Branching Random Walk: Seneta-Heyde norming

J.D.Biggins*
A.E.Kyprianou†
The University of Sheffield, U.K.

Abstract: In the discrete-time branching random walk, the martingale formed by taking the Laplace transform of the nth generation point process is known, for suitable values of the argument, to converge in L_1 under an $X \log X$ condition, and to converge to zero when this moment condition fails. This paper examines the strategy used in Biggins and Kyprianou (1996) to prove that, when the $X \log X$ condition fails, there exists a Seneta-Heyde renormalisation of the martingale that converges in probability to a non-trivial random variable. To bring out how the method works it is first discussed in the context of the Galton-Watson process. The paper is concluded by extending the results to the case of the continuous-time Markov branching random walk.

Résumé: Dans une marche aléatoire de branchement à temps discret, on sait que pour des valeurs convenables de l'argument, la martingale obtenue en prenant la transformée de Laplace du processus ponctuel de la n-ième génération converge dans L_1 sous condition $X \log X$ et converge vers zéro quand cette condition de moment n'est pas vérifiée. Ce papier examine la stratégie développée par Biggins et Kyprianou (1996) pour démontrer que lorsque la condition $X \log X$ ne tient plus, il existe une renormalisation de Seneta-Heyde de la martingale qui converge en probabilité vers une variable aléatoire non triviale. Pour éclairer le fonctionnement de la méthode, celle-ci est d'abord discutée dans le cadre d'un processus de Galton-Watson. Le papier se conclut en étendant les résultats au cas d'une marche aléatoire de branchement markovienne à temps continu.

1 Introduction

This paper outlines the strategy used by Biggins and Kyprianou (1996) to prove that martingales arising in the branching random walk have a Seneta-Heyde norming when the appropriate version of an $X \log X$ condition fails. To bring out how the method works it will first be discussed in the context of the Galton-Watson process (though for that case there are several simpler ways to obtain the desired result). Next the case of the discrete-time branching random walk is examined. The paper finishes by extending the

*Some of this work was carried out whilst visiting the Institute of Mathematics and its Applications, University of Minnesota and the Mittag-Leffler Institute, Stockholm. I am grateful to both for their support.

†Supported by an EPSRC studentship

Keywords and phrases. Martingales, functional equations, Seneta-Heyde norming, branching random walk

AMS 1991 subject classification. Primary 60J80.

Probability and Statistics Section, School of Mathematics and Statistics, The Hicks Building, The University of Sheffield, Sheffield, S3 7RH, U.K.

results to the continous-time Markov branching random walk. The reader should note that the emphasis in the discussion of the discrete time case here is on methodology; rigorous proofs for that case are contained in Biggins and Kyprianou (1996).

As usual, individuals are labelled by their line of descent, so if $u = i_1 \ldots i_n$ then u is the i_nth child of the i_{n-1}th child of ... the i_1th child of the initial ancestor, and $|u|$ is the generation in which u is born. Sums and products involving u will always be over only those that are ever born, but this will not be explicitly stated each time.

Let $Z^{(n)}$ be the number of people in the nth generation, defined recursively by

$$Z^{(n+1)} = \sum_{|u|=n} Z_u,$$

where, given $\mathcal{F}^n$, which is the σ-field describing the ancestry of every individual in the nth generation, $\{Z_u : |u| = n\}$ are independent copies of Z, the family size distribution. In general, a subscript u will be used to indicate quantities associated with u. Let the mean family size, EZ, be $m \in (1, \infty)$. It is well known that $W^{(n)}$, given by $W^{(n)} = m^{-n} Z^{(n)}$, is a (non-negative) martingale, The Kesten-Stigum Theorem, see Asmussen and Hering (1983), shows that this martingale converges in L_1, so that its limit W has mean 1, exactly when $EZ \log Z < \infty$, and the limit is degenerate otherwise. When W is non-degenerate the set $\{W > 0\}$ coincides, almost surely, with the set on which the process survives, that is with $\{Z^{(n)} \to \infty\}$. In general, even when the limit is degenerate, it is possible to find constants $\{l_n\}$ such that

$$l_n W^{(n)} \to \Delta < \infty \quad \text{almost surely,}$$

where $\{\Delta > 0\}$ coincides with the survival set; a result that goes back to Seneta (1968) and Heyde (1970). (Of course when $EZ \log Z < \infty$ the l_n can all be taken to be 1.) There are several ways to prove this result; the neatest is probably that due to Grey (1980), in which an independent copy of the branching process supplies the normalization. However, it does not seem possible to use Grey's method for more complicated branching models.

In the next section a variation of the method employed by Cohn (1982a, 1982b, 1983, 1985) for obtaining a Seneta-Heyde renormalization will be explored in the context of the Galton-Watson process, and the significant steps summarised. The discussion in that section is arranged to lead into, and to facilitate, the corresponding one for the branching random walk, so this presentation of a well-known result is worthwhile. The third section starts with a description of the branching random walk and the martingales which are natural counterparts of $W^{(n)}$. The three key steps, identified during the Galton-Watson discussion, needed to establish a Seneta-Heyde renormalization of these martingales are then examined in turn. The methods used here are strong enough to obtain a Seneta-Heyde renormalization for a general (Crump-Mode-Jagers) branching process. This result, which illustrates that the techniques employed improve on the known results, is discussed briefly in the fourth section.

The main stimulus for the study described in the third and fourth sections was the paper by Neveu (1988) on branching Brownian motion. The fifth section describes briefly how that paper suggested the main result given here. The final section shows how this main result, which is for a discrete-time model, is easily extended to cover the continuous-time Markov case, which includes branching Brownian motion.

2 The Galton-Watson case

The following well-known result is the subject of this section. It only asserts convergence in probability since it is in this form that the result generalizes fully, but, as will become clear, the method actually yields almost sure convergence in the Galton-Watson framework.

Theorem 2.1 *There exist positive constants l_n such that*

$$l_n W^{(n)} \to \Delta \text{ in probability,}$$

as $n \to \infty$, where Δ is finite and strictly positive on the survival set, almost surely.

Note first that, as an elementary consequence of the weak law of large numbers,

$$\frac{W^{(n+1)}}{W^{(n)}} = \frac{Z^{(n+1)}}{mZ^{(n)}} = \frac{1}{Z^{(n)}} \sum_{|u|=n} \frac{Z_u}{m} \to 1 \tag{2.1}$$

in probability on the survival set, as $n \to \infty$.

