n-1} \frac{1}{m_{\lambda(i)}(\theta)}, \quad H(n, \partial) = 0 \tag{7-1}$$

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

The next lemma is a straightforward application of definitions.

**Lemma 7.1** *Given $\lambda$, $E_{(n,s)}[f(X)] = E^{\lambda(n)}[f(X)]$. Let $\overline{E}$ be the expectation over $\lambda$, 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 $\mathbb{B}$ is a conditional law, the conclusions are conditional ones, holding almost surely as $\lambda$ varies over realisations.

**Theorem 7.2** *Assume that the environment $\lambda$ 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 $E_{\mathbb{B}}W = 0$.*

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

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

*Proof.* Let the real random variable $Y_\eta$ have the law with density $e^{-\theta x}/m_\eta(\theta)$ with respect to $\mu_\eta$. In the same way as in Section 4,

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

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 $\kappa$ is less than 0, Theorem 2.1(iii) applies to show that $W$ is zero, proving (i).

When $\kappa$ is greater than zero, $H(\zeta_n)$ is eventually contained in an interval of the form $(0, d^n)$, with $d < 1$. Then (2-1) in Theorem 2.1 is finite when

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

and, by Lemma 7.1,

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

which is finite when $\int E^{\eta}\left[X \log X\right] P^*(d\eta) < \infty$. 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, Theorem 2.1(iii) again shows that $W$ is zero.

When $0 < \kappa < \infty$, $H(\zeta_n)$ is ultimately contained in an interval of the form $(d^n, \infty)$, with $d < 1$. Then, using Lemma 7.1, the series in (2-2) in Theorem 2.1 is infinite when

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

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

In fact, when $\kappa = 0$ the conditions can be relaxed. It is be enough that $(\lambda(i), Y_{\lambda(i)})$ is a Harris chain, for then $-\log H(\zeta_n)$ is a zero mean random walk when sampled at visits to the 'base' state.

The example can be taken further, allowing reproduction to depend on the node, not just its generation. Each node $\nu$ has a law $\lambda(\nu)$ attached to it, with $\lambda$ 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 $\lambda$, and the arguments leading to Theorem 7.2 continue to apply, but the martingale $W_n$ is probably too complicated to be interesting.

This example is fairly simple because the mean-harmonic function, given at (7-1), 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.

## 8 Multitype branching random walk

In this process people have labels drawn from $\Sigma$ and their reproduction is defined by a point process on $\Sigma \times \mathbb{R}$ with a distribution depending on the label of the parent. These labels are usually called types, but, for clarity, we want to reserve 'types' for members of $\mathcal{S}$. The first component of the point process determines the distribution of that child's reproduction point process and the second component gives the child's birth position relative to the parent's. In the notation here, the type space is $\mathcal{S} = \Sigma \times \mathbb{R}$.

Let $Z$ be a reproduction point process, with points $\{(\sigma_i, z_i)\}$ and let $\widetilde{P}_\sigma$ and $\widetilde{E}_\sigma$ be the probability and expectation associated with reproduction from a parent with label $\sigma \in \Sigma$. Bearing in mind the reduction described at the end of Section 4, suppose there is non-negative function $\widetilde{H}$ on $\Sigma$ such that

$$\widetilde{E}_\sigma \left[ \sum_i \widetilde{H}(\sigma_i) e^{-z_i} \right] = \widetilde{H}(\sigma).$$

Then $H(\sigma, z) = \widetilde{H}(\sigma) e^{-z}$ is mean-harmonic and then, when the initial ancestor has label $\sigma$, $X = \widetilde{H}(\sigma)^{-1} \sum_i \widetilde{H}(\sigma_i) e^{-z_i}$.

Write $\zeta_n = (\tilde{\zeta}_n, S_n) \in \Sigma \times \mathbb{R}$ to identify the two components of $\zeta$. Then $(\tilde{\zeta}_n, S_n - S_{n-1})$ is also a Markov chain. We assume that this latter Markov chain has a stationary distribution under which it is ergodic and that this stationary distribution can be written in the form $H(\sigma)p(\sigma, dz)\nu(d\sigma)$, where $\nu$ is a suitable reference measure on $\Sigma$. Let $\pi(d\sigma) = H(\sigma)\nu(d\sigma)$ be the stationary distribution for $\tilde{\zeta}$. Assume $\beta = \int\int zp(\sigma, dz)\pi(d\sigma)$ is well-defined. Let $Y = \sum_i \widetilde{H}(\sigma_i)e^{-z_i}$, so that $E_\sigma \widetilde{H}(\sigma)X = E_\sigma Y$.

**Proposition 8.1** *If $\beta > 0$ and*

$$\int \widetilde{E}_\sigma \left[Y \log^+ Y\right] I(\widetilde{H}(\sigma) > 0)\nu(d\sigma) < \infty \tag{8-1}$$

*then $W_n$ converges in mean for a set of initial labels (in $\Sigma$) with probability one under $\pi$.*

*Proof.* Since $H(\zeta_n) = \widetilde{H}(\tilde{\zeta}_n)e^{-S_n}$, (2-1) becomes

$$\sum_{i=1}^{\infty} \frac{E_{\tilde{\zeta}_i}\left[Y((e^{-S_i}Y) \wedge 1)\right]}{\widetilde{H}(\tilde{\zeta}_i)} < \infty$$

If, $\zeta_0$ is drawn from the stationary distribution then $\{S_n\}$ is an ergodic sequence and $S_n/n \to \beta$ almost surely, $e^{-S_i}$ can be bounded above by an exponentially decaying sequence, and then (2-1) holds when (8-1) holds. □

If the chain $(\tilde{\zeta}_n, S_n - S_{n-1})$ has a stationary distribution with the appropriate properties except that is not ergodic the result will still hold provided the expectation of $S_1 - S_0$ under the stationary measure with respect to the tail $\sigma$-algebra is bounded below by a positive constant, almost surely.

Kyprianou and Rahimzadeh Sani (2001) discuss martingale convergence for multitype branching random walk with finite $\Sigma$ using the measure change argument.

## 9 General branching processes

Olofsson (1998) uses the change of measure argument in the context of the general branching process; the theory developed on optional lines, provides the link between the general results here and that framework.

