

Martingale convergence and the functional equation in the multi-type branching random walk

ANDREAS E. KYPRIANOU¹ and A. RAHIMZADEH SANI²

¹*Department of Mathematics, University of Utrecht, Budapestlaan 6, 3584 CD Utrecht, Netherlands. E-mail: kyprianou@math.uu.nl*

²*Department of Mathematics, Teacher Training University, Tehran, Iran. E-mail: rahimsan@saba.tmu.ac.ir*

A generalization of Biggins's martingale convergence theorem is proved for the multi-type branching random walk. The proof appeals to modern techniques involving the construction of size-biased measures on the space of marked trees generated by the branching process. As a simple consequence we obtain existence and uniqueness of solutions (within a specified class) to a system of functional equations.

Keywords: functional equation; multi-type branching random walk; size-biased measures

1. Introduction

A multi-type branching random walk with p types is defined as follows. An initial ancestor, \mathfrak{O} , of type $i \in \{1, \dots, p\}$ resides at the origin of the real line. This individual gives birth to a random number of offspring scattered on \mathbb{R} according to the point process $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ip})$, where Z_{ij} is the point process counting in Borel sets the number of individuals of type $j \in \{1, \dots, p\}$ that are born to the individual of type i . These offspring, the first generation, reproduce independently such that individuals of type j reproduce according to the point process \mathbf{Z}_j ($j = 1, \dots, p$) and so on.

In this text we shall use an Ulam–Harris labelling notation. By counting siblings from left to right we can identify each individual, $u = (k_1 \dots, k_n)$ with the understanding that u is the k_n th child born to \dots born to the k_1 th child born to the initial ancestor. With this formulation we write $|u|$ for the generation in which u lives, $\tau(u)$ for its type and $\zeta(u)$ for its position in \mathbb{R} (thus $\tau(\mathfrak{O}) = i$ and $\zeta(\mathfrak{O}) = 0$). An individual is identified as uv if it is a descendent of u and, on the tree growing from u , its line of decent looks like v .

Suppose now that μ_{ij} is the intensity measure of the point process Z_{ij} such that for any Borel measurable set A , $\mu_{ij}(A) = E(Z_{ij}(A))$. Define the matrix $M(\theta) = \{m_{ij}(\theta)\}$ satisfying

$$m_{ij}(\theta) = \int_{\mathbb{R}} e^{-\theta x} \mu_{ij}(dx),$$

where $\theta \in \mathbb{R}$. When the entries of $M(\theta)$ are finite and it is positive regular, the Perron–Frobenius theorem tells us that there exists a positive maximum eigenvalue $\rho(\theta)$ and corresponding positive right and left eigenvectors $\mathbf{v}(\theta) = (v_1(\theta), \dots, v_p(\theta))$ and $\mathbf{u}(\theta) = (u_1(\theta), \dots, u_p(\theta))$ respectively whose entries are all finite and strictly positive. The following assumptions will hold throughout this paper:

1. $\theta > 0$;
2. $\theta \in \text{int}\{\phi : M(\phi) < \infty\}$;
3. $M(\theta)$ is positive regular;
4. $P(\{u : |u| = n\}) \neq \emptyset$ for all $n > 0$.

Requiring that $\theta > 0$ is only a matter of convenience. For negative values of θ the arguments given in this paper are easily adjusted. The second condition, together with the third, ensures that $M(\theta)$, $\mathbf{v}(\theta)$, $\mathbf{u}(\theta)$ and $\rho(\theta)$ all have first derivatives that are finite. The fourth condition implies that the process is supercritical and survives for an infinite number of generations with positive probability. This will happen if $M(0)$ is positive definite with $\rho(0) > 1$; see Athreya and Ney (1972, Chapter 5) for further details.

Without loss of generality, we can assume that the left and right eigenvectors of $M(\theta)$ are normalized such that

$$\sum_{i=1}^p u_i(\theta) = \sum_{i=1}^p u_i(\theta)v_i(\theta) = 1.$$

Defining, for each $i \in 1, \dots, p$,

$$W_i^n(\theta) = \sum_{|u|=n} \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)^n},$$