Let the Laplace transform of $W^{(n)}$ be $\Omega_n(x)$. Take l_n to be such that $\Omega_n(l_n) = \kappa$, where κ is fixed to be greater than the extinction probability but less than one. Choose a subsequence, $\{n(i) : i = 1, 2, \ldots\}$, tending to infinity, such that, along this subsequence, $l_n W^{(n)}$ converges in distribution, with the Laplace transform of the limit being Ψ, so that

$$E \exp(-x l_{n(i)} W^{(n(i))}) \to \Psi(x) \quad \text{as } i \to \infty.$$

The limit might be an improper variable, but, as a result of the definition of l_n, it's mass cannot be concentrated solely at zero or infinity. Thus $\{l_n\}$ is a candidate sequence for the Seneta-Heyde norming. It is worth stressing that, in principle, the limit depends on the subsequence selected.

Let Δ be a variable with transform Ψ. Then, as a consequence of (2.1), Δ must satisfy a distributional equation, and Ψ must satisfy a corresponding functional equation. To see this, note that the branching property implies that each person in the first generation can be viewed as the initial ancestor of an independent copy of the process, so that

$$W^{(n)} = \sum_{|u|=1} m^{-1} W_u^{(n-1)},$$

where, given $\mathcal{F}^1$, $\{W_u^{(n-1)} : |u| = 1\}$ are independent copies of $W^{(n-1)}$. Thus

$$l_n W^{(n)} = \sum_{|u|=1} m^{-1} l_n W_u^{(n)} \frac{W_u^{(n-1)}}{W_u^{(n)}}. \tag{2.2}$$

By taking limits along $\{n(i)\}$ and using (2.1), it follows that

$$\Delta \stackrel{\mathcal{D}}{=} \sum_{|u|=1} m^{-1} \Delta_u,$$

where $\stackrel{\mathcal{D}}{=}$ means equality in distribution and, of course, given $\mathcal{F}^1$, $\{\Delta_u : |u| = 1\}$ are independent copies of Δ. Hence, in terms of Ψ,

$$\Psi(x) = E \prod_{|u|=1} \Psi(xm^{-1}).$$

If f is the generating function of the family-size distribution then this equation can be written in the more familiar form

$$\Psi(x) = f(\Psi(xm^{-1})),$$

so that Ψ is a solution to the Poincaré functional equation associated with f. (This argument shows that the existence of a solution to the Poincaré functional equation for a generating function f is a simple consequence of (2.1).)

Letting $x \downarrow 0$ in the Poincaré functional equation shows that $\Psi(0^+)$ satisfies $f(s) = s$, and so must equal either the extinction probability or 1. By arrangement $\Psi(1)$ exceeds the extinction probability, and Ψ is necessarily decreasing so $\Psi(0^+) = 1$; thus Δ is a proper random variable. In a similar way, but letting $x \uparrow \infty$, it follows that Δ has an atom at zero with probability equal to the extinction probability. So, although the variable Δ may depend on the subsequence selected, it is always proper and, once convergence in probability is established, it must be non-zero on the survival set.

The split in (2.2) does not have to be made in the first generation. Splitting on generation k gives (for $n > k$)

$$l_n W^{(n)} = \sum_{|u|=k} m^{-k} l_n W_u^{(n)} \frac{W_u^{(n-k)}}{W_u^{(n)}},$$

so that, taking Laplace transforms conditional on $\mathcal{F}^k$, and using (2.1) again,

$$E[\exp(-x l_{n(i)} W^{(n(i))}) | \mathcal{F}^k] \to \prod_{|u|=k} \Psi(xm^{-k}) \quad \text{as } i \to \infty, \tag{2.3}$$

for all positive x, almost surely. The existence of these limiting conditional Laplace transforms and their properties allow the results described next to be brought into play.

Suppose that $\{Y_n\}$ is a sequence of non-negative random variables adapted to the increasing σ-fields $\{\mathcal{G}^n\}$ and that along a fixed subsequence $\{n(i) : i = 1, 2, \ldots\}$, for each k and $x > 0$, the conditional Laplace transform $E[\exp(-xY_{n(i)}) | \mathcal{G}^k]$ converges as $i \to \infty$. Denote the limit by $\psi_k(x)$. The first lemma is a simple exercise in conditional expectations.

Lemma 2.2 *For each $x > 0$, $\{\psi_k(x)\}$ forms a bounded non-negative martingale with respect to $\{\mathcal{G}^k\}$.*

The martingale $\{\psi_k(x)\}$ converges almost surely and in mean to a limit which will be denoted by $\psi(x)$. It turns out that when this limit has a suitable simple form the convergence in distribution of $\{Y_{n(i)} : i\}$ can be strengthened to convergence in probability. This is the content of the next lemma.

Lemma 2.3 *If, for $x > 0$, $\psi(x) = e^{-xX}$ for a finite random variable X (that does not depend on x) then $Y_{n(i)} \to X$ in probability, as $i \to \infty$.*

To explain the idea of the proof let $X_k = -\log \psi_k(1)$, so that the convergence of the martingale $\{\psi_k(1)\}$ implies that $X_k \to X$ almost surely. The convergence in probability is proved by showing that, for a suitable sequence $\{k(i)\}$, $Y_{n(i)} - X_{k(i)}$ converges in distribution to zero. To see why this should work, compute the appropriate Laplace transform:

$$\begin{aligned}
E[\exp(-x(Y_{n(i)} - X_{k(i)}))] &= E[E[\exp(-x(Y_{n(i)} - X_{k(i)}))|\mathcal{G}^{k(i)}]] \\
&= E[\exp(xX_{k(i)})E[\exp(-xY_{n(i)})|\mathcal{G}^{k(i)}]] \\
&\approx E[\exp(xX_{k(i)})\psi_{k(i)}(x)] \qquad \text{(large } n(i)) \\
&\approx E[\exp(xX)\exp(-xX)] = 1 \quad \text{(large } k(i)),
\end{aligned}$$

where it is the last step that uses the special form of the limit of the martingales $\psi_k(x)$. Convergence in distribution to a constant implies convergence in probability, so a rigorous version of this argument will provide a proof of the Lemma.

These lemmas can be compared with Theorem 3.1 in Cohn (1985), which plays a similar role in that study. That theorem is phrased in terms of random variables rather than Laplace transforms; the formulation adopted here is better suited to the present approach.