Consider a homogeneous branching random walk, as described in Section 4, with $E\int e^{-\alpha t}Z(dt) = 1$ for some $\alpha > 0$ and $\beta = E\int te^{-\alpha t}Z(dt) \geq 0$; 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 $W$. Then

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

is the line formed by picking out individuals born to the right of $t$ but with all their antecedents born to the left $t$, which was introduced at the end of Section 5. It is just a rewriting of $\mathcal{I}[\alpha t]$, introduced at the end of Section 6, for this model, The general branching process arises when $Z$ gives birth times and so is concentrated on $(0, \infty)$; then $\mathcal{C}[t]$ is called the coming generation.

**Proposition 9.1** *$W_{\mathcal{C}[t]}$ is a martingale converging to $W$.*

*Proof.* Here, $\zeta$ is a random walk with mean $\beta \geq 0$; hence, $\liminf_n e^{-\zeta_n} = 0$. Therefore Proposition 6.4 applies. $\square$

Let $X = \int e^{-\alpha t} Z(dt)$. Theorem 7.2 gives (i) $E_{\mathbb{B}} W = 1$ if $\beta > 0$ and $E[X \log X] < \infty$ and (ii) $E_{\mathbb{B}} W = 0$ if $\beta = 0$ or both $\beta < \infty$ and $E[X \log X] = \infty$. When combined with Proposition 9.1, this includes the conclusion of Theorem 2.1 in Olofsson (1998), which deals with the case where $Z$ is concentrated on $(0, \infty)$ and $0 < \beta < \infty$; that paper should be consulted for references to earlier treatments of this result and more context on the general branching process.

The multitype branching random walk can be tackled in a similar way. All the general assumptions and notation of Section 8 apply here. In particular, $\pi$ is the stationary distribution for $\tilde{\zeta}$, where $\zeta_n = (\tilde{\zeta}_n, S_n) \in \Sigma \times \mathbb{R}$. As previously, let $\mathcal{C}[t]$ be the line formed by picking out individuals born to the right of $t$ but with all their antecedents born to the left $t$. Note that $\mathcal{C}[t]$ is no longer the same as $\mathcal{I}[t]$, introduced in Section 6, since the latter is defined in terms of $H$.

**Proposition 9.2** *If $\beta > 0$, $W_{\mathcal{C}[t]}$ is a martingale converging to $W$ almost surely for a set of initial types (in $\Sigma$) with probability one under $\pi$.*

*Proof.* To use Lemma 6.1 we need to show that $\liminf e^{-S_n} = 0$; this is so under the stationary distribution for $\zeta$, for then $S_n$ is ergodic, and hence for a set of initial types (in $\Sigma$) with probability one under $\pi$. Lemma 6.3 applies to complete the verification of the hypotheses of Theorem 6.2. $\square$

As in the one-type case, when the random walk is actually on $(0, \infty)$ these positions can be interpreted as birth times, $\beta > 0$ is automatic, and then Propositions 8.1 and 9.2 becomes results about the general branching process on a general type space. Much further discussion of this process can be found in Jagers (1989). The results here have close connections with Theorems 6.1 and 6.5 there.

## 10 Branching random walk with a barrier

The homogeneous branching random walk based on the point process $Z$ with points at $\{z_i\}$, described in Section 4, is 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). The results are developed to use in the study of the derivative martingale, but they may be of independent interest.

Formally, the branching random walk with a barrier 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 $\partial$ and the associated details are omitted from the discussion and sums over $|\nu| = n$ are over the nodes not of type $\partial$.

Let the intensity measure of $Z$ be $\mu$ and assume that (5-1) holds. Now let $Y_n$ be independent identically distributed variables with their law having density $e^{-x}$ with respect to $\mu$ 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 condition for the weak descending ladder variable to have a finite mean is supplied by Doney (1980). 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 (2003), which develops the random walk results needed to prove Theorem 10.6 stated at the end of this section.

**Lemma 10.1** *(i) As $x \to \infty$, $V(x)/x$ converges to a positive constant, which is finite if $\int x^2 I(x < 0)e^{-x}\mu(dx) < \infty$. (ii) When $V(x)/x$ has a finite limit, $a(x+1) \le V(x) \le b(x+1)$for suitable $a > 0$ and $b < \infty$. (iii) $E[V(Y_1 + s)I(Y_1 + s > 0)] = V(s)$.*

**Lemma 10.2** *When (5-1) 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 10.1(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 Markov chain $\zeta$ associated with this harmonic function is considered next.

**Lemma 10.3** *Transitions of the Markov chain $\zeta$ from $s$ have 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 (1-2). $\square$

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). 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 $\zeta$, which will be described in Theorem 10.7, but first the following simple consequence of his results is recorded. Technically, Tanaka's construction gives the result when $\zeta_0 = 0$; the extension to other starting states is covered in Biggins (2003).

**Lemma 10.4** *$\zeta_n \to \infty$ as $n \to \infty$, almost surely.*

This is enough for the application of the ideas on optional lines. In the same way as in Section 9, $\mathcal{C}[t]$ is the line formed by picking out individuals born to the right of $t$ but with all their antecedents born to the left of $t$.

**Theorem 10.5** *$(W_{\mathcal{C}[t]}, \mathcal{G}_{\mathcal{C}[t]})$ is a martingale. It converges to $W$, which is the limit of the martingale $(W_n, \mathcal{G}_n)$.*

*Proof.* As before, $\mathcal{C}[t]$ are increasing very simple optional lines, and so in Lemma 6.1, $N$ is the first time $\zeta$ exceeds $t$; Lemma 10.4 now shows that $N(\zeta) < \infty$ almost surely. Furthermore, $V(S(\nu))e^{-S(\nu)}I(S(\nu) > t) \to 0$ as $t \to \infty$ and so Lemma 6.3 applies. Hence Theorem 6.2 applies to give the result. $\square$

The next result is the main one about the martingale $W_n$.

**Theorem 10.6** *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 $s$, $E[\tilde{X}_3(s) L_4(\tilde{X}_3(s))]$ is infinite then $W_n \to 0$ almost surely.*

The proof is an application of Corollary 2.5(ii), Corollary 2.7 and the following series of results. It is assumed throughout the remainder of this section that, in addition to (5-1), $\int x^2 e^{-x}\mu(dx) < \infty$.