it can be shown that $\{W_i^n(\theta)\}_{n \geq 0}$ is a mean-one martingale with respect to $\{\mathcal{F}_i^n\}_{n \geq 0}$, the σ -algebras generated by the first $n (\geq 0)$ levels of the multi-type branching random walk initiated by an individual of type i ; see, for example, Rahimzadeh Sani (1999) or Bramson *et al.* (1992). Being a positive martingale, it has an almost sure limit which we shall refer to as $W_i(\theta)$. When there is just one type, then we have a branching random walk and this martingale is the same as the one which is the subject of the martingale convergence theorem first proved by Biggins (1977). Biggins's theorem gives necessary and sufficient conditions for the aforementioned martingale to converge in mean. Further, it generalizes an older result of Kesten and Stigum (1966) which says that if $\{\alpha_n\}_{n \geq 0}$ is a finite-mean, supercritical Galton–Watson process then the positive martingale $E(\alpha_1)^{-n}\alpha_n$ converges in mean if and only if $E(\alpha_1 \log \alpha_1) < \infty$, and when this condition fails then it converges almost surely to zero; Athreya and Ney (1972) has a full account. Both the results of Kesten and Stigum and Biggins were recently reproved by Lyons *et al.* (1995) and Lyons (1997) respectively using a method involving a change of measure on the space of (marked) trees in which realizations of the branching process exist. The change of measure corresponds to size-biasing the reproduction distribution on a randomly chosen line of decent. This method improved on the existing proofs by shortening their length and using probabilistic considerations alone. The robustness of this method has also been demonstrated given the number of other Kesten–

Stigum type theorems for other types of branching processes that have since been reproved using changes of measures on trees; see Kurtz *et al.* (1997), Olofsson (1998), Athreya (2000) and Kyprianou (2001). The present paper now adds to this list with the following theorem.

Theorem 1. *The variable $W_i(\theta)$ is also an L^1 limit if and only if the following two conditions hold:*

$$\log \rho(\theta) - \theta \frac{\rho'(\theta)}{\rho(\theta)} > 0;$$

$$E[W_i^1(\theta) \log^+ W_i^1(\theta)] < \infty, \quad i = 1, \dots, p.$$

Moreover, if either of these two conditions fails, then $W_i(\theta)$ is a.s. zero.

This theorem has already been proved by Rahimzadeh Sani (1999) using methods that generalize the techniques that appeared in the original proof of Biggins’s martingale convergence theorem. The new proof we offer here is considerably quicker. Note also that when $\theta = 0$ this result is the known Kesten–Stigum theorem for multi-type Galton–Watson processes. For this reason assumption 1 has been imposed.

The connection between the limit of this martingale and a certain system of functional equations should not go unmentioned. We include this as a corollary to the theorem. Consider the class of vector functions of the form

$$\Phi = \{(\phi_1, \dots, \phi_p) : \mathcal{L}_{\mu_i} \ni \phi_i : \mathbb{R}^+ \rightarrow [0, 1] \text{ for all } i = 1, \dots, p \text{ and some } \mu_i > 0\},$$

where \mathcal{L}_{μ} is the class of Laplace transforms of positive variables with finite mean μ .

Corollary 2. *When the two conditions in Theorem 1 are satisfied together with the extra condition that $P(Z_{ij}(\mathbb{R}) = \infty) = 0$ for all $i, j \in \{1, \dots, p\}$, then*

$$\phi_i(x) = E(\exp\{-xW_i(\theta)\}), \quad i = 1, \dots, p,$$

is the unique solution in Φ (up to a common multiplicative constant in their arguments) to the system of functional equations

$$\phi_i(x) = E \left[\prod_{|u|=1} \phi_{\tau(u)} \left(x \frac{v_{\tau(u)}(\theta) e^{-\theta \zeta(u)}}{v_i(\theta) \rho(\theta)} \right) \right], \quad i = 1, \dots, p. \tag{1}$$

This system of functional equations can be thought of as a discrete-time analogue of the ordinary differential equations giving travelling wave solutions to a coupled system of K-P-P equations; see Champneys *et al.* (1995). The θ considered in the corollary correspond indirectly to the supercritical wave speeds in the K-P-P analogy. In principle, the method of size-biasing we use here is equally applicable to constructing alternative solutions the problems discussed there. The system (1) may also be considered as a multi-type version of a smoothing transform. Such smoothing transforms and their applications have been studied, for example, by Durrett and Liggett (1983), Liu (1997; 1998), Chauvin and Rouault (1997), Liu and Rouault (1997), Koukiou (1997) and Waymire and Williams (1996).

Remark 3. It follows from Corollary 2 that, for any solution to (1),

$$\frac{1 - \phi_i(x)}{x} \rightarrow \mu \quad \text{as } x \downarrow 0,$$

for $i = 1, \dots, p$, where μ is some constant.

Although this is a trivial observation, it is worth making in view of the comments about (1) in relation to travelling wave solutions to the K-P-P equation and smoothing transforms. In both the latter cases there exist analogous asymptotics to those above. See, for example, the discussion and calculations in Liu (1998), Champneys *et al.* (1995), Harris (1999) and Kyprianou (2001).