Applying Lemma 2.2 to (2.3) yields that, for each $x > 0$, $\prod_{|u|=k} \Psi(xm^{-k})$ is a martingale. Denote the corresponding limits by $M(x)$. At this point in the Galton-Watson case it is quite simple to finish off the argument for Seneta-Heyde norming with almost sure convergence, because this martingale is a simple function of the variables for which a renormalization is sought. Thus

$$\prod_{|u|=k} \Psi(xm^{-k}) = \Psi(xm^{-k})^{Z^{(k)}} = \exp(-Z^{(k)}(-\log \Psi(xm^{-k}))) \to M(x)$$

as $k \to \infty$, so that, for any $x > 0$, $\{-m^k \log \Psi(xm^{-k})\}$ will serve as $\{l_k\}$ and then

$$l_k W^{(k)} \to -\log M(x) \quad \text{almost surely.}$$

Of course different normalizations, arising from different values of x, can only give rise to scalar multiples of a single limit.

Ignoring this simplification, hoping instead to use Lemma 2.3, focuses attention on how the limits $\log M(x)$ depend on x. Note that

$$\begin{aligned}
-\log M(x) = -\log\left(\lim_k \prod_{|u|=k} \Psi(xm^{-k})\right) &= \lim_k \sum_{|u|=k} -\log \Psi(xm^{-k}) \\
&\approx \lim_k \sum_{|u|=k} (1 - \Psi(xm^{-k}))
\end{aligned}$$

so, when $(1 - \Psi(x))/x$ is slowly varying at zero,

$$-\log M(x) \approx \lim_k \sum_{|u|=k} x(1 - \Psi(m^{-k})) \approx -x \log M(1).$$

This indicates that establishing that $(1 - \Psi(x))/x$ is slowly varying is exactly the issue in showing that the martingale limits $M(x)$ have the right form for Lemma 2.3 to apply. (For the Poincaré functional equation it is a simple exercise in analysis to show that, for a solution that is a Laplace transform, the required slow variation obtains.) In this way convergence in probability along the selected subsequence is established. The proof is now completed by using the fact that the solution to the Poincaré functional equation is unique, up to a scale factor, so that all subsequences have the same limit. Uniqueness, based on analytical results about the functional equation, is noted in, for example, Lemma 4.1 and Theorem 4.2 of Seneta (1969).

It is worth identifying explicitly the main steps involved in finding the renormalization of $W^{(n)}$.

Step 1: A law of large numbers. Show that $W^{(n+1)}/W^{(n)}$ converges to 1 in probability. Select candidate norming constants $\{l_n\}$ that give convergence in distribution of $\{l_n W^{(n)}\}$ along a subsequence. Now use the branching property to show two things: the limit of the renormalized subsequence satisfies a distributional equation, so its transform, Ψ, satisfies the corresponding functional equation; and the limit of the conditional Laplace transforms of the renormalized sequence exists, and so provides martingales with associated non-trivial limits by Lemma 2.2.

Step 2: Slow variation. Show, by studying the functional equation, that $(1-\Psi(x))/x$ is slowly varying. Use this property to show that the martingale limits arising from Lemma 2.2 have the properties necessary for Lemma 2.3 to apply, thereby yielding convergence in probability of $\{l_n W^{(n)}\}$ along a subsequence.

Step 3: Uniqueness. Establish the uniqueness of the solution to the functional equation and use it to show that all subsequences have the same limit, so that convergence holds along the full sequence.

3 The branching random walk

The notation used for the branching random walk will subsume that already in use for the Galton-Watson process. An initial ancestor is at the origin of the real line, and the positions of her children are given by a point process Z. Each of these children has children in the same way, in that the positions of each family are, relative to the parent, given by an independent copy of Z, and so on. Let $Z^{(n)}$ be the point process formed by the nth generation, with points $\{z_u : |u| = n\}$; then, by definition, for any set A

$$Z^{(n+1)}(A) = \sum_{|u|=n} Z_u\left(A - z_u\right), \tag{3.1}$$

where, given $\mathcal{F}^n$, $\{Z_u : |u| = n\}$ are independent copies of Z.

Suppose Z has intensity measure μ with Laplace transform m, so that

$$m\left(\phi\right) = E\left[\sum_{|u|=1} \exp(-\phi z_u)\right].$$

Throughout it will be assumed that: $m(\phi)$ is finite on some interval; θ is a value within this interval; the process is supercritical, so that $m(0) > 1$; and family sizes are finite. For convenience, it will also be assumed that $\theta \geqslant 0$; this is no loss of generality as a reflection of the positions about the origin shows.

It is well known, and easily shown, that

$$W^{(n)}(\theta) = m\left(\theta\right)^{-n} \int e^{-\theta x} Z^{(n)}(dx) = \sum_{|u|=n} \frac{\exp(-\theta z_u)}{m\left(\theta\right)^{n}}$$

is a martingale with respect to the σ-fields $\{\mathcal{F}^n\}$. This martingale is positive and so has an almost sure limit $W(\theta)$ which, by Fatou's lemma, satisfies $E[W(\theta)] \leqslant 1$. When $\theta = 0$ this martingale is the Galton-Watson one discussed in the previous section.

The convergence in mean of the martingale $W^{(n)}(\theta)$ or similar martingales arising from slightly different processes is considered in, for example, Kingman (1975), Kahane and Peyrière (1976), Biggins (1977a), Neveu (1988), Lyons (1995). In particular, the next result, which is an analogue of the Kesten-Stigum Theorem, is a consequence of those in Biggins (1977a).

Theorem 3.1 *The martingale $W^{(n)}(\theta)$ converges in L_1, so that $EW(\theta) = 1$, if and only if*

$$\log m(\theta) - \theta m'(\theta)/m(\theta) > 0 \tag{3.2}$$

and

$$E[W^{(1)}(\theta) \log^+ W^{(1)}(\theta)] < \infty, \tag{3.3}$$

and $EW(\theta) = 0$ otherwise.

It is worth noting that the set of θ values satisfying (3.2) intersect with the interior of the domain of finiteness of m to form an open interval.

In the light of this theorem, it is natural to seek a Seneta-Heyde norming for the martingale $W^{(n)}(\theta)$, that is, to prove the following theorem.