First, the simple Lemma 10.4 needs to be supplemented by information on how fast $\zeta$ goes to infinity; the following result, taken from Biggins (2003), provides relevant estimates. It concerns the growth of $D(x) = \sum_n I(\zeta_n \leq x)$.

**Theorem 10.7** *Let $\varphi(x) = \log\log x$ for $x > 3$. For suitable (non-random) $L$ and $U$*

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

*almost surely.*

One consequence of this, or Lemma 10.4, is that in applying the second parts of Theorem 2.4 and Corollary 2.5 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 10.7, 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 10.8** *Let $\tilde{D}(x) = \sum_n V(\zeta_n)^{-1} I(\zeta_n \le x) = \int_0^x V(z)^{-1} D(dz)$ and, as previously, let $\varphi(x) = \log\log x$ for $x > 3$. For suitable (non-random) $\tilde{L}$ and $\tilde{U}$*

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

*Proof.* Lemma 10.1(ii) easily yields $\tilde{D}(x) \ge D(x)/(b(x+1))$ and then the lower bound follows immediately from the lower bound in Theorem 10.7.

Let $U$ be the constant in the upper bound in Theorem 10.7 and let $\epsilon > 0$. Let $D^*(x) = (U+\epsilon)x^2\varphi(x)$ for $x > x_0 > e^e$ and $D^*(x) = D^*(x_0)$ otherwise, with $x_0$, which is random, large enough that $D(x) \le D^*(x)$ for all $x \ge 0$. Then, using Lemma 10.1(ii) and Fubini,

$$\begin{aligned}
\tilde{D}(x) &= \int_0^x V(z)^{-1} D(dz) \\
&\le \int_0^x \frac{1}{a(z+1)} D(dz) \\
&= \frac{1}{a}\int_0^x \left(\int_z^\infty \frac{dy}{(y+1)^2}\right) D(dz) \\
&= \frac{1}{a}\int_0^\infty D(x\wedge y)\frac{dy}{(y+1)^2} \\
&\le \frac{1}{a}\int_0^\infty D^*(x\wedge y)\frac{dy}{(y+1)^2} \\
&= \frac{1}{a}\int_0^x \frac{1}{z+1} D^*(dz).
\end{aligned}$$

Substituting for $D^*$ and recalling that $x_0 > e^e$ gives, for $x > x_0$,

$$\begin{aligned}
\tilde{D}(x) &\le \frac{1}{a}\int_0^x \frac{1}{z+1} D^*(dz) \\
&= \frac{D^*(x_0)}{a} + \frac{U+\epsilon}{a}\int_{x_0}^x \frac{1}{z+1}\left(2z\log\log z + \frac{z}{\log z}\right) dz \\
&\le \frac{D^*(x_0)}{a} + \left(\frac{U+\epsilon}{a}\right) 3x\log\log x,
\end{aligned}$$

which produces the upper bound. □

The next result derives suitable bounding variables for use in Corollary 2.5. It is here that the gap, mentioned already, between using $\tilde{X}_2$ in the upper bound but $\tilde{X}_3(s)$ as a lower bound arises.

**Lemma 10.9** *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 any fixed $s_0$ and $s \ge 2s_0$,*

$$X \ge \frac{a}{2b}\tilde{X}_3(s_0).$$

*Proof.* When the parent is at $s$, applying Lemma 10.1(ii),

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

as required. Similarly

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

and, for $s > 2s_0$,

$$\begin{aligned}
X &\ge \frac{\sum_j a(z_j+s+1)e^{-z_j}I(z_j>-s)}{b(s+1)} \\
&\ge \frac{\sum_j a(z_j+s+1)e^{-z_j}I(z_j>-s/2)}{b(s+1)} \\
&\ge \frac{a}{2b}\sum_j e^{-z_j}I(z_j>-s_0) = \frac{a}{2b}\tilde{X}_3(s_0).
\end{aligned}$$

□

Translating the first of these bounds into the language of Theorem 2.6 and Corollary 2.7, $X_1^* = b\tilde{X}_1$, $g_1(s) = V(s)^{-1}$, $X_2^* = b\tilde{X}_2/a$ and $g_2(s) = 1$. Thus the associated functions are

$$A_1(x) = \sum_{i=1}^{\infty} V(\zeta_i)^{-1} I(H(\zeta_i)x \ge V(\zeta_i)) \text{ and } A_2(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x \ge 1).$$

Exactly the same functions are needed in applying Corollary 2.5(ii). The next lemma makes comparisons between $A_1$ and $A_2$ and suitable slowly varying functions.

**Lemma 10.10**

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

$$\limsup_{x\to\infty}\frac{A_2(x)}{(\log x)^2\log\log\log x}<\infty \quad \textit{and} \quad \liminf_{x\to\infty}\frac{A_2(x)}{(\log x)^2/\log\log\log x}>0.$$

*Proof.* Since

$$\begin{aligned} A_1(x) &= \sum_{i=1}^{\infty} V(\zeta_i)^{-1} I(H(\zeta_i)x \geq V(\zeta_i)) \\ &= \sum_{i=1}^{\infty} V(\zeta_i)^{-1} I(\zeta_i \leq \log x), \end{aligned}$$

the estimates in Lemma 10.8 translate immediately into the stated estimates of $A_1$. For the second part, note that

$$\begin{aligned} A_2(x) = \sum_{i=1}^{\infty} I(H(\zeta_i)x \geq 1) &= \sum_{i=1}^{\infty} I(V(\zeta_i)e^{-\zeta_i}x \geq 1) \\ &= \sum_{i=1}^{\infty} I(\zeta_i \leq \log x + \log V(\zeta_i)). \end{aligned}$$

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

$$\sum_{i=1}^{\infty} I(\zeta_i \leq \log x) \leq A_2(x) \leq \sum_{i=1}^{\infty} I\left((1-\epsilon)\zeta_i \leq \log x + \gamma\right),$$

that is

$$D(\log x) \leq A_2(x) \leq D((1-\epsilon)^{-1}(\log x + \gamma)).$$

Now, the results in Theorem 10.7 complete the required estimation of $A_2$. □

*Proof of Theorem 10.6.* The upper bounds in Lemma 10.10, the first bound in Lemma 10.9 and Corollary 2.7 combine to prove Theorem 10.6(i). The lower bounds in Lemma 10.10, the second and third bounds in Lemma 10.9 and Corollary 2.5(ii) combine to prove Theorem 10.6(ii). The estimates for $A_1$ gives the moment conditions on $\tilde{X}_1$, while those for $A_2$ gives conditions on $\tilde{X}_2$ and $\tilde{X}_3(s)$. □

## 11 Proofs concerning the derivative martingale

The context here is described in Sections 4 and 5. It is a homogeneous branching random walk satisfying (5-1) and $\int x^2 I(x<0)e^{-x}\mu(dx) < \infty$ in which

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

Before starting the main proof, some results concerning $W_n$ are noted for later use.