We conclude this section by giving a brief outline of the paper. In the next section we consider the multi-type branching random walk as a process having sample paths on a measurable space of marked trees with an associated probability measure. It is shown that there exists a new probability measure on this space of marked trees whose Radon–Nikodym derivative (restricted to \mathcal{F}_i^n) with respect to the original measure is precisely $W_i^n(\theta)$. Consequently, the problem of L^1 -convergence is transformed to studying the martingale under the new probability measure. As mentioned earlier, the change of measure corresponds to size-biasing the reproduction distribution along a randomly chosen line of descent called the spine. It turns out that under the new measure, the behaviour of the martingale is dominated by the asymptotic behaviour of the spine. In Section 3 we show that the position and type of individuals on the spine follows a Markov additive process and further discuss some of its basic asymptotic properties. The final section is devoted to the proof of the theorem and its corollary.

2. Measures on trees with spines

The set of possible realizations of the multi-type branching random walk with initial ancestor of type i generates a space of trees with nodes marked in $\mathbb{R} \times \{0, 1, \dots, p\}$. Call this space \mathcal{T}_i , $i = 1, \dots, p$, and note that if \mathcal{T}_i^n is the subspace of \mathcal{T}_i consisting of all marked trees truncated at the n th generation, then \mathcal{F}_i^n is the σ -algebra generated by \mathcal{T}_i^n for all $n \geq 0$. The probability measure η_i on $(\mathcal{T}_i, \sigma(\mathcal{T}_i))$ corresponding to the reproduction laws outlined at the beginning of the previous section satisfies the decomposition

$$d\eta_i^{n+1}(t) = d\eta_i^1(t) \prod_{|u|=1} d\eta_{\tau(u)}^n(t(u)),$$

where $t \in \mathcal{T}_i$, η_i^n is the restriction of η_i to \mathcal{T}_i^n and $\{t(u) : |u| = 1\}$ are the independent subtrees of t initiated by individuals in the first generation.

For each $t \in \mathcal{T}_i$, starting from the initial ancestor, we can distinguish (possibly finite) ancestral lines of decent $\xi = (\xi_0, \xi_1, \dots)$ (so $\mathcal{O} = \xi_0$) which we shall call *spines*. Let $\tilde{\mathcal{T}}_i$ be the space of trees with a distinguished spine ξ , and $\tilde{\mathcal{T}}_i^n$ the subspace of trees with spines truncated at generation n . Call $\tilde{\mathcal{F}}_i^n (\supset \mathcal{F}_i^n)$ the σ -algebra generated by $\tilde{\mathcal{T}}_i^n$. Consider the measure $\tilde{\eta}_i$ on $(\tilde{\mathcal{T}}_i, \sigma(\tilde{\mathcal{T}}_i))$ such that for $(t, \xi) \in \tilde{\mathcal{T}}_i^n$ (using obvious notation)

$$d\tilde{\eta}_i^n(t, \xi) = \prod_{k=0}^{n-1} d\eta_{\tau(\xi_k)}^1(t(\xi_k)) \prod_{\substack{|v|=1 \\ u=\xi_k v \neq \xi_{k+1}}} d\eta_{\tau(u)}^{n-k-1}(t(u)),$$

which decomposes the probability measure η_i^n such that

$$d\eta_i^n(t) = \sum_{|u|=n} I(\xi_n = u) d\tilde{\eta}_i^n(t, \xi). \tag{2}$$

For $n \geq 0$, let $W_{\tau(\xi_n)}^1(\theta, \xi_n)$ be the version of $W_j^1(\theta)$ on the tree growing from ξ_n when $\tau(\xi_n) = j$. Construct a new bivariate probability measure on $\tilde{\pi}_i$ on $(\tilde{\mathcal{F}}_i, \sigma(\tilde{\mathcal{F}}_i))$ whose restriction to $\tilde{\mathcal{F}}_i^n$ satisfies

$$\begin{aligned} d\tilde{\pi}_i^n(t, \xi) &= \frac{v_{\tau(\xi_n)}(\theta) e^{-\theta \zeta(\xi_n)}}{v_i(\theta) \rho(\theta)^n} d\tilde{\eta}_i^n(t, \xi) \\ &= \prod_{k=0}^{n-1} \{p(\xi_{k+1}) \times W_{\tau(\xi_k)}^1(\theta, \xi_k)\} d\tilde{\eta}_i^n(t, \xi), \end{aligned} \tag{3}$$

where $(t, \xi) \in \tilde{\mathcal{F}}_i^n$ and

$$p(\xi_{k+1}) = \frac{v_{\tau(\xi_{k+1})} e^{-\theta(\zeta(\xi_{k+1}) - \zeta(\xi_k))}}{\sum_{|v|=1} v_{\tau(\xi_k v)} e^{-\theta(\zeta(\xi_k v) - \zeta(\xi_k))}}, \quad 0 \leq k \leq n-1.$$