Theorem 3.2 *When (3.2) holds there exists a sequence of constants $\{l_n\}$ such that*

$$l_n W^{(n)}(\theta) \to \Delta \quad \textit{in probability,}$$

as $n \to \infty$, where Δ is a finite random variable which is strictly positive when the process survives. (In general, both $\{l_n\}$ and Δ depend on θ.)

Since θ is fixed it will be omitted in the notation whenever possible. To simplify the notation further, let

$$y_u = \frac{\exp(-\theta z_u)}{m(\theta)^{|u|}}.$$

Thus, with these conventions, $W^{(n)} = \sum_{|u|=n} y_u$. The issues that arise in following through the method outlined in the previous section will now be examined.

Step 1: A law of large numbers

By looking at the branching processes stemming from each kth generation person it is easily seen that, for $n > k$,

$$W^{(n)} = \sum_{|u|=k} y_u W_u^{(n-k)}, \tag{3.4}$$

where, given $\mathcal{F}^k$, $\{W_u^{(n-k)} : |u| = k\}$ are independent copies of $W^{(n-k)}$. Thus

$$\frac{W^{(n+1)}}{W^{(n)}} = \sum_{|u|=n} \frac{y_u}{W^{(n)}} W_u^{(1)}$$

and so is, given $\mathcal{F}^n$, a weighted sum of independent identically distributed variables. Kurtz (1972) provides tools for studying the convergence of such weighted sums; using these leads to the following analogue of (2.1).

Theorem 3.3 *When (3.2) holds*

$$\frac{W^{(n+1)}}{W^{(n)}} \to 1, \quad \textit{in probability,} \tag{3.5}$$

as $n \to \infty$, on the survival set of the process.

The norming constants are defined through the Laplace transform of $W^{(n)}$, just as in the Galton-Watson case, and a suitable subsequence of $l_n W^{(n)}$ converges in distribution. The limit will again be denoted by Δ, with transform Ψ. Then, using (3.4) and Theorem 3.3, Δ satisfies the distributional equation

$$\Delta \stackrel{\mathcal{D}}{=} \sum_{|u|=1} y_u \Delta_u,$$

where, given $\{y_u : |u| = 1\}$, Δ_u are independent copies of Δ. In terms of transforms this becomes

$$\Psi(x) = E\left[\prod_{|u|=1} \Psi\left(x y_u\right)\right]. \tag{3.6}$$

This functional equation, or variants of it, has been much studied, see, for example, Kahane and Peyrière (1976), Biggins (1977a), Durrett and Liggett (1983), Pakes (1992), Liu (1995). The argument that Δ is proper and takes the value zero with the extinction probability is just as in the Galton-Watson case. Furthermore, using (3.4) and Theorem 3.3 again, as n goes to infinity along the selected subsequence

$$E[\exp(-x l_n W^{(n)}) | \mathcal{F}^k] \to \prod_{|u|=k} \Psi(x y_u) =: M^{(k)}(x), \tag{3.7}$$

for all positive x, almost surely. By Lemma 2.2, the right hand side here is a bounded martingale; its limit is denoted by $M(x)$.

The martingale $W^{(n)}$ is obtained by adding contributions from each nth generation person, and so may reasonably be called an additive martingale. In contrast the martingales $M^{(n)}$ result from multiplying terms; it is natural therefore to call these multiplicative martingales. This is the terminology used by Neveu (1988) in his study of similar martingales arising in the context of branching Brownian motion.

Step 2: Slow variation

Theorem 3.4 *When (3.2) holds any solution Ψ to the functional equation (3.6) that is the Laplace transform of a non-trivial random variable is such that $L(x) := x^{-1}(1-\Psi(x))$ is slowly varying as $x \downarrow 0$.*

Under the additional condition that $m(0) < \infty$, this theorem is a consequence of Theorem 2 of Liu (1995). In the context of the Bellman-Harris process, which gives rise to a particular case of equation (3.6), Theorem 3.4 follows from Schuh (1982), as is pointed out in Cohn (1983). The general, or Crump-Mode-Jagers, branching process gives rise to a functional equation just like (3.6), save only that, for all u, $y_u < 1$. Slow variation for this case is stated in Theorem 6.2 of Cohn (1985).

Provided $\sup_{|u|=k} y_u \downarrow 0$ (which does hold), Theorem 3.4 allows us to make the second of the following approximations as k gets large,

$$\begin{aligned} -\log M^{(k)}(x) &\approx \sum_{|u|=k} (1 - \Psi(xy_u)) \\ &\approx \sum_{|u|=k} x(1 - \Psi(y_u)) \approx -x \log M^{(k)}(1). \end{aligned}$$

Thus, in the limit as $k \to \infty$,

$$-\log M(x) = -x \log M(1), \tag{3.8}$$

and Lemma 2.3 applies to give that, along the selected subsequence, $\{l_n W^{(n)}\}$ converges in probability to $-\log M(1)$. This implies, in particular, that Δ is $-\log M(1)$.

Note too that

$$-\log M^{(n)}(1) \approx \sum_{|u|=n} (1 - \Psi(y_u)) = \sum_{|u|=n} L(y_u) y_u$$

so that when $W^{(n)}$ converges in mean, which implies that $L(0^+) = 1$,

$$-\log M^{(n)}(1) \approx \sum_{|u|=n} y_u = W^{(n)};$$

thus the logarithm of the multiplicative martingale and the additive martingale are then asymptotically equivalent.

Step 3: Uniqueness

Theorem 3.5 *When (3.2) holds, the non-trivial solution to the functional equation (3.6) is unique, within the class of Laplace transforms of non-negative variables (up to a multiplicative constant in the argument).*

It is this theorem that provided the greatest difficulties in establishing the main result. It is easy to see that any solution to the functional equation has a multiplicative martingale of the form introduced earlier (the right side of (3.7)) associated with it. Furthermore, as the martingale converges in mean, $EM(x) = \Psi(x)$, which combines with (3.8) to show that Ψ is the Laplace transform of $-\log M(1)$.

