CONVERGENCE RESULTS ON MULTITYPE, MULTIVARIATE BRANCHING RANDOM WALKS

J. D. BIGGINS,* The University of Sheffield
A. RAHIMZADEH SANI,** Teacher Training University of Tehran

Abstract

We consider a multitype branching random walk on \(d\)-dimensional Euclidian space. The uniform convergence, as \(n\) goes to infinity, of a scaled version of the Laplace transform of the point process given by the \(n\)th generation particles of each type is obtained. Similar results in the one-type case, where the transform gives a martingale, were obtained in Biggins (1992) and Barral (2001). This uniform convergence of transforms is then used to obtain limit results for numbers in the underlying point processes. Supporting results, which are of interest in their own right, are obtained on (i) ‘Perron–Frobenius theory’ for matrices that are smooth functions of a variable \(\lambda \in L\) and are nonnegative when \(\lambda \in L^− \subset L\), where \(L\) is an open set in \(\mathbb{C}^d\), and (ii) saddlepoint approximations of multivariate distributions. The saddlepoint approximations developed are strong enough to give a refined large deviation theorem of Chaganty and Sethuraman (1993) as a by-product.

Keywords: Large deviation; local limit theorem; saddlepoint approximation; uniform convergence; multitype branching random walk; Laplace transform; martingale; Perron–Frobenius; nonnegative matrix

2000 Mathematics Subject Classification: Primary 60J80; Secondary 15A48; 60F10

1. Introduction

We consider the multitype branching random walk on the \(d\)-dimensional Euclidian space, \(\mathbb{R}^d\). The process starts with a single particle located at the origin. This particle produces daughter particles, which are scattered in \(\mathbb{R}^d\), to give the first generation; these first-generation particles produce daughter particles to give the second generation; and so on. As usual in branching processes, the \(n\)th-generation particles reproduce independently of each other. Each particle in this process is of one of \(p\) types; the set of possible types is identified with \(\{1, \ldots, p\}\). For each \(i \in \{1, \ldots, p\}\) there is a vector of point processes \((Z_{i1}, Z_{i2}, \ldots, Z_{ip})\). Then, when a particle of type \(i\) reproduces, the positions of its daughter particles of the various types, relative to the parent’s position, are given by a copy of \((Z_{i1}, Z_{i2}, \ldots, Z_{ip})\). The one-type branching random walk has received extensive treatment in the literature. The multitype extension has received less attention, but discussion of it can be found in Mode (1971), Biggins (1976), (1996), Bramson et al. (1992), and Kyprianou and Rahimzadeh Sani (2001).

Received 19 January 2005; revision received 5 April 2005.
* Postal address: Department of Probability and Statistics, The University of Sheffield, Sheffield S3 7RH, UK. Email address: j.biggins@sheffield.ac.uk
** Postal address: Department of Mathematics, Teacher Training University of Tehran, 49 Mofatteh Avenue, Tehran, 15614, Iran. Email address: rahimsan@saba.tmu.ac.ir
We reserve the notation $i$, $j$, and $k$, drawn from $\{1, \ldots, p\}$, for particle types. Let $\mu_{ij}$ be the intensity measure of $Z_{ij}$. We assume throughout that

$$\text{there exists a } \theta \in \mathbb{R}^d \text{ such that } \max_{i,j} \int_{\mathbb{R}^d} e^{-\theta^\top x} \mu_{ij}(dx) < \infty, \quad (1)$$

where $\theta^\top x = \sum_j \theta_i x_i$ is the usual inner product of vectors. This condition is enough to ensure that convolutions of the $\mu_{ij}$ produce well-defined measures.

Technically, the description of the process outlines how to produce the probability measure on the space of trajectories, given the type of the single initial ancestor. There are $p$ such measures, one for each possible initial type. Suppose that the initial ancestor has type $i$ and let $Z_{ij}^n$ be the point process giving the resulting positions of the type-$j$ particles in generation $n$; then $Z_{ij}^1$ is distributed like $Z_{ij}$. The first objective here is to obtain the asymptotic behaviour, as $n \to \infty$, of the Laplace transform of $Z_{ij}^n$. Note that results for $Z_{ij}^n$ are, in fact, results for the measure on trajectories obtained when the single initial ancestor is of type $i$. Incorporating the initial type into the notation in this way, rather than through the measure as $P_i$ and $E_i$, is helpful in calculations.

Before our theorems can be stated, further notation for intensity measures and their transforms is needed. Define $\mu_{ij}^{(n+1)*}$ inductively by

$$\mu_{ij}^{(n+1)*} = \sum_{k=1}^p \mu_{ik} \ast \mu_{kj}^{(n)*}$$

where ‘*’ denotes ordinary convolution of measures. It is easy to confirm, by induction on $n$, that the point process $Z_{ij}^n$ has the intensity measure $\mu_{ij}^{(n)*}$. Furthermore, the counts of the numbers of each type in each generation, given by $(Z_{ij}^1(\mathbb{R}^d), \ldots, Z_{ij}^p(\mathbb{R}^d))$, is a multitype Galton–Watson process, which is discussed in Athreya and Ney (1972), for example. A matrix $A$ of nonnegative entries is called positive regular when, for some positive integer $n$, all the entries of $A^n$ are strictly positive. The multitype Galton–Watson process is positive regular when the matrix $P(Z_{ij}(\mathbb{R}^d) > 0)$ is positive regular. Throughout, the embedded Galton–Watson process is assumed to be positive regular. This Galton–Watson process is supercritical when the largest eigenvalue of its mean matrix, $E(Z_{ij}(\mathbb{R}^d))$, exceeds 1. (For a positively regular process, this eigenvalue will be infinite when the mean matrix has any infinite entries.) A supercritical process survives with positive probability. One of the conditions of our theorems will imply that the process is supercritical.

The $d$-dimensional complex space $\mathbb{C}^d$ is equipped with the maximum metric. Hence, for $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d) \in \mathbb{C}^d$, we have $|x, y| = \max_i |x_i - y_i|$, where $|x| = \max_i |x_i|$ is the usual absolute value in $\mathbb{C}$. Let $B(x, r)$ be the open ball centred at $x$ and of radius $r$ using this metric, and let $\overline{B}(x, r)$ be its closure. (Later, we will also need $S(x, r)$ in $\mathbb{R}^d$, the open and closed balls in $\mathbb{R}^d$.) In all that follows, we reserve the letters $\theta$ and $\eta$ for the real and imaginary parts of $\lambda \in \mathbb{C}^d$, so that, with this convention, $\lambda = \theta + i\eta$. For any $A \subset \mathbb{C}^d$, let $A_-$ be its intersection with $\mathbb{R}^d$; thus

$$A_- = \{ \lambda \in A : \lambda = \theta + i\eta \text{ with } \eta = 0 \}.$$

Define the Laplace transforms $m_{ij}(\lambda)$, for $\lambda \in \mathbb{C}^d$, by

$$m_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda^\top x} \mu_{ij}(dx) = \int_{\mathbb{R}^d} e^{-\lambda^\top x} Z_{ij}(dx).$$

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
Then \(|m_{ij}(\lambda)| \leq m_{ij}(\theta)|\) and, by Hölder’s inequality, \(\{\theta \in \mathbb{R}^d : m_{ij}(\theta) < \infty\}\) is a convex set. Let \(\text{int } A\) be the interior of \(A\) and let
\[
L = \bigcap_{i,j} \text{int}\{\lambda = \theta + i\eta \in \mathbb{C}^d : m_{ij}(\theta) < \infty\}.
\]

Assumption (1) is now strengthened to the following:

\[
\text{Assumption (1) is now strengthened to the following:}
\]

Theorem 1. Suppose that the \(m_{ij}(\lambda)|\) and, by Hölder’s inequality, \(\{\theta \in \mathbb{R}^d : m_{ij}(\theta) < \infty\}\) is a convex set. Let \(\text{int } A\) be the interior of \(A\) and let
\[
L = \bigcap_{i,j} \text{int}\{\lambda = \theta + i\eta \in \mathbb{C}^d : m_{ij}(\theta) < \infty\}.
\]

Assumption (1) is now strengthened to the following:

\[
L = \text{nonempty}.
\]

In fact, \(L\) is a nonempty, open, convex subset of \(\mathbb{C}^d\) and each \(m_{ij}(\lambda)|\) is analytic in \(\lambda \in L\).