(The decomposition (2) can be used to show that $\tilde{\pi}_i$ is really a probability measure, and therefore $\tilde{\eta}_i$ is not.) Let π_i^n be its projection onto \mathcal{F}_i^n . In view of (2), this is a probability measure that satisfies

$$\frac{d\pi_i^n}{d\eta_i^n}(t) = W_i^n(\theta). \tag{4}$$

The construction (3) suggests that $\tilde{\pi}_i^n$ corresponds to a multi-type branching random walk that evolves generation by generation as in Section 1 except for the following modification: along the spine, given the node ξ_n in generation $n \geq 0$, the law of its reproduction with respect to the law of $\mathbf{Z}_{\tau(\xi_n)}$ has Radon–Nikodym derivative $W_{\tau(\xi_n)}^1(\theta, \xi_n)$. Further, the node ξ_{n+1} is chosen from the offspring of ξ_n with probability $p(\xi_{n+1})$. Note that this implies the probability of no offspring for individuals along the spine is zero and therefore, under the measure $\tilde{\pi}_i$, spines which are finite form a null set.

It is worth remarking at this point that this construction of measures on trees with spines, while consistent with Lyons (1997), differs slightly from that of Athreya (2000). In Athreya’s paper, given $\tilde{\mathcal{F}}_i^n$, the next spinal node ξ_{n+1} is chosen randomly from the entire population in the $(n+1)$ th generation. However, to achieve the same Radon–Nikodym derivative as (4), Athreya still uses the same change of measure for the law of reproduction along the spine.

The proof of Theorem 1 in Section 4 will follow by considering the following dichotomy relating the behaviour of $W_i^n(\theta)$ under the measures π_i and η_i . Let $\bar{W}_i(\theta) = \limsup_{n \uparrow \infty} W_i^n(\theta)$, which also equals $W_i(\theta)$ η_i -a.s. Then (Durrett 1991, p. 210)

$$\bar{W}_i(\theta) = \infty \text{ } \pi_i\text{-a.s.} \Leftrightarrow \bar{W}_i(\theta) = 0 \text{ } \eta_i\text{-a.s.} \tag{5}$$

$$\bar{W}_i(\theta) < \infty \text{ } \pi_i\text{-a.s.} \Leftrightarrow \int \bar{W}_i(\theta) d\eta_i = 1. \tag{6}$$

3. Process on the spine

In this section we justify the claim that the process on the spine

$$S_n(\theta) := \{\theta \zeta(\xi_n) + n \log m(\theta), \tau(\xi_n)\}, \quad n \geq 0,$$

is $\tilde{\pi}$ -Markov additive. This property of the spine will prove to be very important in proving Theorem 1 and Corollary 2. To begin with, we shall briefly recall the definition of a Markov additive process and demonstrate some of its properties.

Suppose we have a family of independent random variables $\{X_1, \dots, X_p\}$ and an ergodic Markov process $\Lambda = \{\Lambda_n\}_{n \geq 0}$ on the integers $\{1, \dots, p\}$. Define the process S as follows:

$$S_0 = 0,$$

$$S_n = \sum_{i=1}^n Y_i$$

where $Y_i \stackrel{\text{i.i.d.}}{\sim} X_j$ if $\Lambda_i = j$. The pair $\{(S_n, \Lambda_n)\}_{n \geq 0}$ is called a Markov additive process. The following result follows from an easy application of the classical properties of random walks and renewal processes.

Lemma 4. *Suppose that $E|X_i| < \infty$ and let $\mu_i = E(X_i)$ for all $i = 1, \dots, p$. Suppose that (Π_1, \dots, Π_p) is the stationary distribution of Λ and define $\chi = \sum_{i=1}^p \mu_i \Pi_i$. Then $\chi \geq 0$ implies that*

$$\limsup_{n \uparrow \infty} S_n = \infty,$$

and $\chi \leq 0$ implies that

$$\liminf_{n \uparrow \infty} S_n = -\infty.$$

Note that this is not the strongest statement we can make about the limiting behaviour of the spatial part of a Markov additive process, but it will suffice for our purposes. Rahimzadeh Sani (1999) gives further results.