One tempting way forward is to try to use the slow variation of L and the approximation $-\log M^{(n)}(1) \approx \sum_{|u|=n} L(y_u) y_u$, mentioned at the end of Step 2, to show that

$$-\log M^{(n)}(1) \approx \sum_{|u|=n} L(y_u) y_u \approx L(a_n) \sum_{|u|=n} y_u = L(a_n) W^{(n)}$$

for some deterministic sequence a_n. Then the almost sure convergence of the multiplicative martingale would imply the almost sure convergence of $L(a_n)W^{(n)}$, which is a stronger result than that claimed in Theorem 3.2. Furthermore, much as in the uniqueness proof for the Poincaré equation given in Theorem III.5.2 in Asmussen and Hering (1983), two supposedly different solutions to the functional equation, Ψ_1 and Ψ_2, with associated multiplicative martingale limits M_1 and M_2, and sequences $a_n^{(1)}$ and $a_n^{(2)}$ respectively, then give

$$\frac{-\log M_1(1)}{-\log M_2(1)} = \lim_n \frac{L_1(a_n^{(1)})W^{(n)}}{L_2(a_n^{(2)})W^{(n)}} = \lim_n \frac{L_1(a_n^{(1)})}{L_2(a_n^{(2)})}.$$

The right hand limit thus exists and must be equal to some constant in $(0,\infty)$ because $-\log M_1(1)$ and $-\log M_2(1)$ are strictly positive on the survival set. Hence M_1 is a scalar multiple of M_2, so that Ψ_1 and Ψ_2 differ only by a scale factor, giving uniqueness.

Unfortunately the vital step in this line of reasoning, that the terms $\{L(y_u) : |u| = n\}$ can all be well approximated by $L(a_n)$ for some a_n because L is slowly varying, fails because the values of $\{y_u : |u| = n\}$ are too disperse. (It is easy to use the results in Biggins (1977b) to see that the ratio between the smallest and largest term in $\{y_u : |u| = n\}$ grows geometrically in n.) The way round this problem is to construct new martingales that are better behaved in this respect.

Let $\mathcal{I}(s)$ be the set of individuals who are the first in their line of descent to have y_u less than e^{-s}, so

$$\mathcal{I}(s) = \{u : y_u < e^{-s}, \text{ but } y_v \geqslant e^{-s} \text{ for } v < u\},$$

where $v < u$ if v is a strict ancestor of u. Note that

$$\mathcal{I}(s) = \{u : \theta z_u + |u| \log m(\theta) > s, \text{ but } \theta z_v + |v| \log m(\theta) \leqslant s \text{ for } v < u\},$$

so that this set of individuals are the first in their lines of descent to cross the space-time line $\theta p + t \log m(\theta) = s$, where (p,t) are (space,time) coordinates.

The sets $\{\mathcal{I}(s) : s\}$ form a collection that is totally ordered as s varies, in the sense that all members of any of these sets have ancestors in any earlier one. The collection $\{u : |u| = n\}$ is similarly ordered as n varies. By developing a suitable general theory, most of which already exists in the work of Chauvin (1988, 1991) and Jagers (1989), it can be shown that

$$M^{(\mathcal{I}(s))}(x) := \prod_{u \in \mathcal{I}(s)} \Psi(x y_u)$$

is a martingale with the same limit as the martingale $M^{(n)}(x)$; thus

$$M^{(\mathcal{I}(s))}(x) \to M(x) \text{ almost surely, as } s \to \infty.$$

The martingales $M^{(\mathcal{I}(s))}$ and $M^{(n)}$ have the same form, differing only in the individuals that the products are taken over. Taking the product over $\mathcal{I}(s)$ offers the big advantage that it ensures that, by arrangement, the terms in $\{y_u : u \in \mathcal{I}(s)\}$ are all quite near to e^{-s}. Thus the argument mentioned previously, where we replace $L(y_u)$ by something deterministic, has a much better chance of working when applied to this martingale.

To see what more is required suppose first that when a line of descent crosses a space-time line the exceedance is bounded. That is, in terms of the basic point process, the set $\{z_u : |u| = 1\}$ is bounded above, so that, for some c, $\sup\{\theta z_u + \log m(\theta) : |u| = 1\} < c$. Expressing this in terms of y_u for $u \in \mathcal{I}(s)$, there is a c such that $e^{-(s+c)} < y_u < e^{-s}$. Thus, as L is monotone,

$$L(e^{-s}) \sum_{u \in \mathcal{I}(s)} y_u \leqslant \sum_{u \in \mathcal{I}(s)} L(y_u) y_u \leqslant L(e^{-(s+c)}) \sum_{u \in \mathcal{I}(s)} y_u,$$

and the slow variation of L, and a little analysis, yields the almost sure asymptotic equivalence

$$L(e^{-s})\sum_{u\in\mathcal{I}(s)} y_u \sim -\sum_{u\in\mathcal{I}(s)} \log\Psi\left(y_u\right) \quad \text{as } s\to\infty; \tag{3.9}$$

the second expression is the logarithm of the multiplicative martingale $M^{(\mathcal{I}(s))}(1)$, which has, as has already been noted, the non-trivial limit $-\log M(1)$. The argument for uniqueness suggested earlier works without further problems in this case.

To cover the general case an argument is needed to show that the possibility of arbitrarily large exceedance over the space-time line does not disturb the limiting behaviour. The approach used is based on results obtained by Nerman (1981). Suppose we define $\mathcal{I}(s,c)$ to be those individuals in $\mathcal{I}(s)$ but with an exceedance over the line greater than c. Then the argument leading to the equivalence in (3.9) applies if sums are taken over $u\in\mathcal{I}(s)\setminus\mathcal{I}(s,c)$. The issue then becomes to show that the part omitted can be made as small as desired or, more precisely, after some straightforward manipulations, to show that

$$\lim_{c\to\infty}\lim_{t\to\infty}\frac{\sum_{\mathcal{I}(t,c)} e^t y_u L(y_u)/L(e^{-t})}{\sum_{\mathcal{I}(t)\setminus\mathcal{I}(t,c)} e^t y_u} = 0. \tag{3.10}$$

As slowly varying functions grow more slowly than any power, it will, in place of (3.10), be enough to show that, for sufficiently small $\epsilon>0$,

$$\lim_{c\to\infty}\lim_{t\to\infty}\frac{\sum_{\mathcal{I}(t,c)}(e^t y_u)^{1-\epsilon}}{\sum_{\mathcal{I}(t)\setminus\mathcal{I}(t,c)} e^t y_u} = 0. \tag{3.11}$$

The proof now proceeds by identifying an embedded general branching process, with birth times $\{-\log y_u\}$, thereby allowing Theorem 6.3 of Nerman (1981), on the convergence of the ratios of a general branching process counted by two different characteristics, to be invoked to prove (3.11) and hence the uniqueness claimed in Theorem 3.5.