Let \(M(\lambda)\) be the matrix given by \(M(\lambda) = (m_{ij}(\lambda)|)\) and let \(M^n(\lambda)\) be its \(n\)th power, with \((i, j)\)th entry \(m^n_{ij}(\lambda)|\). Furthermore, let
\[
M^n_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda^\top x} Z^n_{ij}(dx).
\]

Then
\[
m^n_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda^\top x} \mu^n_{ij}(dx) = E M^n_{ij}(\lambda).
\]

By analogy with the case in which the entries are nonnegative, the eigenvalue \(\rho\) of \(M\) is called the maximum-modulus eigenvalue if it is a simple root of \(\det[z I - M]\) and its modulus is strictly larger than that of all the other roots (\(I\) being the identity matrix). Note that, when a maximum-modulus eigenvalue exists, it is automatically unique.

When \(\theta \in L_- (= L \cap \mathbb{R}^d\), the entries of \(M(\theta)|\) are finite, nonnegative real numbers. The positive regularity of the embedded Galton–Watson process clearly implies that \(M(\theta)|\) is then positively regular. This means that the following extension of Perron–Frobenius theory which, is discussed in the final section, applies to \(M\). In essence, in the present context, it says that Perron–Frobenius properties extend smoothly to complex arguments with suitably small imaginary parts. Part (iii) of Theorem 1, below, gives an asymptotic estimate, \((2)\), of \(M^n(\lambda)|\) as \(n \to \infty\). The key point about \((2)\) is that it gives a uniform bound on the rate of convergence in a suitable neighbourhood. This uniformity has no direct parallel in Perron–Frobenius theory and requires a little care to establish.

Theorem 1. Suppose that the \(p \times p\) matrix \(M = (m_{ij})\) of functions defined on the open set \(L \subset \mathbb{C}^d\) satisfies the following conditions: for all \(i\) and \(j\), \(m_{ij}(\lambda)|\) are analytic functions in \(\lambda \in L\), and, for all \(\lambda \in L \subset L\), \(M(\lambda)|\) is positive regular. Then there is an open set \(\Omega \subset L\) containing \(L\) such that, for \(\lambda \in \Omega\), the following statements hold.

(i) \(M(\lambda)|\) has a maximum-modulus eigenvalue, \(\rho(\lambda)|\), that is analytic in \(\lambda\).

(ii) The left- and right-eigenvectors associated with \(\rho(\lambda)|\), namely \(u(\lambda)|\) and \(v(\lambda)|\), normalised so that \(\sum_{i=1}^p u_i(\lambda) = 1\) and \(\sum_{i=1}^p u_i(\lambda)v_i(\lambda) = 1\), are analytic in \(\lambda\) and, for all \(i\), \(u_i(\lambda) \neq 0\) and \(v_i(\lambda) \neq 0\).

(iii) For any \(\theta \in L\), there is a neighbourhood, \(B \in \Omega\), containing \(B(\theta, \delta)\) for some \(\delta > 0\), and constants \(K < \infty\) and \(\gamma \in (0, 1)\) such that, for all \(n\), \(i\), and \(j\),
\[
\sup_{\lambda \in B} |\rho(\lambda)^{-n}(M(\lambda)\rho))_{ij} - v_i(\lambda)u_j(\lambda)| \leq K \gamma^n.
\]
All the notation introduced in Theorem 1 is now applied to $M$, the matrix of Laplace transforms, with $L = L$. Thus, $M$ has maximum-modulus eigenvalue $\rho$ and eigenvectors $u$ and $v$ defined on the set $\Omega$, with the properties described in Theorem 1. For all $\lambda \in \Omega$, $n = 1, 2, \ldots$, and $i$ and $j$, define

$$W^n_{ij}(\lambda) = v_j(\lambda)/v_i(\lambda) \rho(\lambda)^{-n} M^n_{ij}(\lambda) = \frac{v_j(\lambda)}{v_i(\lambda)} \rho(\lambda)^{-n} \int_{\mathbb{R}^d} e^{-\lambda x} Z^n_{ij}(dx),$$

(3)

which is a scaled version of the Laplace transform of the point process $Z^n_{ij}$. Establishing the uniform convergence of such transforms in the one-type case is an integral part of the approach to results on the distribution of the points described in Biggins (1992). For the multitype case, results on the distribution of points were obtained in Bramson et al. (1992), where particles were confined to the integer lattice of $\mathbb{R}$. More precise results, of the kind given in Biggins (1992), should hold and should not be limited to the lattice case. One such result is given in Theorem 7, below.

Let $\mathcal{F}_i$ be the $\sigma$-algebra that contains all information on the multitype branching random walk up to generation $n$. In the one-type case, (3) defines a martingale with respect to $\mathcal{F}_n$, which makes aspects of the study simpler. In the multitype case, it is the sequence

$$W^n_i(\lambda) = \sum_j W^n_{ij}(\lambda) = \sum_j \frac{v_j(\lambda)}{v_i(\lambda)} \rho(\lambda)^{-n} \int_{\mathbb{R}^d} e^{-\lambda x} Z^n_{ij}(dx),$$

(4)

that is a martingale; this is well known and proved here in Lemma 1. Our approach to the convergence of $\{W^n_i(\lambda)\}$ involves first considering the convergence of the martingale $\{W^n_i(\lambda)\}$. Note that, given any sample path of the process, each $W^n_i(\lambda)$ and $W^n_i(\lambda)$ is analytic in $\lambda \in \Omega$.

We now introduce certain sets that will be used to define the region where convergence occurs. For $\alpha \in (1, 2]$, let

$$\Omega^\alpha_\omega = \left\{ \lambda \in \Omega : \alpha \lambda \in \Omega, \frac{\rho(\alpha \theta)}{\rho(\lambda)^{\alpha}} < 1 \right\},$$

(5)

$$\Omega^\alpha_\omega = \text{int} \left\{ \lambda = \theta + i\eta \in \Omega : \max_i \mathbb{E}[|W^n_i(\theta)^\alpha|] < \infty \right\},$$

(6)

$$\Lambda_\alpha = \Omega^\omega_\omega \cap \Omega^\alpha_\omega \quad \text{and} \quad \Lambda = \bigcup_{1<\alpha\leq2} \Lambda_\alpha.$$ 

These are all open sets in $\mathbb{C}^d$. Error estimates we derive involve $(\rho(\alpha \theta)/\rho(\lambda)^{\alpha})^n$. Hence, having $\lambda \in \Omega^\omega_\omega$ ensures that such bounds decay quickly with $n$. Obviously, having $\lambda \in \Omega^\alpha_\omega$ imposes a moment condition. The approach requires $\lambda$ to satisfy both these conditions for the same $\alpha \in (1, 2]$, which leads to the definition of $\Lambda_\alpha$ and then $\Lambda$.