Proof. Consider first the ratio

$$\frac{S_n}{n} = \sum_{j=1}^p \left\{ \frac{\sum_{i=1}^n Y_i I(\Lambda_i = j)}{\sum_{i=1}^n I(\Lambda_i = j)} \right\} \left\{ \frac{\sum_{i=1}^n I(\Lambda_i = j)}{n} \right\}.$$

The law of large numbers and ergodicity of Λ imply that

$$\frac{S_n}{n} \rightarrow \chi := \sum_{i=1}^p \mu_i \Pi_i \text{ a.s.}$$

as n tends to infinity.

Suppose now that $\Lambda_0 = j$ and $\{N_j^n\}_{n \geq 0}$ are the times that the Markov chain Λ is in state $j \in \{1, \dots, p\}$. Note that $S^j = \{S_{N_j^n}\}_{n \geq 0}$ is a random walk. Further, by the previous observation and the ergodicity of Λ , we have the following law of large numbers:

$$\frac{S_{N_j^n}}{n} = \frac{S_{N_j^n}}{N_j^n} \frac{N_j^n}{n} \rightarrow \frac{\chi}{\Pi_j} \text{ a.s.,}$$

for all $j = 1, \dots, p$. Hence the mean increment of each S^j is χ/Π_j . (Note that this can also be verified by direct computation using standard techniques.) Consequently, the random walks $\{S^j : j = 1, \dots, p\}$ are simultaneously transient or recurrent according to the value of χ . Since, for any $j = 1, \dots, p$, $\limsup_{n \uparrow \infty} S_n \geq \limsup_{n \uparrow \infty} S_{N_j^n}$ and $\liminf_{n \uparrow \infty} S_n \leq \liminf_{n \uparrow \infty} S_{N_j^n}$, the result follows. \square

Consider the increments $Y_n = \theta \zeta(\xi_n) - \theta \zeta(\xi_{n-1}) + \log \rho(\theta)$, which by construction are independent. Given $\tau(\xi_{n-1}) = j$, the mean increment from type j ,

$$\begin{aligned} \mu_j(\theta) &:= E_{\tilde{\pi}_j}(Y_1) \\ &= \theta E_{\eta_j} \left[\sum_{|u|=1} \zeta(u) \frac{v_{\tau(u)}(\theta) e^{-\theta \zeta(u)}}{\sum_{|v|=1} v_{\tau(v)}(\theta) e^{-\theta \zeta(v)}} W_j^1(\theta) \right] + \log \rho(\theta) \\ &= -\theta \sum_{k=1}^p \frac{m'_{jk}(\theta) v_k(\theta)}{\rho(\theta) v_j(\theta)} + \log \rho(\theta). \end{aligned}$$

A similar calculation shows that, under assumption 2, the absolute expectation of the increments is finite; see also Biggins (1977) for similar reasoning. Also, if we consider the process of types along the spine $\{\tau(\xi_n)\}_{n \geq 0}$, we see that it is a $\tilde{\pi}_i$ -Markov chain with transition probability $p_{jk}(\theta)$ which is equal to

$$\begin{aligned} P_{\tilde{\pi}_j}(\tau(\xi_1) = k) &= E_{\eta_j} \left[\sum_{|u|=1} I(\tau(u) = k) \frac{v_{\tau(u)}(\theta) e^{-\theta \zeta(u)}}{\sum_{|v|=1} v_{\tau(v)}(\theta) e^{-\theta \zeta(v)}} W_j^1(\theta) \right] \\ &= \frac{m_{jk}(\theta) v_k(\theta)}{\rho(\theta) v_j(\theta)}. \end{aligned}$$

Hence it is easy to check that the stationary distribution $(\Pi_1(\theta), \dots, \Pi_p(\theta))$ satisfies $\Pi_j(\theta) = v_j(\theta) u_j(\theta)$ for all $j = 1, \dots, p$. Consequently, we conclude that the process $\{\theta \zeta(\xi_n) + n \log \rho(\theta), \tau(\xi_n)\}_{n \geq 0}$ is Markov additive. Its drift, $\chi(\theta)$, can be written more neatly as $\log \rho(\theta) - \theta \rho'(\theta)/\rho(\theta)$. To see this, recall that

$$\begin{aligned} \chi(\theta) &= \sum_{j=1}^p \mu_j(\theta) \Pi_j(\theta) \\ &= \log \rho(\theta) - \theta \frac{\mathbf{u}(\theta)M'(\theta)\mathbf{v}(\theta)^T}{\rho(\theta)}. \end{aligned}$$

Differentiate both sides of the equality $\mathbf{u}(\theta)M(\theta)\mathbf{v}(\theta)^T = \rho(\theta)$ and the result follows.