Both Theorem 6.3 of Nerman (1981) and (3.9) with sums taken over $u\in\mathcal{I}(s)\setminus\mathcal{I}(s,c)$, are results that hold almost surely, so the proof just outlined shows that the asymptotic equivalence in (3.9) holds almost surely in the general case. This is a Seneta-Heyde renormalization of the additive martingale $\sum_{u\in\mathcal{I}(s)} y_u$ that converges almost surely. In contrast, the renormalization of the original additive martingale $W^{(n)}$ $(=\sum_{|u|=n} y_u)$ is proved to converge using Lemma 2.3, which yields only convergence in probability.

It is worth giving a little more detail on how the embedded general branching process that is at the heart of the proof of uniqueness arises. Note first that $\mathcal{I}(0)$ is the set of individuals that are the first in their line of descent to have $-\log y_u>0$. To construct a suitable embedded process, regard these individuals as the children of the initial ancestor, with birth times $\{-\log y_u : u\in\mathcal{I}(0)\}$. Of course these individuals need not be in the first generation of the original process and so need not be true children of the initial ancestor. Consider the space-time lines $\theta p + t\log m(\theta) = s$ as s varies. When $s=0$ the collection $\mathcal{I}(0)$ are the individuals who are the first in their line of descent to be on the positive side of the line. This is illustrated in Figure 1 where the individuals coloured black belong to the set $\mathcal{I}(0)$. (Note that this figure shows a sample path of only the first six generations of some branching random walk).

As s increases the space-time line sweeps past individuals. Suppose that at s', the line passes through an individual $u\in\mathcal{I}(0)$, so that, for any $s<s'$, $u\in\mathcal{I}(s)$ but $u\notin\mathcal{I}(s')$. As u drops out of $\mathcal{I}(s)$ it is replaced by a collection of its descendants; this is formed by

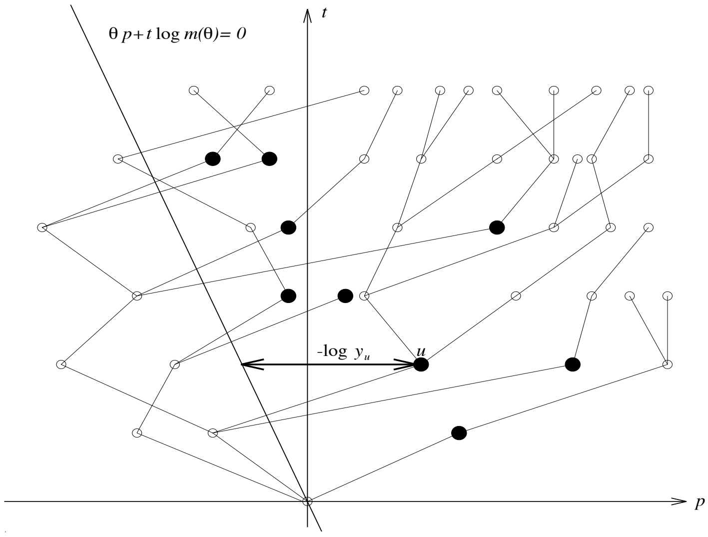

Figure 1:

the copy of $\mathcal{I}(0)$ associated with the descendants of u. It is these individuals that are to be regarded as the children of u in the next (embedded) generation. An example of this can be seen in Figure 2. The darkened branches indicate the tree generated by u and the individuals in this tree coloured black make up u's copy of $\mathcal{I}(0)$.

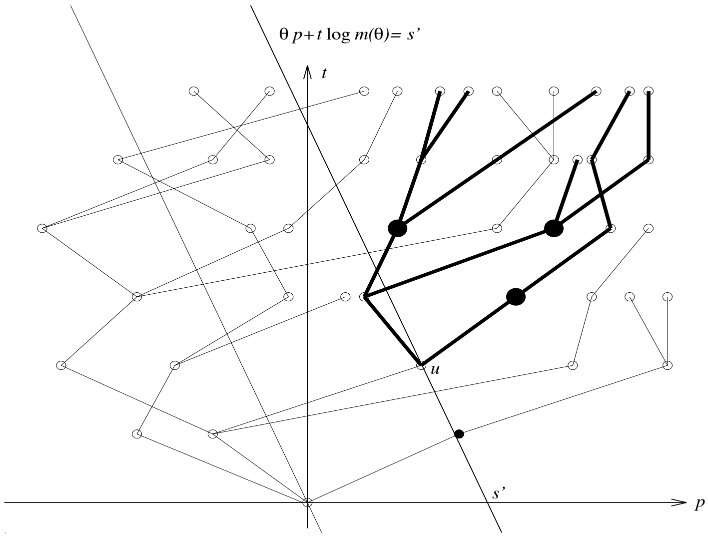

Figure 2:

In this way it is shown that $\{-\log y_u : u \in \mathcal{I}(s) \text{ for some } s > 0\}$ are the birth times for a general (Crump-Mode-Jagers) branching process, and $\mathcal{I}(s)$ is what is, in that context, often called the coming generation, which is, at time s, the individuals not yet born but whose mothers are. See Figure 3 in which individuals coloured black are members of the coming generation, $\mathcal{I}(s)$.

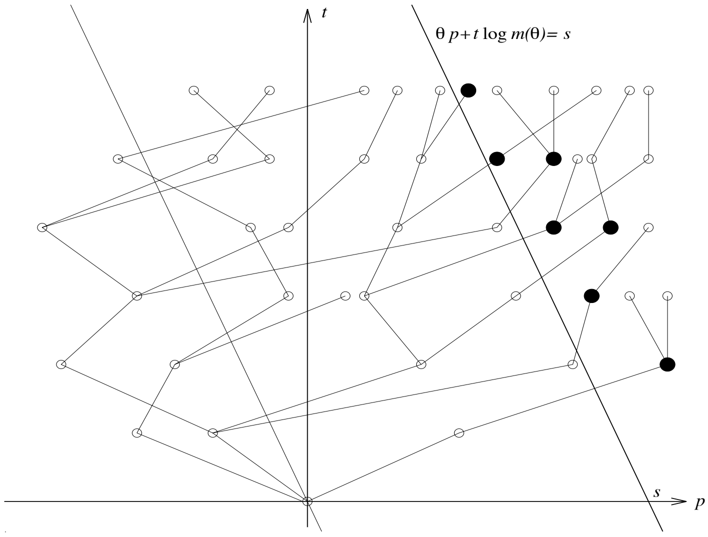

Figure 3:

4 Seneta-Heyde norming in the general branching process

By identifying position and birth-time, any general branching process can be treated as a branching random walk with the special feature that children are always born to the right of the parent. In this special case the embedded general branching process described in the previous section is the whole process. Then (3.9) gives the Seneta-Heyde normalization of Nerman's martingale, see Nerman (1981), with almost sure convergence. As Cohn (1985) points out, this is sufficient, because of Theorem 6.3 of Nerman (1981), to establish the Seneta-Heyde norming for any way of counting the process, providing the counting characteristic ϕ is sufficiently well behaved. (Jagers (1975) can be consulted for a discussion of the notion of characteristics.) Expressing this accurately gives the following result.

Theorem 4.1 *Consider a general (C-M-J) branching process with finite family sizes, reproduction intensity measure η, Malthusian parameter $\alpha > 0$, and with birth times $\{b_u\}$. Suppose that there is a $\beta \in [0, \alpha)$ such that $\int e^{-\beta\sigma}\eta(d\sigma) < \infty$. Then there is a slowly varying function L such that, for any characteristic $\phi(t)$ (with D-paths) and $E(\sup_t e^{-\beta t}\phi(t)) < \infty$,*

$$L(e^{-\alpha t})\sum_u \phi_u(t - b_u)$$

has an almost sure limit as $t \to \infty$. The limit is finite, and non-zero when the process survives.

This is an improvement on the result obtained by Cohn (1985) in that he assumed a finite mean family size, so that η is a finite measure, which is not needed in the approach outlined here. As mentioned in the Introduction, Theorem 4.1 illustrates that, though Theorem 3.2 gives only convergence in probability, the study yields, as a by-product, almost sure convergence, under improved conditions, in the nearest previously considered problem. This suggests that to strengthen the convergence in Theorem 3.2, if such a thing is possible, some new idea will be needed.

5 Branching Brownian motion

The purpose of this section is to outline the results that were the main stimulus for embarking on the developments described in the previous two sections. It starts by giving an account of a few of the results described in Neveu (1988).

Consider a branching Brownian motion with binary splitting starting from a single particle at the origin. Let $Z^{(t)}$ be the point process giving the occupied positions at time t. For $a \geqslant \sqrt{2}$, the differential equation

$$\frac{1}{2}\Phi'' + a\Phi' + (\Phi^2 - \Phi) = 0 \quad \text{with } \Phi(-\infty) = 0,\ \Phi(\infty) = 1, \tag{5.1}$$

has a unique solution, up to a translation. (Note that the more usual formulation, used in Neveu (1988), reflects the problem about the origin.) Then, for any x,

$$\exp\left(\int \log \Phi(at - x - p) Z^{(t)}(dp)\right) \tag{5.2}$$

forms a (multiplicative) martingale. For $a > \sqrt{2}$, as $x \to \infty$,

$$1 - \Phi(x) \sim ce^{bx} \text{ for suitable } b \text{ and } c. \tag{5.3}$$

This estimate implies, again for $a > \sqrt{2}$, that the logarithm of the martingale (5.2) is asymptotically equivalent to the (additive) martingale,

$$\int e^{-b(p-ta)} Z^{(t)}(dp). \tag{5.4}$$

Though the notation is different this is the same equivalence noted at the end of Step 2 in Section 3.

The product in (5.2) is over all particles present at time t. This can be regarded as a product over the positions where lines of descent first hit this time line. Neveu makes the pretty observation that if, instead, the product is taken over the positions where lines of descent first meet a suitable space-time line, specifically (p, t) such that $s = at - p$, all terms in the product are equal, and, as parallel space-time lines are considered, that is, as s changes, the number of terms in the product develops as a Markov branching process. (The theory needed to make this observation rigorous is given by Chauvin (1991).) This Markov branching process is the counterpart of the embedded general branching process that plays an important part in Step 3 in Section 3. As Brownian paths are continuous, overshoot of space-time lines is impossible, making some arguments easier.

Once the assumption of binary splitting is abandoned, in which case Φ^2 in (5.1) is replaced by $f(\Phi)$, where f is the family-size generating function, the equivalence Neveu noted between the additive and multiplicative martingales can break down. Assume the mean family size is finite. Then, because the differential equation has a solution, the multiplicative martingales exist and have non-trivial limits. However the continuous analogue of Theorem 3.1 shows that when the family size does not have a finite $X \log X$ moment the additive martingale converges to zero. Theorem 2.1 in Uchiyama (1978) can be consulted to see that, once $X \log X$ fails, the vital property (5.3) no longer holds. (Uchiyama's condition is, effectively, in terms of f, but its equivalence to the moment condition is standard.) Analogy with the Galton-Watson case suggested that when $X \log X$ fails the limit of a multiplicative martingale should provide the non-trivial limit for a suitable renormalization of the additive martingale. This study shows that this is indeed the case.

6 Continuous-time BRW

It is fairly easy, using a skeleton argument, to extend the discrete-time result given in Theorem 3.2 to the continuous-time Markov case, which includes branching Brownian motion discussed in the previous section. The initial ancestor lives for an exponential length of time, during which she moves according to a process with independent increments (with D-paths), at the end of this time she divides into children which are scattered about her present position according to a point process, X. Each of these children then lives, moves and divides in a similar way to the initial ancestor, independently of each other, and so on. In branching Brownian motion the movement is Brownian and X attaches weight only to the origin.

Let $Z^{(t)}$ be the point process at time t, and let $m(\phi) = E \int e^{-\phi x} Z^{(1)}(dx)$. The process obtained by sampling the continuous-time process only at integer times is a discrete-time branching random walk, so this definition of m fits with the earlier one. The Laplace transform m will be assumed to satisfy the same conditions as in the discrete case. Then

$$W^{(t)}(\theta) = m(\theta)^{-t} \int e^{-\theta x} Z^{(t)}(dx)$$

is a continuous-time martingale, which in branching Brownian motion is the martingale (5.4) in another notation. A discussion of the L_1 convergence of this martingale and of its convergence as a function of θ can be found in the final section of Biggins (1992).