**Lemma 11.1** *(i) $E_{\mathbb{B}}W = 0$, (ii) $\mathbb{B}(W = 0) = 1$ and (iii) $\inf\{S(\nu) : |\nu| = n\} \to \infty$, almost surely as $n \to \infty$.*

This result is contained in Lemma 5 and the discussion at the end of Section 3 in Biggins (1977); the second and third assertions are immediate consequences of $E_{\mathbb{B}}W = 0$. Informally, the last part says that every line of descent goes to infinity. A proof that $E_{\mathbb{B}}W = 0$ without the side condition that $\int x^2 I(x < 0)e^{-x}\mu(dx)$ is finite was given by Lyons (1997), can also be obtained fairly directly from Theorem 2.1(iii) and is contained in Theorem 7.2(iii)(a).

*Proof of Theorem 5.1.* Let $\mathcal{E}_{\mathfrak{b}}$ be the event that no node in the branching random walk has a position to the left of $-\mathfrak{b}$; then, by Lemma 11.1(iii), $\mathcal{E}_{\mathfrak{b}}$ increases to an event with probability one as $\mathfrak{b} \to \infty$. Use the homogeneous branching random walk to construct a branching random walk with a barrier at $-\mathfrak{b}$; on $\mathcal{E}_{\mathfrak{b}}$ the processes with and without a barrier agree. To make the coupling precise, let $\tilde{I}_{\mathfrak{b}}(\nu)$ be one if the node $\nu$ is retained in the process with a barrier at $-\mathfrak{b}$ and zero otherwise. Now, by Lemma 10.2, $V(\mathfrak{b})^{-1}\sum_{|\nu|=n} V(\mathfrak{b} + S(\nu))e^{-S(\nu)}\tilde{I}_{\mathfrak{b}}(\nu)$ is a positive martingale, which must converge to a finite limit, denoted by $B_{\mathfrak{b}}$. Hence, using Lemma 10.1(i) and Lemma 11.1,

$$
\begin{aligned}
B_{\mathfrak{b}} &= \frac{1}{V(\mathfrak{b})}\lim_{n\to\infty}\sum_{|\nu|=n} V(\mathfrak{b} + S(\nu))e^{-S(\nu)}\tilde{I}_{\mathfrak{b}}(\nu) \\
&\leq \frac{C}{V(\mathfrak{b})}\lim_{n\to\infty}\sum_{|\nu|=n} (\mathfrak{b} + S(\nu))e^{-S(\nu)}I(\mathfrak{b} + S(\nu) > 0) \\
&= \frac{C}{V(\mathfrak{b})}\lim_{n\to\infty}(W_n\mathfrak{b} + \partial W_n) \\
&= \frac{C}{V(\mathfrak{b})}\lim_{n\to\infty}\partial W_n;
\end{aligned}
$$

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

Let $S^\nu$ be the function $S$ on the sub-tree rooted at $\nu$. Splitting on the first generation shows that

$$\partial W_n(S) = \sum_{|\nu|=1} S(\nu)e^{-S(\nu)}W_{n-1}(S^\nu) + \sum_{|\nu|=1} e^{-S(\nu)}\partial W_{n-1}(S^\nu).$$

Since $\int |x|e^{-x}\mu(dx) < \infty$,

$$\sum_{|\nu|=1} e^{-S(\nu)} < \infty \quad \text{and} \quad \sum_{|\nu|=1} |S(\nu)|e^{-S(\nu)} < \infty$$

almost surely. Now, letting $n$ go infinity, straightforward analysis drawing on Lemma 11.1(i) gives that

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

which is another way to write (5-2). Hence, $\mathbb{B}(\Delta = 0)$ is a fixed point of the generating function of the family size and so must have the stated property.

*Proof of Theorem 5.2.* It has already been shown that $V(\mathfrak{b})B_{\mathfrak{b}} \le C\Delta$ with equality on $\mathcal{E}_{\mathfrak{b}}$. When Theorem 10.6(i) holds, $E_{\mathbb{B}}B_{\mathfrak{b}} = 1$ and then $V(\mathfrak{b}) \le CE_{\mathbb{B}}\Delta$ for any $\mathfrak{b}$; thus $E_{\mathbb{B}}\Delta = \infty$. Similarly, when the conditions of Theorem 10.6(ii) hold $B_{\mathfrak{b}}$, and hence $\Delta$, is zero on $\mathcal{E}_{\mathfrak{b}}$ for every $\mathfrak{b}$. $\square$

*Proof of Theorem 5.3.* Note first that Proposition 9.1 shows that $W_{\mathcal{C}[t]}$ and $W_n$ have the same limit; by Lemma 11.1, this limit is zero and so $\inf\{S(\nu) : \nu \in \mathcal{C}[t]\} \to \infty$ as $t \to \infty$. Now, applying Theorem 10.5 shows that on $\mathcal{E}_{\mathfrak{b}}$

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

Letting $\mathfrak{b} \to \infty$ completes the proof. $\square$

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).

## 12 Measure change and mean convergence

The method used to determine conditions for mean convergence of $W_n$ 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 key idea in all these papers is to exploit a change of measure to establish moment conditions for the martingale to converge in mean; the actual measure change has much longer history, as can be seen from the references in Lyons (1997). 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 the classical Kesten-Stigum Theorem; moreover, their framework, is at first sight, rather different from here and so some points of contact are noted later in this section.