4. Proofs

Proof of Theorem 1. Suppose that $E(W_i^1(\theta)\log^+ W_i^1(\theta)) = \infty$ for some $i \in \{1, \dots, p\}$, thus implying that, for all $c > 0$,

$$\sum_{k \geq 1} P_{\tilde{\pi}_i}(\log^+ W_i^1(\theta) > cn) = \infty. \tag{7}$$

Using previous notation, consider the \mathcal{S}_{n+1} -measurable events

$$\{\log^+ W_i^1(\theta, \xi_{N^n}) > cn\},$$

where $\mathcal{S}_{n+1} := \tilde{\mathcal{F}}_i^{N^{n+1}}$. In view of (7), an application of (the conditional version of) the Borel–Cantelli lemma to this adapted sequence of events implies that

$$P_{\tilde{\pi}_i}(\log^+ W_i^1(\theta, \xi_{N^n}) > cn \text{ infinitely often}) = 1$$

for all $c > 0$, and hence

$$\limsup_{n \uparrow \infty} \frac{\log^+ W_{\tau(\xi_n)}^1(\theta, \xi_n)}{n} \geq \text{const.} \times \limsup_{n \uparrow \infty} \frac{\log^+ W_i^1(\theta, \xi_{N^n})}{n} = \infty$$

$\tilde{\pi}_i$ -a.s. where, if $\tau(\xi_n) = j$, $W_{\tau(\xi_n)}^1(\theta, \xi_n)$ is the version of $W_j^1(\theta)$ on the tree growing from ξ_n .

Conversely, if we assume that $E(W_i^1(\theta)\log^+ W_i^1(\theta)) < \infty$ for all $i \in \{1, \dots, p\}$ then a similar argument, using again the Borel–Cantelli lemma, implies that

$$P_{\tilde{\pi}_i}(\log^+ W_i^1(\theta, \xi_{N^n}) > cn \text{ infinitely often}) = 0$$

for all $c > 0$ and $i = 1, \dots, p$, so that

$$\limsup_{n \uparrow \infty} \frac{\log^+ W_{\tau(\xi_n)}^1(\theta, \xi_n)}{n} \leq \text{const.} \times \sum_{i=1}^p \limsup_{n \uparrow \infty} \frac{\log^+ W_i^1(\theta, \xi_{N^n})}{n} = 0$$

$\tilde{\pi}_i$ -a.s.

Now that the $\tilde{\pi}_i$ -almost sure asymptotic behaviour of $S_n(\theta)$ and $W_{\tau(\xi_n)}^1(\theta, \xi_n)$ have been established with respect to the conditions of the theorem, the proof follows that of the case $p = 1$ given in Lyons (1997) (but for some minor alterations). We shall include it here for completeness.

By decomposing $W_i^{n+1}(\theta)$ into contributions from individuals in the n th generation it can be seen that

$$W_i^{n+1}(\theta) \geq \exp\{-\theta\zeta(\xi_n) + n \log \rho(\theta)\} \frac{v_{\downarrow}(\theta)}{v_{\uparrow}(\theta)} W_{\tau(\xi_n)}^1(\theta, \xi_n), \tag{8}$$

where

$$v_{\downarrow}(\theta) = \min\{v_i(\theta) : i = 1, \dots, p\} > 0,$$

$$v_{\uparrow}(\theta) = \max\{v_i(\theta) : i = 1, \dots, p\} > 0.$$

From the previous discussion, if $\chi(\theta) \leq 0$ then the diverging exponential term in (8) is sufficient to guarantee that

$$\limsup_{n \uparrow \infty} W_i^n(\theta) = \infty \text{ } \pi_i\text{-a.s.}$$

If, however, $\chi(\theta) > 0$ and $E(W_i^1(\theta) \log^+ W_i^1(\theta)) = \infty$ for some $i \in \{1, \dots, p\}$, then the lim sup of the right-hand side of (8) is dominated by the final term and we conclude again that $\bar{W}(\theta) = \infty$ π_i -a.s.

Now suppose that both $\chi(\theta) > 0$ and $E(W_i^1(\theta) \log^+ W_i^1(\theta)) < \infty$ for all $i \in \{1, \dots, p\}$. Let \mathcal{S} be the σ -field generated by the sequence $\{\mathbf{Z}_{\tau(\xi_n)}\}_{n \geq 0}$; then

$$E_{\tilde{\pi}_i}(W_i^n(\theta) | \mathcal{S}) = \sum_{k=1}^{n-1} \frac{v_{\tau(\xi_k)}(\theta) e^{-\theta\zeta(\xi_k)}}{v_i(\theta) \rho(\theta)^k} W_{\tau(\xi_k)}^1(\theta, \xi_k) - \sum_{k=0}^{n-1} \frac{v_{\tau(\xi_k)}(\theta) e^{-\theta\zeta(\xi_k)}}{v_i(\theta) \rho(\theta)^k}.$$