Assume (3.2) holds. Then, by Theorem 3.2, which applies to the skeleton, there exists a sequence of constants $\{l_n\}$ such that $l_n W^{(n)}$ converges in probability to Δ. Let $[t]$ be the largest integer smaller than t and let $l_t = l_{[t]+1}$. It will now be shown that $l_t W^{(t)}$ also converges in probability to Δ.

Note first that, because $l_n W^{(n)}$ converges, for any finite x,

$$(l_{n+1} W^{(n+1)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x) \to 0 \text{ in probability.}$$

Using dominated convergence, this can be written in terms of Laplace transforms as

$$E[\exp(-u(l_{n+1} W^{(n+1)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x))] \to 1.$$

As $W^{(t)}$ is a martingale, two applications of Jensen's inequality give, for $n < t < n+1$,

$$\begin{aligned}
&E[\exp(-u(l_{n+1} W^{(n+1)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x))] \\
&\qquad = \; E[E[\exp(-u(l_{n+1} W^{(n+1)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x)) | \mathcal{F}^t]] \\
&\qquad \geqslant \; E[E[\exp(-u(l_{n+1} W^{(t)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x)) | \mathcal{F}^n]] \\
&\qquad \geqslant \; E[\exp(-u(l_{n+1} W^{(n)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x))] \\
&\qquad \to \; 1 \text{ as } n \to \infty.
\end{aligned}$$

Thus

$$(l_t W^{(t)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x))$$

converges in probability to 0 as $t \to \infty$. Now

$$\begin{aligned} |l_t W^{(t)} - \Delta| &\leqslant |l_n W^{(n)} - \Delta| + |(l_t W^{(t)} - l_n W^{(n)}) I(l_n W^{(n)} \leqslant x)| \\ &\quad + |(l_t W^{(t)} - l_n W^{(n)}) I(l_n W^{(n)} > x)|; \end{aligned}$$

the first two terms converge to zero in probability, whilst the final one is non zero with a probability close to $P(\Delta > x)$ for large n. As Δ is finite and x is arbitrary this last probability can be made as small as desired, thereby showing that, in probability,

$$l_t W^{(t)} \to \Delta.$$

When branching Brownian motion is approached in this way the functional equation (3.6) for the first generation of the discrete skeleton is used to introduce the multiplicative martingales, rather than the differential equation

$$\frac{1}{2}\Phi'' + a\Phi' + (f(\Phi) - \Phi) = 0 \quad \text{with } \Phi(-\infty) = 0,\ \Phi(\infty) = 1. \tag{6.1}$$

If Ψ solves the functional equation it is easy to show that it also solves the equation

$$\Psi(x) = E \exp \int \log \Psi\left(x \frac{e^{-\theta p}}{m(\theta)^t}\right) Z^{(t)}(dp).$$

It is then straightforward to show, following Appendix A of Bramson (1978), that, writing $\Phi(x) := \Psi(e^{-\theta x})$,

$$u(x,t) = E \exp \int \log \Phi\left(x + p\right) Z^{(t)}(dp) = \Phi\left(x - \frac{\log m(\theta)}{\theta} t\right)$$

solves the KPP equation $u_t = (1/2)u_{xx} + f(u) - u$ and consequently Φ solves (6.1) with $a = \theta^{-1} \log m(\theta)$.

References

[1] Asmussen, S. and Hering, H. (1983). *Branching Processes.* Birkhäuser, Boston.

[2] Biggins, J.D. (1977a). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14**, 25-37.

[3] Biggins, J.D. (1977b). Chernoff's Theorem in the branching random walk. *J. Appl. Probab.* **14**, 630–636.

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

[5] Biggins, J.D. and Kyprianou, A.E. (1996). Seneta-Heyde norming in the branching random walk. *Ann. Probab.* To appear.

[6] Bramson, M. (1979). Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31**, 531–581.

[7] Chauvin, B. (1988). *Arbres et Processus de Branchement.* These de Doctorat, Université Paris 6.

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

[9] Cohn, H. (1982a). Another look at the finite mean supercritical Bienaymé-Galton-Watson process. *J. Appl. Probab.* **19A**, 307–312.

[10] Cohn, H. (1982b). Norming constants for the finite mean supercritical Bellman-Harris process. *Z. Wahrsch. verw. Gebeite* **61**, 189–205.

[11] Cohn, H. (1983). On the convergence result for the supercritical Bellman-Harris process. *Aust. J. Statist.* **25**, 249–255.

[12] Cohn, H. (1985). A martingale approach to supercritical (CMJ) branching processes. *Ann. Probab.* **13**, 1179–1191.

[13] Grey, D.R. (1980). A new look at convergence of branching processes. *Ann. Probab.* **8**, 377–380.

[14] Heyde, C.C. (1970). Extension of a result of Seneta for the supercritical branching process. *Ann. Math. Statist.* **41**, 739–742.

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

[16] Kahane and Peyrière (1976). Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.* **22**, 131–145.

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

[18] Kurtz, T.G. (1972). Inequalities for the law of large numbers. *Ann. Math. Statist.* **43**, 1874–1883.

[19] Liu, Q. (1995). Fixed points of a generalized smoothing transformation and applications to branching processes. *J. Appl. Probab.* Submitted

[20] Lyons, R. (1996). A simple path to Biggins' martingale convergence. In *Classical and Modern Branching processes*, eds K.B. Athreya, P. Jagers, IMA proceedings **84**, 217–222. Springer-Verlag, New York.

[21] Nerman, O. (1981). On the convergence of supercritical general (C-M-J) branching process. *Z. Wahrsch. verw. Gebeite* **57**, 365–395.

[22] 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.

[23] Pakes (1992). On characterizations via mixed sums. *Austral. J. Statist.* **34**, 323–339.

[24] Seneta, E. (1968). On recent theorems concerning the supercritical Galton-Watson process. *Ann. Math. Statist.* **39**, 2098–2102.

[25] Seneta, E. (1969). Functional equations and the Galton-Watson process. *Adv. Appl. Probab.* **1**, 1–42.

[26] Schuh, H. J. (1982) Seneta constants for the supercritical Bellman-Harris process. *Adv. Appl. Probab.* **14**, 732–751.

[27] Uchiyama, K. (1978) The behaviour of solutions of some non-linear diffusion equations for large time. *J. Math. Kyoto Univ.* **18**, 453–508.