The process is supercritical, that is, $\rho(\theta) > 1$, whenever there exists a $\theta \in \Omega^\omega_\omega$ for some $\alpha \in (1, 2]$. To check this, first note that $\log \rho(\theta)$ is a convex function; see Kingman (1961), Miller (1961), or Seneta (1973, Theorem 3.7). Hence,

$$\frac{(\alpha - 1)}{\alpha} \log \rho(\theta) + \frac{1}{\alpha} \log \rho(\alpha \theta) \geq \log \rho(\theta),$$

and $\log(\rho(\alpha \theta)/\rho(\theta)^{\alpha}) < 0$ when $\theta \in \Omega^\omega_\omega$. Then $\log(\theta) > 0$ and, so, $\rho(\theta) > 1$. 

**Theorem 2.** Let $\alpha \in (1, 2]$ and $\lambda = \theta + i\eta \in \Lambda_\alpha$. Then $\{W^n_i(\lambda)\}$ converges almost surely and in $\alpha$th mean, as $n \to \infty$, for each $i$. 

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
This gives the following result on mean convergence as an immediate consequence.

**Corollary 1.** Suppose that $\lambda \in \Lambda$. Then $\{W^n_i(\lambda)\}$ converges almost surely and in mean, as $n \to \infty$, for each $i$.

The mean convergence of $\{W^n_i(\lambda)\}$ for real $\lambda$ was also discussed in Bramson *et al.* (1992), Kyprianou and Rahimzadeh Sani (2001), and, rather briefly, in Biggins and Kyprianou (2004).

**Theorem 3.** Assume that $\Lambda_{-} \neq \emptyset$. Then there is an open set $\Gamma_1$, with $\Lambda_{-} \subset \Gamma_1 \subset \Lambda$, such that, for each $i$, the martingale $\{W^n_i(\lambda)\}$ converges uniformly in any compact subset of $\Gamma_1$, almost surely and in mean, as $n \to \infty$. Furthermore, the limit $W_i(\lambda)$ is analytic in $\lambda \in \Gamma_1$.

By building on these convergence results for the martingale, we will obtain analogous results for $\{W^n_i(\lambda)\}$.

**Theorem 4.** Let $\alpha \in (1, 2]$ and assume that $(\Lambda_{a})_{-} \neq \emptyset$. Then there is an open set $\Gamma_1$, with $(\Lambda_{a})_{-} \subset \Gamma_1 \subset \Lambda_{a}$, such that, for all $\lambda \in \Gamma_1$, $\{W^n_i(\lambda)\}$ converges to $u_j(\lambda)v_j(\lambda)W_i(\lambda)$, almost surely and in mean, as $n \to \infty$, where $W_i(\lambda)$ is the limit of the martingale $\{W^n_i(\lambda)\}$ as $n \to \infty$.

**Theorem 5.** Assume that $\Lambda_{-} \neq \emptyset$. Then there is an open set $\Gamma_2$, with $\Lambda_{-} \subset \Gamma_2 \subset \Lambda$, such that, for all $i, j$, $\{W^n_{ij}(\lambda)\}$ converges uniformly in any compact subset of $\Gamma_2$, almost surely and in mean, as $n \to \infty$, to the random variable $u_j(\lambda)v_j(\lambda)W_i(\lambda)$, where $W_i(\lambda)$ is the limit of the martingale $\{W^n_i(\lambda)\}$ as $n \to \infty$.

The sequences $\{W^n_{ij}(\lambda)\}$ and $\{W^n_{ij}(\lambda)\}$ converge on the set $\Gamma_2$ introduced in Theorem 5. However, to move from information on transforms to information on the associated measures, a result on the behaviour of $M^n_{ij}(\lambda)$ is needed for $\theta \in \Lambda_{-}$, but in the case where $\lambda$ need not be in $\Gamma_2$.

The branching random walk is strongly nonlattice when it is positively regular and, for some $(k, l)$ and some $\theta \in \{\theta: m_{kl}(\theta) < \infty\}$,

$$\left|\frac{m_{k\bar{i}}(\theta + i\eta)}{m_{kl}(\theta)}\right| = 1 \quad \text{only when } \eta = 0. \quad (7)$$

This follows the usage of ‘strongly nonlattice’ in Bhattacharya (1977, Equation (1.64)), rather than that in Stone (1965).

**Theorem 6.** For a strongly nonlattice branching random walk, for any set $\mathcal{K} \subset \Lambda_{-}$ that is compact in $\mathbb{R}^d$ and any $\epsilon \in (0, 1)$, there is an $\epsilon < 1$ such that

$$\epsilon^{-n} \sup_{i, j} \sup_{\theta \in \mathcal{K}} \sup_{\epsilon \leq |\eta| \leq \infty - 1} \left|\frac{M^n_{ij}(\theta + i\eta)}{\rho^n(\theta)}\right| \to 0 \quad \text{as } n \to \infty, \quad (8)$$

almost surely.

Between them, Theorems 5 and 6 provide enough information on the behaviour of the transforms $M^n_{ij}(\lambda)$ to develop good estimates of the associated measures. In the one-type case, in Biggins (1992), this step relies heavily on the corresponding results for sums of independent, identically distributed random variables, obtained in Stone (1967). Unfortunately, this approach does not extend directly, since the existing results for Markov additive processes, which are here the analogue of independent, identically distributed random variables, are not suitable. Instead,
Theorem 7. Assume that \( \log \rho \) is strictly convex and that the process is strongly nonlattice. Let \( \mathcal{K} \) be a compact subset of \( \Lambda_- \). Let \( h \) be such that, for some finite \( G \) and all \( \theta \in \mathcal{K} \), the function \( e^{\theta \top h(x)} \) is in \( \mathcal{G}(G) \). Then, with

\[
\xi(\theta) = -\theta \top v_1(\theta) - \log \rho(\theta)
\]

and \( C \) any convex set in \( \mathbb{R}^d \),

\[
n^{d/2}e^{\xi(\theta)} \int_C h(x)Z^n_{ij}(dx + n v_1(\theta)) \to \frac{v_j(\theta)u_j(\theta)W_i(\theta)}{(2\pi)^d \det[v_2(\theta)]} \int_C e^{\theta \top x} h(x) \, dx
\]

uniformly in \( \theta \in \mathcal{K} \) and \( C \), almost surely on \( \mathcal{S} \), where \( \mathcal{S} \) is the survival set of the underlying Galton–Watson process.

The condition that \( e^{\theta \top h(x)} \) be in \( \mathcal{G}(G) \) becomes more restrictive as \( \mathcal{K} \) becomes larger. In particular, it forces \( h(x) \) to decay rapidly with \( |x| \) when the origin is in the interior of \( \mathcal{K} \).