Referring to the previous discussion, the summands in both terms above decay at worst exponentially, and thus Fatou's lemma implies that

$$E_{\tilde{\pi}_i}(\liminf_{n \uparrow \infty} W_i^n(\theta) | \mathcal{S}) < \infty \text{ } \tilde{\pi}_i\text{-a.s.}$$

As $[W_i^n(\theta)]^{-1}$ is a positive π_i -martingale we thus have that $\lim_{n \uparrow \infty} W_i^n(\theta)$ and hence $\bar{W}(\theta)$ are finite π_i -a.s. □

Proof of Corollary 2. The proof we will give uses ideas from Doney (1972) that were also employed in Biggins (1977) for the one-type branching random walk.

Decomposing $W_i^n(\theta)$ into contributions from individuals in the n th generation and taking limits as n tends to infinity gives the distributional identity

$$W_i(\theta) = \sum_{|u|=1} \frac{v_{\tau(u)}(\theta) e^{-\theta\zeta(u)}}{v_i(\theta) \rho(\theta)} W_{\tau(u)}(\theta, u)$$

where, if $\tau(u) = j$, $W_{\tau(u)}(\theta, u)$ is the version of $W_j(\theta)$ on the tree rooted at u . Taking Laplace transforms of this identity (paying particular attention to the extra assumption in the statement of the corollary) gives a solution to the functional equation (1) in Φ .

Suppose now we take two solutions (ϕ_1, \dots, ϕ_p) and $(\varphi_1, \dots, \varphi_p)$ in Φ that satisfy the functional equation (1). Solutions to equation (1) in Φ can equivalently be seen as random variables $(\Delta_1, \dots, \Delta_p)$ satisfying the smoothing transform

$$\Delta_i \stackrel{d}{=} \sum_{|u|=1} \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)} \Delta_{\tau(u)}(u), \quad i = 1, \dots, p,$$

where $\stackrel{d}{=}$ is equality in distribution and, given \mathcal{F}_i^1 , $\Delta_{\tau(u)}(u)$ are independent copies of Δ_j when $\tau(u) = j$. Taking expectations in this smoothing transform, one obtains

$$\mu_i = \sum_{j=1}^p P_{ij}(\theta)\mu_j, \quad i = 1, \dots, p, \tag{9}$$

where $P_{ij}(\theta) = m_{ij}(\theta)v_j(\theta)/\rho(\theta)v_i(\theta)$ is the transition matrix of the individual type on the spine under $\tilde{\pi}$. Since $M(\theta)$ is positive regular, it can be easily checked that $P(\theta)$ is, too. Thus it follows from Kemeny and Snell (1976) that the solution to (9) is $\mu_i = \mu > 0$ for all $i = 1, \dots, p$.

Since we are to prove uniqueness up to a multiplicative constant in the argument, we may now take it for granted that, for $i = 1, \dots, p$, ϕ_i and φ_i are in \mathcal{L}_1 . For all $x > 0$, let $g_i(x) = x^{-1}|\phi_i(x) - \varphi_i(x)|$ and

$$g(x) = \max \left\{ \frac{|\phi_i(x) - \varphi_i(x)|}{x} : i = 1, \dots, p \right\},$$

so that g is bounded and positive with $g(0^+) = 0$. We have

$$\begin{aligned} g_i(x) &\leq \frac{1}{x} E_{\eta_i} \left(\sum_{|u|=1} \left| \phi_{\tau(u)} \left(x \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)} \right) - \varphi_{\tau(u)} \left(x \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)} \right) \right| \right) \\ &\leq E_{\eta_i} \left(\sum_{|u|=1} g_{\tau(u)} \left(x \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)^n} \right) \frac{v_{\tau(u)}(\theta)e^{-\theta\zeta(u)}}{v_i(\theta)\rho(\theta)^n} \right) \\ &= E_{\tilde{\pi}_i} \left(g_{\tau(\xi_1)} \left(x \frac{v_{\tau(\xi_1)}(\theta)e^{-\theta\zeta(\xi_1)}}{v_i(\theta)\rho(\theta)} \right) \right). \end{aligned}$$