Recall that $\mathfrak{B}_n$ is the projection of the sample space $\mathfrak{B}$ onto the first $n$ generations. A branching process is a Markov chain with a state in $\mathfrak{B}_n$ at time $n$ and transition probabilities defined by the $n$th generation nodes producing independent families, with the distribution of the family of a node of type $s$ being $P_s$. 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 $\mathfrak{B}$; this gives $\mathbb{B}$.

However, to describe the measure change neatly, it is 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 $\xi$ defines a line of descent starting from the initial ancestor. Let $\Xi$ be the set of possible $\xi$. 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 $\xi$ 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 $\sigma$-algebra generated by the first $n$ steps of the Markov chain, that is, the information on the development for the first $n$ generations, including the trunk. Let $\mathcal{F}_n^*$ be the $\sigma$-algebra generated by $\mathcal{F}_{n-1}$ and $\mathcal{G}_n$, so that $\mathcal{F}_n^*$ is generated by the trunk up to generation $n-1$ and the tree up to generation $n$; hence $\mathcal{F}_{n-1} \subset \mathcal{F}_n^* \subset \mathcal{F}_n$.

The new measure depends on $H$, the particular mean-harmonic function under consideration. 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.

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

$$P(\xi_{n+1} = \nu | \mathcal{F}_{n+1}^*) = \frac{h(\nu) I(\nu \in c(\xi_n))}{\sum_{\sigma \in c(\xi_n)} h(\sigma)}, \text{ when } \sum_{\sigma \in c(\xi_n)} h(\sigma) \in (0, \infty), \tag{12-1}$$

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 $\xi$ maps $(\mathfrak{T}, \mathbb{P})$ to $(\mathfrak{B}, \mathbb{B})$. Since the theorems in Section 2 involve $E_{\mathbb{B}} W$ it is worth noting explicitly that $E_{\mathbb{B}} W = E_{\mathbb{P}} W$.

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 $\xi$.

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 states for the first $n$ generations are known. Then, reproduction from $n$th 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 $\mathbb{B}$). On the trunk, when $S(\xi_n) = s$, the types of the children of $\xi_n$ 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 of the children of $\xi_n$, $\xi_{n+1}$ is chosen exactly as previously, that is as in $\mathbb{P}$ — see (12-1). Call the resulting measure $\mathbb{Q}$. To express the derivative more neatly, and for later developments, let $X(\nu)$ be defined as $X$ for the tree initiated by $\nu$. More precisely, let $S^\nu$ be the function $S$ on the sub-tree rooted at $\nu$ and let

$$X(\nu) = X(S^\nu) = I(h(\nu) > 0) \frac{\sum_{\sigma \in c(\nu)} h(\sigma)}{h(\nu)} + I(h(\nu) = 0).$$

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

It is easy to confirm that under $\mathbb{Q}$ the types on the trunk, given by $\zeta_n = S(\xi_n)$, develop as a Markov chain on $\mathcal{S}^H$ with the transition kernel (1-2) when the initial ancestor has a type in $\mathcal{S}^H$.

The approach is based on the following theorem which is a corollary of a result in Durrett (1996, Theorem 4.3.3) — see also Athreya (2000). The notation employed suggests how the result will be used.

**Theorem 12.1** *Suppose $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures and $\mathcal{G}_n$ are increasing $\sigma$-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) $E_{\mathbb{P}}W = 1$ if and only if $\mathbb{Q}(W < \infty) = 1$;*
*(iii) $E_{\mathbb{P}}W = 0$ 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) on $\mathcal{G}_n$; clearly, if $\mathbb{Q}$ is tractable it can be used to study the mean convergence of the original martingale through the last two parts of the lemma. Note that this measure change only concerns $\mathbb{P}$ and $\mathbb{Q}$ on the $\sigma$-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 $h(\nu) = 0$. Therefore, 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, when $H(s) = 0$, $E_{\mathbb{P}}\left[X(\nu)|S(\nu) = s\right] = 1$ by definition. By exploiting the trunk, a simpler martingale than $W_n$ can be constructed, by forming a product (down $\xi$) using these adapted positive terms with expectation one. In fact it is useful to define these products for any node. To do this, recall that $\{\nu_i : i = 0, 1, \ldots, |\nu|\}$ is the ancestry of $\nu$. 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 12.1, $\mathbb{P}$ 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 12.2** *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 12.3** *$\overline{W}(\xi_n)$ 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 $\xi_r$ use the same law as in $\mathbb{P}$ and $\xi_r$ 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 12.2,

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

as required. □

Recall that $\mathcal{G}_n$ is the $\sigma$-algebra generated by the first $n$ generations without the trunk. The idea now is to integrate out $\xi$ 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 (12-1) 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 12.4** *For a fixed $\nu$, let $n = |\nu| \le r$. Then*

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

*Proof.* Fix $r$. The result is true for $n=0$. Suppose it is true for $(n-1)$. Let $\sigma\{\mathcal{F}^*_n, \mathcal{G}_r\}$ be the $\sigma$-algebra generated by the two components; recall that $\mathcal{F}^*_n$ is information on the first $n-1$ generations including the trunk and on the $n$ generation without the trunk. Then

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

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 $(n-1)$,

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

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 (12-1),

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

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

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

$h(\nu) = 0$ almost surely. Hence the formula holds in both these cases. Finally,

$$E_{\mathbb{P}}[h(\nu_{n-1})X(\nu_{n-1})] \le E_{\mathbb{P}}[W_{n-1}] = h(0) < \infty$$

so that $I(h(\nu_{n-1})X(\nu_{n-1}) = \infty)$ is $\mathbb{P}$-null. □

**Proposition 12.5** *$W_n/h(0)$ is the derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{G}_n$. Hence Theorem 12.1 applies.*

*Proof.* By Lemma 12.3, the derivative sought equals $E_{\mathbb{P}}[\overline{W}(\xi_n)|\mathcal{G}_n]$. For a fixed $\nu$ 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 12.4; thus

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

□

The next theorem underpins all the results given in Section 2, but, unlike them, it requires knowledge of the measure change that goes beyond defining the Markov chain $\zeta$.

Let $c'(\xi_n)$ be the children of $\xi_n$ excluding $\xi_{n+1}$ 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.