Corollary 2. Let \( b < \infty \). Under the conditions of Theorem 7,

\[
n^{d/2}e^{\xi(\theta)} Z^n_{ij}(C + n v_1(\theta)) \to \frac{v_j(\theta)u_j(\theta)W_i(\theta)}{(2\pi)^d \det[v_2(\theta)]} \int_C e^{\theta \top x} \, dx
\]

uniformly in convex \( C \subset [x : |x| \leq b] \) and \( \theta \in \mathcal{K} \), almost surely on \( \mathcal{S} \).

It is worth pointing out that if, for some \( y \), \( \mathcal{K} \) lies inside the half-space \( \{ \theta : \theta \top y > 0 \} \), then the result of this corollary can also hold for some sets \( C \) that are not bounded.

The relationships between \( \rho \), \( v_1 \), and \( \xi \) are the usual ones associated with large deviation and saddlepoint approximations; see Jensen (1995, Section 2.2), for example. The assumption that \( \rho \) is strictly convex means that compact subsets of \( \Lambda_- \) translate, under \( y = v_1(\theta) \), into compact subsets of \( \text{int}(v_1(\theta) : \theta \in L_-) \).
Theorem 7 has the following analogue for the measures $\mu_{ij}^{\alpha}$, with the proof requiring only obvious changes.

**Theorem 8.** Under the conditions of Theorem 7,

$$n^{d/2}e^{\alpha x} \int_C h(x)\mu_{ij}^{\alpha}(dx + n\nu_1(\theta)) \to \frac{v_{ij}(\theta)\mu_j(\theta)}{(2\pi)^d \det[\nu_2(\theta)]} \int_C e^{\theta x}h(x)\,dx$$

uniformly in $\theta \in \mathcal{K}$ and $C$.

2. Proofs of Theorems 2 and 3

For any $\lambda = \theta + i\eta \in \Omega$, define the functions

$$\bar{\nu}(\lambda) = \max_{i,j} \left| \frac{v_j(\lambda)}{v_i(\lambda)} \right|, \quad \nu(\lambda) = \min_{i,j} \left| \frac{v_j(\lambda)}{v_i(\lambda)} \right|, \quad v(\lambda) = \frac{\nu(\lambda)}{\bar{\nu}(\lambda)},$$

and

$$\phi(\lambda) = \frac{\rho(\theta)}{|\rho(\lambda)|^a}.$$

All of these are strictly positive, continuous functions of $\lambda \in \Omega$, where $\Omega$ is as in Theorem 1. Also, since $|m_{ij}(\lambda)| \geq m_{ij}(\theta)$, we know that $|\rho(\lambda)| \leq \rho(\theta)$ (Lancaster and Tismenetsky (1985, Theorem 15.2.1)) and, so, $\phi(\lambda) \geq 1$.

For each $\alpha \in (1, 2]$, define $\Omega_\alpha$ by $\Omega_\alpha = \{\lambda \in \Omega : a\lambda \in \Omega\}$. Then $\Omega_\alpha$ is an open subset of $\Omega$. For $\alpha \in (1, 2]$ and $\lambda \in \Omega_\alpha$, let

$$v(\lambda) = \bar{\nu}(\lambda)^a \quad \text{and} \quad \kappa(\lambda) = \frac{\rho(\alpha \theta)}{|\rho(\lambda)|^a}.$$

Then $v_1$ and $\kappa$ are strictly positive, continuous functions of $\lambda \in \Omega_\alpha$; they depend on $\alpha$, but this has been suppressed in the notation. Note that $\Omega_\alpha^2$ is defined in (5), can now be written as $\Omega_\alpha^2 = \{\lambda \in \Omega_\alpha : \kappa(\lambda) < 1\}$ and is an open subset of $\Omega_\alpha$. Define

$$\beta(\theta) = \max_i \mathbb{E}[|W_i^1(\theta)|^a].$$

Then $\Omega_\alpha^3$, defined in (6), can be written as $\Omega_\alpha^3 = \text{int}(\lambda = \theta + i\eta \in \Omega : \beta(\theta) < \infty)$ and $\beta$ is a real-valued, continuous function on $\Omega_\alpha^3$.

Let $\alpha \geq 1$. Define the $\alpha$th absolute central moment of a random (complex-valued) variable $X$, $\sigma^\alpha(X)$, by

$$\sigma^\alpha(X) = \mathbb{E}[|X - \mathbb{E}[X]^a|],$$

and the $\alpha$th absolute central moment conditional on the $\sigma$-algebra $\mathcal{G}$ by

$$\sigma^\alpha(X \mid \mathcal{G}) = \mathbb{E}[|X - \mathbb{E}[X \mid \mathcal{G}]|^a \mid \mathcal{G}].$$

Let $\{z_{i_k, l, s}^j : s\}$ be the positions of the particles making up $Z_{i_k}^l$, where by $z_{i_k, l, s}^j$ we denote the position of the $s$th particle of type $k$ in generation $l$ when the initial ancestor is of type $i$. Now, by looking at the particles in the $n$th generation as the $(n - l)$th-generation children of the particles in generation $l$, we can introduce $C_{i_k}^l(\cdot | l, s)$ as the point process giving the positions of the type-$j$ particles in generation $n$ descended from the $s$th particle of type $k$ in generation $l$, relative to the
latter particle’s position. Given $\mathcal{F}_t$, $\{Z^n_{k_j}(\cdot | l, s) : s\}$ are independent copies of the point process $Z^{n-l}_{k_j}$. Thus, conditional on $\mathcal{F}_t$, the random variables

$$W^n_{k-1}(\lambda | l, s) = \sum_j v_j(\lambda) \frac{1}{\rho(\lambda)^{n-l}} \int_{\mathbb{R}^d} e^{-x \sum_{k}^n Z^n_{k_j} (dx | l, s)}$$

are, as $s$ varies, independent and identical copies of $W^n_{k-1}(\lambda)$. Furthermore,

$$W^n_l(\lambda) = \sum_k v_k(\lambda) \frac{1}{\rho(\lambda)^l} \sum_s \exp(-\lambda s) W^{n-l}_{k}(\lambda | l, s).$$

**Lemma 1.** $\{W^n_l(\lambda)\}$ is a martingale with respect to $\{\mathcal{F}_t\}$.

*Proof.* From Theorem 1, $E W^n_{1}(\lambda) = 1$. Therefore, decomposition (10) with $l = n - 1$ gives the result.

The following result, which is an extension of von Bahr and Esseen (1965, Equation (4)) to complex-valued random variables, was proved in Biggins (1992). Barral (2001) provided a very pretty approach to uniform convergence in the one-type case by applying a lemma like this to Banach-space-valued random variables.

**Lemma 2.** If $\{X_i\}$ are independent, complex-valued random variables with $E X_i = 0$ or, more generally, martingale differences, then for $\alpha \in (1, 2]$ there exists a constant $C > 0$, depending on $\alpha$ but independent of $\lambda$ and the sequence $\{X_n\}$, such that

$$E\left[\sum_{j=1}^n X_j|^\alpha\right] \leq C \sum_{j=1}^n E|X_j|^\alpha.$$

The next lemma, which is the multitype version of Biggins (1992, Lemma 2), gives some bounds related to the martingale $\{W^n_i(\lambda)\}$. In bounding formulae like these, the argument $\lambda$ will often be omitted; thus, for example, $\phi \equiv \phi(\lambda)$.