Iterating yields

$$\begin{aligned} g_i(x) &\leq E_{\tilde{\pi}_i} \left(g_{\tau(\xi_n)} \left(x \frac{v_{\tau(\xi_n)}(\theta)e^{-\theta\zeta(\xi_n)}}{v_i(\theta)\rho(\theta)^n} \right) \right) \\ &\leq E_{\tilde{\pi}_i} \left(g \left(x \frac{v_{\tau(\xi_n)}(\theta)e^{-\theta\zeta(\xi_n)}}{v_i(\theta)\rho(\theta)^n} \right) \right) \end{aligned}$$

for all $n \geq 1$. As has already been demonstrated, under the conditions of the theorem,

$$\lim_{n \uparrow \infty} \frac{v_{\tau(\xi_n)}(\theta)}{v_i(\theta)} \exp\{-\theta\zeta(\xi_n) + n \log \rho(\theta)\} = 0 \quad \tilde{\pi}_i\text{-a.s.},$$

and hence, since g is bounded, $0 \leq g_i(x) \leq g(0^+) = 0$ for all $x > 0$. □

Acknowledgements

This work was initiated following a visit by A.E. Kyprianou to the Teacher Training University, Tehran. Both authors would like to express their thanks to the Teacher Training University for their support.

References

- Athreya, K.B. (2000) Change of measures for Markov chains and the LlogL theorem for branching processes. *Bernoulli*, **6**, 323–338.
- Athreya, K.B. and Ney, P. (1972) *Branching Processes*. Berlin: Springer-Verlag.
- Biggins, J.D. (1977) Martingale convergence in the branching random walk. *J. Appl. Probab.*, **14**, 25–37.
- Bramson, M., Ney, P. and Tao, J. (1992) The population composition of a multi-type branching random walk. *Ann. Appl. Probab.*, **2**, 575–596.
- Champneys, A., Harris, S.C., Toland, J.F., Warren, J. and Williams, D. (1995) Algebra, analysis and probability for a coupled system of reaction–diffusion equations. *Philos. Trans. Roy. Soc. Lond. Ser. A*, **350**, 69–112.
- Chauvin, B. and Rouault, A. (1997) Boltzmann–Gibbs weights in the branching random walk. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes*, IMA Vol. Math. Appl. **84**, pp. 41–50. New York: Springer-Verlag.
- Doney, R.A. (1972) A limit theorem for a class of supercritical branching processes. *J. Appl. Probab.*, **9**, 707–724.
- Durrett, R. (1991) *Probability: Theory and Examples*. Pacific Grove, CA: Wadsworth & Brooks/Cole.
- Durrett, R. and Liggett, M. (1983) Fixed points of the smoothing transform. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **64**, 275–301.
- Harris, S.C. (1999) Travelling-waves for the F-K-P-P equation via probabilistic arguments. *Proc. Roy. Soc. Edinburgh Sect. A*, **129**, 503–517.
- Kemeny, J.G. and Snell, J.L. (1976) *Finite Markov Chains*. Berlin: Springer-Verlag.
- Kesten, H. and Stigum, B.P. (1966) A limit theorem for multi-type Galton–Watson processes. *Ann. Math. Statist.*, **37**, 1211–1223.
- Koukiou, F. (1997) Directed polymers in random media and spin glass models on trees. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes*, IMA Vol. Math. Appl. **84**, pp. 171–179. New York: Springer-Verlag.
- Kurtz, T., Lyons, R., Pemantle, R., Peres, Y. (1997) A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes*, IMA Vol. Math. Appl. **84**, pp. 181–185. New York: Springer-Verlag.
- Kyprianou, A.E. (2001) Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris’ probabilistic analysis. Submitted.
- Lyons, R. (1997) A simple path to Biggins’ martingale convergence for branching random walk. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes*, IMA Vol. Math. Appl. **84**, pp. 217–221. New York: Springer-Verlag.
- Lyons, R., Pemantle, R. and Peres, Y. (1995) Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes. *Ann. Probab.*, **23**, 1125–1138.
- Liu, Q. (1997) Sur une équation fonctionnelle et ses applications: une extension du théorème de Kesten–Stigum concernant des processus de branchement. *Adv. Appl. Probab.*, **29**, 353–373.

- Liu, Q. (1998) Fixed points of a generalized smoothing transform and its applications to the branching random walk. *Adv. Appl. Probab.*, **30**, 85–112.
- Liu, Q. and Rouault, A. (1997) On two measures defined on the boundary of a branching tree. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes*, IMA Vol. Math. Appl. **84**, pp. 187–201. New York: Springer-Verlag.
- Olofsson, P. (1998) The $x \log x$ condition for general branching processes. *J. Appl. Probab.*, **35**, 537–554.
- Rahimzadeh Sani, A. (1999) Martingale convergence in the multitype branching random walk. In *4th Iranian International Statistics Conference*, Vol. 1. Tehran: Shahid Beheshti University Press.
- Waymire, E. and Williams, S. (1996) A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.*, **348**, 585–632.

Received June 2000 and revised May 2001