**Theorem 12.6**

*(i) If*

$$\mathbb{Q}\left(\liminf_n h(\xi_n) < \infty,\ \sum_{i=1}^{\infty} h(\xi_i)X'(\xi_i) < \infty\right) > 0 \tag{12-2}$$

*or*

$$\mathbb{Q}\left(\sum_{i=1}^{\infty} h(\xi_i)X(\xi_i) < \infty\right) > 0, \tag{12-3}$$

*which implies (12-2), 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 (12-2) or (12-3) is one.*

*(ii) If*

$$\mathbb{Q}\left(\limsup_n h(\xi_n)X(\xi_n) = \infty\right) > 0 \tag{12-4}$$

*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.* Recall that $S^\nu$ is the function $S$ on the sub-tree rooted at $\nu$. 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 $\sigma$-algebra generated by the reproduction of the members of the trunk. (Technically, in the language of Jagers (1989), with $\mathcal{L}$ the optional line formed by all non-trunk children of the nodes forming the trunk, $\mathcal{H}$ is the pre-$\mathcal{L}$ $\sigma$-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 (12-3) implies (12-2) and either implies that $\mathbb{Q}(W < \infty) > 0$; in addition, $\mathbb{Q}(W < \infty) = 1$ when either probability is one. Theorem 12.1 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

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

A further application of Theorem 12.1 completes the proof. □

Waymire and Williams (1996) consider multiplicative cascades, which are described briefly at the end of Section 4, 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 $m(1) = b$. 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 for the cascade, corresponding directly to (4-1) with $\theta = 1$. Now Theorem 12.6 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.

Observe that $h(\xi_n)$, $X'(\xi_{n-1})$ and $X(\xi_{n-1})$ are all $\mathcal{F}_n$-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 Theorem 12.6(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, Theorem 2.1.

**Lemma 12.7** *Suppose $Y_i$ are non-negative variables that are adapted to the increasing $\sigma$-algebras $\mathcal{F}_i$. Then*

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

**Lemma 12.8** *Let $Y$ be a non-negative function on $\mathfrak{B}_1$ and $Y(\nu)$ the corresponding function of the reproduction from node $\nu$. 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 $g$, $E_{\mathbb{Q}}[g(X(\xi_i))|\mathcal{F}_i] = E_{\zeta_i}[Xg(X)]$*

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

*Proof of Theorem 2.1.* Apply Lemma 12.7 and Lemma 12.8 to the series in (12-3) for the first part; for the second part, conditional Borel-Cantelli and Lemma 12.8 show that (2-2) implies that the probability in (12-4) is one. Since $H(\zeta_n) = H(S(\xi_n)) = h(\xi_n)$, the final part follows from Theorem 12.6(ii) and the inequality $h(\xi_i)X(\xi_i) \geq h(\xi_{i+1})$. □

*Proof of Proposition 2.2.* First, integrate out $\xi_1$ to get, for any non-negative $g$,

$$E_{\mathbb{Q}}[g(X')|\mathcal{G}_1] = \left(\frac{1}{\sum_{|\nu|=1} h(\nu)}\right) \sum_{|\nu|=1} g\left(\frac{\sum_{|\sigma|=1} h(\sigma) - h(\nu)}{h(0)}\right) h(\nu)$$

Then, with this as $Y$, Lemma 12.8 gives,

$$\begin{aligned} E_{\mathbb{Q}}[g(X'(\xi_n))|\mathcal{F}_n] &= E_{\mathbb{Q}}[Y(\xi_n)|\mathcal{F}_n] \\ &= E_{\zeta_n}[XY] \\ &= \frac{E_{\zeta_n}\left[\sum_i H(f_i) g\left(\sum_{j\neq i} H(f_j)/H(\zeta_n)\right)\right]}{H(\zeta_n)}. \end{aligned}$$

Now apply Lemma 12.7 to the series in (12-2) to give the first part. Since

$$h(\xi_i)\left(X(\xi_i) - X'(\xi_i)\right) = h(\xi_{i+1}) = H(\zeta_{i+1})$$

convergence of $\sum_n H(\zeta_n)$ implies that the series in (12-2) and (12-3), and hence (2-1) and (2-3), converge together. □

*Proof of Proposition 2.3.* This is like that of Theorem 2.1, but now translating Theorem 12.6 when the probabilities in (12-2), (12-3) and (12-4) are positive, rather than one. □

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 the discussion here takes multitype branching process.

## 13 Stochastic domination

The conditions in Theorem 2.1 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 proofs are independent of the discussion in Section 12. The next two elementary lemmas establish the framework. The first is well-known and proved by integration by parts.

**Lemma 13.1** *If for all $x$ $P(X > x) \leq P(Y > x)$ then, for any increasing non-negative function $f$, $E[f(X)] \leq E[f(Y)]$.*

**Lemma 13.2** *Suppose $\eta$ is a measure on $(0, \infty)$ and $A(x) = \eta(0, x]$. Then*

$$\int ((x/y) \wedge 1)\eta(dy) = \int_1^\infty \frac{A(wx)}{w^2} dw.$$

*Proof.*

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

as required. □

*Proof of Theorem 2.4.* Since, for $h > 0$, $x((hx) \wedge 1) = x(I(hx \geq 1) + hxI(hx < 1))$ is an increasing function of $x > 0$, applying Lemma 13.1 shows that, when $\zeta_i \in F$,

$$E_{\zeta_i} \left[ X((H(\zeta_i)X) \wedge 1) \right] \leq E \left[ g(\zeta_i) X^* ((H(\zeta_i) g(\zeta_i) X^*) \wedge 1) \right],$$

where the expectation on the right is with respect to $X^*$ only. By assumption, $\zeta_i \in F$ eventually, with probability one, and so (2-1) in Theorem 2.1 holds when

$$E \left[ \sum_i g(\zeta_i) X^* ((H(\zeta_i) g(\zeta_i) X^*) \wedge 1) \right] < \infty.$$

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

$$E \left[ \sum_i g(\zeta_i) X^* ((H(\zeta_i) g(\zeta_i) X^*) \wedge 1) \right] = E \int X^* ((X^*/y) \wedge 1) \eta(dy).$$

Applying Lemma 13.2 completes the proof of (i).