**Lemma 3.** Let $\alpha \in (1, 2]$, $\lambda = \theta + i \varphi \in \Lambda_\alpha$, and $\psi = v_1 \varphi \psi \theta$. Then there are constants $c_1$, $c_2$, and $c_3$, depending on $\alpha$ but not on $\lambda$, $i$, or $n$, such that the following statements hold.

(i) $E |W^{n+1}_i(\lambda) - W^n_i(\lambda)|^\alpha \leq E[|W^{n+1}_i(\lambda) - W^n_i(\lambda)|^\alpha] \leq c_1 \psi \kappa^n$.

(ii) $\sigma^\alpha(W^n_i(\lambda)) \leq c_2 \psi (1 - \kappa)^{-1}$.

(iii) $\sum_{k=0}^n E |W^{k+1}_i(\lambda) - W^n_i(\lambda)| \leq c_3 \psi^{1/\alpha} \kappa^{n/\alpha} (1 - \kappa^{1/\alpha})^{-1}$.

(iv) For $\theta \in (\Lambda_\alpha)_-$, there exists a $\delta > 0$ such that

$$\sup\{E |W^n_i(\lambda)|^\alpha : \lambda \in B(\theta, \delta), i, n\} < \infty.$$

*Proof.* The first inequality in part (i) is just $E[|X|^\alpha] \geq (E[X]^\alpha)$. Since $E[|W^1_i(\lambda)|^\alpha] \geq 1$, we obtain

$$E[|W^1_i(\lambda) - 1|^\alpha] \leq 4 E[|W^1_i(\lambda)|^\alpha] \leq 4 \phi^\alpha \psi^\alpha \theta.$$
Using (10), Lemma 2, and the fact that, given $\mathcal{F}_n$, $\mathcal{W}_k^n(\lambda|n, s)$ are independent copies of $\mathcal{W}_k^n(\lambda)$, we have

$$
\sigma^\alpha(\mathcal{W}_{i+1}^n(\lambda) \mid \mathcal{F}_n) = E[|\mathcal{W}_{i+1}^n(\lambda) - \mathcal{W}_i^n(\lambda)|^\alpha \mid \mathcal{F}_n]
$$

\[ \leq C \sum_k \frac{e_k(\lambda)}{v_i(\lambda)} \beta(k^n) \sum_j |\exp(-\lambda^T z_i^{k;\lambda})|^\alpha E[|\mathcal{W}_k^n(\lambda) - 1|^\alpha]
\]

\[ \leq c_1 \psi \kappa^n \mathcal{W}_i^n(\alpha \theta) \phi^n v^\alpha
\]

Taking expectations now gives part (i).

Note that, since $\lambda \in \Lambda_\alpha$, we have $\kappa < 1$. Part (ii) now follows directly from summing over $n$ in part (i) and using Lemma 2. To prove part (iii), sum the inequalities in part (i).

The continuity of the functions $\nu$, $\nu_1$, $\phi$, and $\kappa$ on $\Lambda_\alpha$ and part (ii) mean that there exists a constant $C > 0$, depending only on $\alpha$ and $\delta > 0$, such that, for any $\lambda \in \overline{B}(\theta, \delta)$, we have $\sigma^\alpha(\mathcal{W}_i^n(\lambda)) \leq C$. Since $E[|\mathcal{W}_i^n(\lambda)|^\alpha] \leq 2 + 2\sigma^\alpha(\mathcal{W}_i^n(\lambda))$, this implies part (iv).

**Proof of Theorem 2.** From part (iii) of Lemma 3, $E \sum_{n=0}^{\infty} |\mathcal{W}_i^{n+1}(\lambda) - \mathcal{W}_i^n(\lambda)|$ is finite and, so, $\{\mathcal{W}_i^n(\lambda)\}$ is a Cauchy sequence almost surely. Thus, as $n \to \infty$, $\{\mathcal{W}_i^n(\lambda)\}$ converges to $W_i(\lambda)$ almost surely. Applying Fatou’s lemma, Lemma 2, and Lemma 3(i) gives

$$
E[|W_i(\lambda) - \mathcal{W}_i^n(\lambda)|^\alpha] \leq \lim inf_{N \to \infty} E[|\mathcal{W}_i^{N+1}(\lambda) - \mathcal{W}_i^n(\lambda)|^\alpha]
$$

\[ \leq C \sum_{j=0}^{\infty} E[|\mathcal{W}_i^{n+j+1}(\lambda) - \mathcal{W}_i^{n+j}(\lambda)|^\alpha
\]

\[ \leq C c_1 \psi \kappa^n \frac{k^n}{1 - k}.
\]

Since $\kappa < 1$, $\{\mathcal{W}_i^n(\lambda)\}$ converges in $\alpha$th mean.

The distinguished boundary (see Hörmander (1973)) of $B(x, r)$, which is a subset of the topological boundary, is defined by

$$
D(x, r) = \{ y \in \mathbb{C}^d: |x_s - y_s| = r, s = 1, \ldots, d \}.
$$

We assume $D(x, r)$ to be parametrized such that

$$
D(x, r) = \{ z(t) = (z_1(t), \ldots, z_d(t)): z_s(t) = x_t + re^{i\theta}, s = 1, \ldots, d, t \in I \},
$$

where $I = [0, 2\pi]$ is a $d$-dimensional closed cube in $\mathbb{R}^d$. The next lemma, which is Biggins (1992, Lemma 3), is the key to obtaining bounds that hold uniformly.

**Lemma 4.** If $f$ is analytic on the open ball $B = B(x, 2r)$ and $r < r'$, then

$$
\sup_{\lambda} |f(\lambda)| \leq \pi^{-d} \int_I |f(z(t))| dt,
$$

where $z(t) \in D(x, 2r)$.
Proof of Theorem 3. The proof closely follows that of the one-type result of Biggins (1992). Let \( x \in \Lambda_\gamma \) be fixed. There exist an \( \alpha \in (1, 2] \) and a \( B = B(x, 3r) \subset \Lambda_\alpha \), which means that \( B \subset \Omega^2_\alpha \) and \( B \subset \Omega^3_\alpha \). For any two positive integers \( n \) and \( N \geq n \), we apply Lemma 4 to \( W^{N+1}_i(\lambda) - W^m_i(\lambda) \) to obtain

\[
\pi^d \sup_{\lambda \in B(x,r)} |W^{N+1}_i(\lambda) - W^m_i(\lambda)| \leq \int \sum_{m=n}^{N} |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt,
\]

\[
\leq \sum_{m=n}^{\infty} \int |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt,
\]

where \( z(t) \in D(x, 2r) \subset B \). Applying Fubini’s theorem and Lemma 3(iii), we obtain

\[
E \sum_{m=n}^{\infty} \int |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt \leq \int C \rho^{1/\alpha} \frac{\kappa n/\alpha}{1 - \kappa 1/\alpha} \, dt,
\]

where the argument \( z(t) \) has been suppressed on the right-hand side. Recall that \( \kappa < 1 \) throughout \( \Omega^2_\alpha \) and \( \beta < \infty \) throughout \( \Omega^3_\alpha \). Therefore, using the continuity of the various functions, there exist constants \( \delta < 1 \) and \( K < \infty \) such that the integrand on the right-hand side can be bounded by \( K \delta n \) throughout \( B \) and, hence, throughout \( D(x, 2r) \). Thus,

\[
E \sum_{m=n}^{\infty} \int |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt \leq K \delta n \int \, dt < \infty
\]

and so,

\[
\sum_{m=n}^{\infty} \int |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt < \infty \quad \text{almost surely.}
\]

Therefore, \( \{W^m_i(\lambda)\} \) is a Cauchy sequence under the supremum norm \( \| \cdot \| \) on \( B(x, r) \), which implies uniform convergence on \( B(x, r) \), almost surely.

For the remainder of the proof, the supremum norm is defined over \( B(x, r) \). From the almost-sure uniform convergence of the martingale, for fixed \( n \), as \( N \to \infty \),

\[
\|W^N_i - W^m_i\| \to \|W_i - W^m_i\|
\]

almost surely. Let \( n \) be fixed; then, by taking the expectation of both sides and applying first Fatou’s lemma and then Lemma 4, we obtain

\[
E \|W_i - W^m_i\| \leq \liminf_{N \to \infty} E \|W^N_i - W^m_i\|
\]

\[
\leq \pi^{-d} E \int \sum_{m=n}^{\infty} |W^{m+1}_i(z(t)) - W^m_i(z(t))| \, dt
\]

\[
\leq \pi^{-d} K \delta n \int \, dt < \infty
\]

\[
\to 0 \quad \text{as } n \to \infty.
\]

This proves uniform convergence on \( B(x, r) \), in mean.

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
Finally, define

$$
\Gamma = \bigcup_{x \in \Lambda} B(x, r).
$$

Then, by a compactness argument, the uniform convergence of \( \{ \omega^n_l(\lambda) \} \) in the open balls in \( \Gamma \) implies the uniform convergence in any compact subset of \( \Gamma \).

### 3. Proofs of Theorems 4 and 5

Let \( 1 \leq l \leq n \) and let \( i \) and \( j \) be fixed. We start with a decomposition that is similar to (10),

$$
\omega^n_{ij}(\lambda) = \sum_k \sum_s v_k(\lambda) \frac{1}{\rho(\lambda)} \exp(-\lambda \tau^l_{ik,s}) \omega^n_{kj}(\lambda;l,s),
$$

(11)

where, given \( \mathcal{F}_l \), \( \omega^n_{kj}(\lambda;l,s) \) are independent copies of \( \omega^n_{kj}(\lambda) \) as \( s \) varies. Hence,

$$
\mathbb{E}[\omega^n_{ij}(\lambda) \mid \mathcal{F}_l] = \frac{v_j(\lambda)}{v_i(\lambda)\rho(\lambda)} \sum_k \sum_s \exp(-\lambda \tau^l_{ik,s}) \omega^n_{kj}(\lambda).
$$

(12)

Therefore, given any sample path of the process, \( \mathbb{E}[\omega^n_{ij}(\lambda) \mid \mathcal{F}_l] \) is an analytic function of \( \lambda \in \Omega \).

For all \( \lambda \in \Omega \), define

$$
g^n_{ij}(\lambda;l) = \omega^n_{ij}(\lambda) - \mathbb{E}[\omega^n_{ij}(\lambda) \mid \mathcal{F}_l]
$$

and

$$
h^n_{ij}(\lambda;l) = \mathbb{E}[\omega^n_{ij}(\lambda) \mid \mathcal{F}_l] - u_j(\lambda)v_j(\lambda)\omega^n_{ij}(\lambda).
$$

The functions \( g^n_{ij}(\lambda;l) \) and \( h^n_{ij}(\lambda;l) \) are analytic functions of \( \lambda \in \Omega \). Furthermore,

$$
|\omega^n_{ij}(\lambda) - u_j(\lambda)v_j(\lambda)\omega^n_{ij}(\lambda)| \leq |g^n_{ij}(\lambda;l)| + |h^n_{ij}(\lambda;l)| + |u_j(\lambda)v_j(\lambda)|(|\omega^n_{ij}(\lambda) - \omega^n_{ij}(\lambda)|).
$$

Hence, \( |\omega^n_{ij}(\lambda) - u_j(\lambda)v_j(\lambda)\omega^n_{ij}(\lambda)|^a \) is less than or equal to

$$
3(|g^n_{ij}(\lambda;l)|^a + |h^n_{ij}(\lambda;l)|^a + |u_j(\lambda)v_j(\lambda)|^a |\omega^n_{ij}(\lambda) - \omega^n_{ij}(\lambda)|^a).
$$

(13)

The idea now is to let \( l = \lceil \frac{x}{n} \rceil \), where \( \lfloor x \rfloor \) is the greatest integer not exceeding \( x \), and let \( n \) tend to infinity. This motivates the lemmas we now give before returning to the main proof. It is in these results that we need the uniformity proved in Theorem 1(iii).

**Lemma 5.** Suppose that \( \theta \in (\Lambda_\varphi) \). Then there are positive constants \( c_1, c_2, \) and \( c_3 \) depending on \( \alpha \), a constant \( \gamma < 1 \), and a neighbourhood \( B \) of \( \theta \) such that, for all \( \lambda \in B \), \( n = 1, 2, \ldots, \) and all \( i \) and \( j \), we have

(i) \( |h^n_{ij}(\lambda;l)| \leq c_1 \gamma^{n-l} \nu \phi^n \omega^n_{ij}(\theta) \),

(ii) \( E \sum_{l=0}^n |h^n_{ij}(\lambda;l)|^a \leq c_2 \phi^n \), and

(iii) \( \sigma^n(\omega^n_{ij}(\lambda)) \leq c_3 \phi^{a\alpha} \).

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
Therefore, by Theorem 1(iii), there exist constants $c_1 > 0$ and $\gamma \in (0, 1)$ and a neighbourhood $B$ of $\Theta$ such that, for all $\lambda \in B$ and all $n$,

$$|h_{ij}^n(\lambda)| \leq c_1 v_1^\phi \gamma^{n-\ell} \sum_{k} \sum_{s} \left[ \frac{v_k(\theta)}{v_1(\theta)} \phi(\theta)^{-\ell} \exp(-\theta^T \tilde{z}_{k,s}^l) \right].$$

as required. This and Lemma 3(iv) combine, after making $B$ smaller if necessary, to show that, for some constant $c_1'$, $E[|h_{ij}^n(\lambda)|^\alpha] \leq c_1' v_1^\alpha (\gamma/\phi)^{(n-\ell)\alpha}$. Since $\gamma < 1$ and $\phi \geq 1$, we have $(\gamma/\phi) < 1$ for all $\lambda \in B$. Thus,

$$E \sum_{l=0}^\infty |h_{ij}^n(\lambda)|^\alpha \leq c_1' \frac{1}{1 - (\gamma/\phi)\alpha}. $$

The continuity of the functions $v$ and $\phi$ now implies part (ii). Finally, using Lemma 2,

$$\sigma^\alpha(W_{ij}^n(\lambda)) \leq C \sum_{l=1}^n E[|E[W_{ij}^n(\lambda) | F_l] - E[W_{ij}^n(\lambda) | F_{l-1}]|^\alpha]$$

$$\leq 3^\alpha C \left( |u_j(\lambda)v_j(\lambda)|^\alpha \sum_{l=1}^n E[|W_{ij}^l(\lambda) - W_{ij}^{l-1}(\lambda)|^\alpha] + 2 E \sum_{l=0}^n |h_{ij}^n(\lambda)|^\alpha \right).$$

By suitable bounding of the continuous functions involved in Lemma 3(i), for all $\lambda \in B(\Theta, \delta)$ with $\delta$ sufficiently small, we have

$$\sum_{l=1}^{\infty} E[|W_{ij}^l(\lambda) - W_{ij}^{l-1}(\lambda)|^\alpha] \leq K_1 < \infty$$

and, so, using this and part (ii), we have

$$\sigma^\alpha(W_{ij}^n(\lambda)) \leq 3^\alpha C (|u_j(\lambda)v_j(\lambda)|^\alpha K_1 + 2c_2\phi^{n\alpha}).$$

Now the boundedness of $u_j(\lambda)v_j(\lambda)$ on $B$ and the fact that $\phi \geq 1$ combine to prove part (iii).