In a similar way, considering the series in (2-2),

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

giving (ii). □

*Proof of Corollary 2.5.* 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 $\epsilon$ 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 (2-4) 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. Now, suppose that $A(x)/(x^\delta L(x))$ 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 (2-5). □

*Proof of Theorem 2.6.* For any positive $h$ and $x_j$, crude bounding gives,

$$\left(\sum_j x_j\right)\left(\left(h\sum_j x_j\right)\wedge 1\right) \leq J^2\sum_j x_j((hx_j)\wedge 1)$$

and so, for any $s\in F$,

$$E_s\left[X((hX)\wedge 1)\right] \leq J^2\sum_j E\left[g_j(s)X_j^*((hg_j(s)X_j^*)\wedge 1)\right].$$

Hence, using this and Lemma 13.1, for $\zeta_i\in F$

$$\begin{aligned}E_{\zeta_i}\left[X((H(\zeta_i)X)\wedge 1)\right] &\leq E\left[\left(\sum_j g_j(\zeta_i)X_j^*\right)\left(\left(H(\zeta_i)\sum_j g_j(\zeta_i)X_j^*\right)\wedge 1\right)\right] \\ &\leq J^2\sum_j E\left[(g_j(\zeta_i)X_j^*)((H(\zeta_i)(g_j(\zeta_i)X_j^*))\wedge 1)\right].\end{aligned}$$

Hence (2-1) in Theorem 2.1 holds when the sum over $i$ of each of the terms on the right here is finite. These translate to tests on the $A_j$ as in the proof of Theorem 2.4(i). □

*Proof of Corollary 2.7.* This uses the same method as that used for Corollary 2.5(i). □

## 14 Proofs relating to Optional lines

The proof of the first result here, which is the promised generalization of Lemma 6.1, relies in an important way on the measure change discussed in Section 12; the rest of the discussion is independent of Section 12 except for the notational convention that $h(\nu)$ is $H(S(\nu))$ and the assumption that the initial type is in $\mathcal{S}^H$ so that $W_0$, which is also $h(0)$, is positive.

Let $N$ be the generation in which $\xi$ hits the line $\mathfrak{L}$; more precisely, define $N$ by $\mathfrak{L}(\xi_N) = 1$, with $N = \infty$ when there is no such $N$. Often it will be easy to see when $N$ is finite under $\mathbb{Q}$. This definition is consistent with that of $N(\zeta)$ used in Lemma 6.1.

It is worth noting that the definition of an optional line used in Kyprianou (2000) to prove a particular case of Lemma 14.1 is intermediate between those of simple and very simple.

**Lemma 14.1** *When $\mathfrak{L}$ is a simple optional line, $\mathbb{Q}(N < \infty) = E_{\mathbb{B}} W_{\mathfrak{L}}/h(0)$, and so $E_{\mathbb{B}}[W_{\mathfrak{L}}] = h(0)$ 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 12.3 to move from $E_{\mathbb{Q}}$ to $E_{\mathbb{P}}$; conditioning on $\mathcal{G}_n$ and using that $\mathfrak{L}$ is a simple optional line; and, finally, using Lemma 12.4.

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

Summing over $n$ now gives the result since $E_{\mathbb{B}}[W_{\mathfrak{L}}] = E_{\mathbb{P}}[W_{\mathfrak{L}}]$. □

Recall that $\mathfrak{L}_n$ is the line formed by members of $\mathfrak{L}$ in the first $n$ generations and the $n$th generation nodes with no ancestor in $\mathfrak{L}$.

**Lemma 14.2** *For any (not necessarily simple) optional line $\mathfrak{L}$, $E_{\mathbb{B}}[W_n | \mathcal{G}_{\mathfrak{L}}] = W_{\mathfrak{L}_n}$.*

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

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

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

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

□

In general $\mathfrak{L}_n$ need not be optional when $\mathfrak{L}$ is and so $W_{\mathfrak{L}_n}$ 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 14.3** *Let $\mathfrak{L}$ be a simple optional line. Then $\mathfrak{L}_n$ is a simple optional line and $(W_{\mathfrak{L}_n}, \mathcal{G}_n)$ is a positive martingale with a limit at least $W_{\mathfrak{L}}$. When $E_{\mathbb{B}}[W_{\mathfrak{L}}] = h(0)$, (i) $W_{\mathfrak{L}_n} = E_{\mathbb{B}}[W_{\mathfrak{L}}|\mathcal{G}_n]$, (ii) the martingale $(W_{\mathfrak{L}_n}, \mathcal{G}_n)$ converges in mean to $W_{\mathfrak{L}}$, and (iii) $E_{\mathbb{B}}[W_n|\mathcal{G}_{\mathfrak{L}}] = E_{\mathbb{B}}[W_{\mathfrak{L}}|\mathcal{G}_n]$.*

*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 $n$th generation that are neither in $\mathfrak{L}$ nor have an ancestor in $\mathfrak{L}$, so that

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

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

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

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

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

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

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

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

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

$E_{\mathbb{B}}[W_{\mathfrak{L}}] = h(0)$ forces equality here, which in turn implies that $W_{\mathfrak{L}_n}$ converges to $W_{\mathfrak{L}}$. Hence, $W_{\mathfrak{L}_n} = E_{\mathbb{B}}[W_{\mathfrak{L}}|\mathcal{G}_n]$ and, by Lemma 14.2, $W_{\mathfrak{L}_n} = E_{\mathbb{B}}[W_n|\mathcal{G}_{\mathfrak{L}}]$, proving (iii). □

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

**Lemma 14.4** *Let $\mathfrak{L}'$ and $\mathfrak{L}$ be simple optional lines with $\mathfrak{L}' \le \mathfrak{L}$ and $E_{\mathbb{B}}[W_{\mathfrak{L}}] = h(0)$. Then $E_{\mathbb{B}}[W_{\mathfrak{L}}|\mathcal{G}_{\mathfrak{L}'}] = W_{\mathfrak{L}'}$.*