**Lemma 6.** Suppose that $\alpha \in (1, 2]$ and $\Theta \in (\Lambda_\alpha)_\sim$. Then there exists a $B = B(\Theta, \delta) \subset \Lambda_\alpha$ such that, for $l = \left\lfloor \frac{1}{n} \right\rfloor$,

(i) for all $\lambda \in B$, $h_{ij}^n(\lambda)|l| \to 0$ as $n \to \infty$, almost surely and in $\alpha$th mean; and

(ii) $\sup(|h_{ij}^n(\lambda)|: \lambda \in B) \to 0$ as $n \to \infty$, almost surely and in mean.

**Proof.** Let $B = B(\Theta, \delta)$ be a neighbourhood of $\Theta$ throughout which (14) holds. Take $\delta$ to be smaller if necessary, so that $\phi^{\gamma} \leq 1$ throughout $B$. Now application of Theorems 2 and 3 to $W_{ij}(\Theta)$ together with simple bounding in (14) gives all the claimed results.
Lemma 7. Suppose that \( \alpha \in (1, 2) \) and \( \vartheta \in (\Lambda_\alpha) \). Then there exists a neighbourhood of \( \vartheta \), say \( B = B(\vartheta, r) \subset \Lambda_\alpha \), such that, for \( l = \lfloor \frac{1}{2} n \rfloor \),

(i) for all \( \lambda \in B \), \( g^n_{ij}(\lambda l) \to 0 \) as \( n \to \infty \), almost surely and in \( \text{arth} \) mean; and

(ii) \( \sup \{|g^n_{ij}(\lambda l)|: \lambda \in B\} \to 0 \) as \( n \to \infty \), almost surely and in mean.

Proof. First note that, using (11) and Lemma 2,

\[
\sigma^\alpha(W^n_{ij}(\lambda) \mid F_l) \leq C u(\lambda)^\alpha |\rho(\lambda)|^{-l\alpha} \sum_k \sum_s \exp(-\alpha \theta \zeta_{k,s}^l) \sigma^\alpha(W^n_{kl}(\lambda)). \tag{15}
\]

Let \( B = B(\vartheta, \delta) \) be inside the neighbourhood of \( \vartheta \) introduced in Lemma 5. Then, for some positive constant \( c \) and for all \( \lambda \in B, n = 1, 2, \ldots, \) and all \( i \) and \( j \), applying Lemma 5(iii) to the bound (15) gives

\[
\sigma^\alpha(W^n_{ij}(\lambda) \mid F_l) \leq cv_1 W^n_{ij}(\alpha \theta) \phi^{\alpha(n-l)} l^l. \tag{16}
\]

Now, since \( E W^n_{ij}(\alpha \theta) = 1 \), we have

\[
E\{g^n_{ij}(\lambda l)|F_l\} = E\{\sigma^\alpha(W^n_{ij}(\lambda) \mid F_l)\} \leq cv_1 \phi^{\alpha(n-l)} l^l.
\]

Choose \( \delta \) to be smaller if necessary, so that \( \phi^{\alpha\delta} < 1 \) for all \( \lambda \in B \). Then the right-hand side of the above inequality converges to 0 geometrically quickly, as \( n \to \infty \), when \( l = \lfloor \frac{1}{2} n \rfloor \). This implies that \( g^n_{ij}(\lambda l) \to 0 \) in \( \text{arth} \) mean and almost surely.

Jensen’s inequality and (16) give

\[
E |g^n_{ij}(\lambda l)| \leq (E \sigma^\alpha(W^n_{ij}(\lambda) \mid F_l))^{1/\alpha} \leq c^{1/\alpha} v_1^{1/\alpha} \phi^{\alpha-l} \delta^{l/\alpha}.
\]

For \( l = \lfloor \frac{1}{2} n \rfloor \), \( \phi^{\alpha-l} \delta^{l/\alpha} \) is asymptotically equivalent to \( (\phi \delta)^{1/\alpha} l^l, \) and \( \phi \delta^{1/\alpha} \leq \gamma_1 < 1 \) in \( B \).

Hence, there exists a \( C' < \infty \) such that, for all \( \lambda \in B, n = 1, 2, \ldots, l = \lfloor \frac{1}{2} n \rfloor \), and all \( i \) and \( j \), we have \( E |g^n_{ij}(\lambda l)| \leq C' \gamma_1^l \). Let \( 2r < \delta \) and apply Lemma 4 to the analytic functions \( g^n_{ij}(\lambda l) \) to obtain

\[
\sup_{k \in B(\vartheta, r)} |g^n_{ij}(\lambda l)| \leq \pi^{-d} \int_t |g^n_{ij}(z(t)|l)| \, dt,
\]

where \( z(t) \in D(\vartheta, 2r) \). The expectation of the right-hand side here goes to 0 geometrically quickly. Hence, \( g^n_{ij}(\lambda l) \) converges to 0 uniformly in \( \lambda \in B(\vartheta, r) \), almost surely and in mean.

Proof of Theorems 4 and 5. In (13), let \( l = \lfloor \frac{1}{2} n \rfloor \) and then let \( n \to \infty \). In a suitable neighbourhood of \( \vartheta \in (\Lambda_\alpha) \), the almost-sure and \( \text{arth} \) and mean convergence of \( W^n_{ij}(\lambda) - W_{ij}(\lambda), g^n_{ij}(\lambda l) \), and \( h^n_{ij}(\lambda l) \) are contained in Theorem 2, Lemma 7(i), and Lemma 6(i), respectively. Now take the union of these neighbourhoods to be \( \Gamma_1 \). This proves Theorem 4.

Similarly, the almost-sure and mean uniform convergence of \( W^n_{ij}(\lambda) - W_{ij}(\lambda), g^n_{ij}(\lambda l) \), and \( h^n_{ij}(\lambda l) \) are contained in Theorem 3, Lemma 7(ii) and Lemma 6(ii), respectively. A union of suitable neighbourhoods now provides \( \Gamma_2 \), proving Theorem 5.