*Proof.* Let $N'$ and $N$ be the generations where $\xi$ hits $\mathfrak{L}'$ and $\mathfrak{L}$ respectively. Then $N' \le N$ and so, by Lemma 14.1, $E_{\mathbb{B}}[W_{\mathfrak{L}}] = h(0)$ implies that $E_{\mathbb{B}}[W_{\mathfrak{L}'}] = h(0)$. Since $\mathfrak{L}' \le \mathfrak{L}$, $\mathcal{G}_{\mathfrak{L}'} \subset \mathcal{G}_{\mathfrak{L}}$ and so, by Lemma 14.2,

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

Letting $n$ go to infinity and applying Lemma 14.3(ii) completes the proof. $\square$

*Proof of Lemma 6.3.* Let $\mathcal{G}_*$ be the $\sigma$-algebra generated by $\{\mathcal{G}_{\mathfrak{L}[t]} : t \geq 0\}$. Lemma 14.2 implies that $(W_{\mathfrak{L}[t]_n}, \mathcal{G}_{\mathfrak{L}[t]})$ is a positive martingale, and so $W_{\mathfrak{L}[t]_n}$ converges as $t \to \infty$, to $E_{\mathbb{B}}\left[W_n | \mathcal{G}_*\right]$. Now

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

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,

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

which implies that $E_{\mathbb{B}}\left[W_n | \mathcal{G}_*\right] = W_n$, as required. $\square$

*Proof of Theorem 6.2.* The martingale property follows immediately from Lemma 14.4. Let $W'$ be the limit of $W_{\mathfrak{L}[t]}$. By Lemma 14.3(i), $E_{\mathbb{B}}\left[W_{\mathfrak{L}[t]} | \mathcal{G}_n\right] = W_{\mathfrak{L}[t]_n}$; letting $t \to \infty$, Fatou gives $E_{\mathbb{B}}\left[W' | \mathcal{G}_n\right] \leq W_n$ and then letting $n \to \infty$ gives $W' \leq W$. Again, let $\mathcal{G}_*$ be the $\sigma$-algebra generated by $\{\mathcal{G}_{\mathfrak{L}[t]} : t \geq 0\}$. By Lemma 14.3(iii), $E_{\mathbb{B}}\left[W_{\mathfrak{L}[t]} | \mathcal{G}_n\right] = E_{\mathbb{B}}\left[W_n | \mathcal{G}_{\mathfrak{L}[t]}\right]$; letting $n$ and then $t$ go to infinity shows that $W' \geq E_{\mathbb{B}}\left[W | \mathcal{G}_*\right]$. Hence $E_{\mathbb{B}}\left[W' - W\right] \geq 0$, but $W' \leq W$. Hence $W' = W$, completing the proof. $\square$

The conclusions of Theorem 6.2 are much easier to obtain when $W_n$ converges in mean to $W$ as the proof of the next result illustrates.

**Theorem 14.5** *Suppose $W_n$ converges in mean to $W$. Let $\{\mathfrak{L}[t] : t \geq 0\}$ be optional lines that are increasing with $t$ and satisfy $E_{\mathbb{B}}\left[W_{\mathfrak{L}[t]}\right] = h(0)$ for every $t$. Then $(W_{\mathfrak{L}[t]}, \mathcal{G}_{\mathfrak{L}[t]})$ is a positive martingale, $W_{\mathfrak{L}[t]}$ converges in mean. The martingale's limit is $W$ if, for each $n$, $W_{\mathfrak{L}[t]_n}$ tends to $W_n$, almost surely, as $t \to \infty$.*

*Proof.* Lemma 14.2 gives $E_{\mathbb{B}}\left[W_n | \mathcal{G}_{\mathfrak{L}[t]}\right] = W_{\mathfrak{L}[t]_n}$. Now, letting $n \to \infty$ shows that $E_{\mathbb{B}}[W | \mathcal{G}_{\mathfrak{L}[t]}] \geq W_{\mathfrak{L}[t]}$; both sides have expectation $h(0)$, forcing equality. Standard martingale theory now gives that $W_{\mathfrak{L}[t]}$ converges to $E_{\mathbb{B}}\left[W | \mathcal{G}_*\right]$. When $W_{\mathfrak{L}[t]_n}$ tends to $W_n$, letting $t \to \infty$ and then $n \to \infty$ in $E_{\mathbb{B}}\left[W_n | \mathcal{G}_{\mathfrak{L}[t]}\right] = W_{\mathfrak{L}[t]_n}$ gives $E_{\mathbb{B}}\left[W | \mathcal{G}_*\right] = W$. $\square$

# 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. (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14**, 25–37.

[7] Biggins, J.D. (1991). 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. **18**, 159–173.

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

[9] Biggins, J.D. (2003). Random walk conditioned to stay positive. *Jnl. Lond. Math. Soc.* **67** 259–272.

[10] Biggins, J.D., Cohn, H. and Nerman, O. (1999). Multitype branching in varying environment. *Stoc. Proc. Appl.* **83**, 357–400.

[11] Biggins, J.D. and Kyprianou, A.E. (1997). Seneta-Heyde norming in the branching random walk. *Ann. Probab.* **25**, 337–360.

[12] Biggins, J.D. and Kyprianou, A.E. (2001). The smoothing transform; the boundary case. Preprint 516/01, Department of probability and Statistics, University of Sheffield. Available at:
`http://www.shef.ac.uk/~st1jdb/tsttbc.html`

[13] 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.

[14] Chauvin, B. (1991). Product martingales and stopping lines for branching Brownian motion. *Ann. Probab.* **19**, 1195–1205.

[15] Doney, R.A. (1980). Moments of ladder heights in random walks. *Jnl. Appl. Probab.* **17**, 248–252.

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

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

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

[19] Goettge, R.T. (1975). Limit Theorems for the Supercritical Galton-Watson Process in Varying Environments, *Math. Biosci.* **28**, 171–190.

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

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

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

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

[24] Kingman, J.F.C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3**, 790–801.

[25] 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.

[26] 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.

[27] Kyprianou, A.E. (2000). Martingale Convergence and the Stopped Branching Random Walk. *Probab. Theory Rel. Fields* **116**, 405–419.

[28] Kyprianou, A.E. (2004). Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris' probabilistic analysis. *Ann. Inst. H. Poincaré* **40** 53–72.

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

[30] 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

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

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

[33] 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.

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

[35] 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.

[36] 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.

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

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

[39] 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