4. Proof of Theorem 6

Lemma 8. If the branching random walk is strongly nonlattice then, for any compact set \( K \subset \{ \theta \in \mathbb{R}^d : \sup_{i,j} m_{ij}(\theta) < \infty \} \) and all \( a \in (0, 1) \), there exists an \( \varepsilon < 1 \) such that

\[
\varepsilon^{-n} \sup_{i,j} \sup_{\theta \in K, a \leq |\theta| \leq a^{-1}} \left| \frac{m^n_{ij}(\theta + i \eta)}{\rho^n(\theta)} \right| \to 0 \text{ as } n \to \infty. \tag{17}
\]
Proof. Let \((k, l)\) be such that (7) holds. As a function of \(\eta\), \(m_{kl}(\theta + i\eta)/m_{kl}(\theta)\) is the characteristic function of the probability measure \(m_{kl}(\theta)^{-1}e^{-\theta^\top x}\mu_{kl}(dx)\) and (7) implies that this measure is strongly nonlattice. Then, using this and continuity, we have
\[
\tilde{\epsilon} = \sup_{\theta \in \mathcal{K}} \sup_{a \leq |\eta| \leq a^{-1}} \left| \frac{m_{kl}(\theta + i\eta)}{m_{kl}(\theta)} \right| < 1.
\]

Let \(M^\dagger(\theta)\) be the matrix with \((k, l)\)th entry \(\tilde{\epsilon}m_{kl}(\theta)\) and with other entries the same as those in \(M(\theta)\), and let \((M^\dagger(\theta)^n)_{ij}\) be the \((i, j)\)th entry of its \(n\)th power. Let \(\rho^\dagger(\theta)\) be the eigenvalue of \(M^\dagger(\theta)\) of maximum modulus, with right-eigenvector \(v^\dagger(\theta)\). Then \(\rho^\dagger(\theta) < \rho(\theta)\) (see Seneta (1973, Theorem 1.1)). Furthermore, since Theorem 1 applies to \(M^\dagger(\theta)\), there exists an \(\epsilon < 1\) such that \(\rho^\dagger(\theta) < \epsilon \rho(\theta)\) for all \(\theta \in \mathcal{K}\), and
\[
\sup_{\theta \in \mathcal{K}} \max_{i,j} |v^\dagger_i(\theta) v^\dagger_j(\theta)| = c < \infty.
\]

Now,
\[
\sup_{a \leq |\eta| \leq a^{-1}} |(M(\theta + i\eta)^n)_{ij}| \leq (M^\dagger(\theta)^n)_{ij} \leq c(\epsilon \rho(\theta))^n,
\]
proving (17).

Proof of Theorem 6. It is sufficient to prove (8) with the compact set \(\mathcal{K}\) being some closed ball in \(\mathbb{R}^d\) centred on \(\theta \in \Lambda_{-\alpha}\), since a simple covering argument extends the result to general \(\mathcal{K}\). We begin the proof by introducing various bounds leading to the appropriate ball to consider. Recall that \(S(x, r)\) is a ball in \(\mathbb{R}^d\) centred on \(x\) with radius \(r\), and \(\overline{S}(x, r)\) is its closure.

Choose \(\theta \in \Lambda_{-\alpha}\); then \(\theta \in \Lambda_\alpha\) for some \(\alpha\), which is now fixed. Now take \(\delta_1\) such that \(S(\theta, 2\delta_1) \subset \Lambda_\alpha\) and \(S_1 = \overline{S}(\theta, \delta_1)\). Let \(\epsilon_1\) be an \(\epsilon\) such that (17) holds when \(\mathcal{K} = \{\theta : \theta \in S_1\}\) and let \(a = \frac{1}{2} \epsilon\) (where \(\epsilon\) is as defined in the statement of the theorem). Recall that
\[
\frac{\rho(\alpha \theta)^{1/\alpha}}{\rho(\theta)} < 1, \quad \theta \in \Lambda_\alpha.
\]

Let
\[
\epsilon_2 = \max \left\{ \epsilon_1, \sup_{\theta \in S_1} \frac{\rho(\alpha \theta)^{1/\alpha}}{\rho(\theta)} \right\}.
\]

Now let \(S_2 = S(\theta, 2\delta_2)\) with \(2\delta_2 < \delta_1\), and let
\[
\overline{\rho} = \sup_{\theta \in S_2} \rho(\theta) \quad \text{and} \quad \underline{\rho} = \inf_{\theta \in S_2} \rho(\theta).
\]

Take \(\delta_2\) small enough that
\[
\epsilon_3 := \epsilon_2 \frac{\sup_{\theta \in S_2} \rho(\theta)}{\inf_{\theta \in S_2} \rho(\theta)} = \epsilon_2 \frac{\overline{\rho}}{\underline{\rho}} < 1.
\]

Now let \(S_3 = \{\theta : \theta \in S_2\}\) and \(K = \{\theta : \theta \in S_3\}\), and let \(\epsilon_4\) be such that (18) holds when \(\mathcal{K} = \{\theta : \theta \in S_3\}\) and let \(a = \frac{1}{2} \epsilon\) (where \(\epsilon\) is as defined in the statement of the theorem). Recall that
\[
\frac{\rho(\alpha \theta)^{1/\alpha}}{\rho(\theta)} < 1, \quad \theta \in \Lambda_\alpha.
\]

Let
\[
\epsilon_2 = \max \left\{ \epsilon_1, \sup_{\theta \in S_1} \frac{\rho(\alpha \theta)^{1/\alpha}}{\rho(\theta)} \right\}.
\]

Now let \(S_2 = S(\theta, 2\delta_2)\) with \(2\delta_2 < \delta_1\), and let
\[
\overline{\rho} = \sup_{\theta \in S_2} \rho(\theta) \quad \text{and} \quad \underline{\rho} = \inf_{\theta \in S_2} \rho(\theta).
\]

Take \(\delta_2\) small enough that
\[
\epsilon_3 := \epsilon_2 \frac{\sup_{\theta \in S_2} \rho(\theta)}{\inf_{\theta \in S_2} \rho(\theta)} = \epsilon_2 \frac{\overline{\rho}}{\underline{\rho}} < 1.
\]

Note that, for \(\theta \in S_2\),
\[
\rho(\alpha \theta)^{1/\alpha} \leq \epsilon_2 \rho(\theta) \leq \epsilon_2 \overline{\rho}.
\]
These preliminaries provide a suitably small neighbourhood of \( \theta, S_2 \), in which various inequalities hold. Now let \( S = \overline{S(\theta, \delta_2)} \subset S_2 = \overline{S(\theta, 2\delta_2)} \) and define the two regions
\[
R = \{ \theta + i\eta: \theta \in S, \epsilon \leq |\eta| \leq \epsilon^{-1} \} \quad \text{and} \quad R_2 = \{ \theta + i\eta: \theta \in S_2, \frac{\epsilon}{2} \leq |\eta| \leq (\frac{1}{2}\epsilon)^{-1} \}.
\]
We will prove (8) for \( \theta \in S \), that is, for \( \lambda \in R \). Clearly,
\[
\sup_{\lambda \in R} |\mathcal{M}_{ij}^{\alpha}(\lambda)| \leq \sup_{\lambda \in R} |\mathcal{M}_{ij}^{\alpha}(\lambda) - m_{ij}^{\alpha}(\lambda)| + \sup_{\lambda \in R} |m_{ij}^{\alpha}(\lambda)|. \tag{20}
\]
Furthermore,
\[
\sup_{\lambda \in R} |m_{ij}^{\alpha}(\lambda)| \leq \sup_{\lambda \in R_2} |m_{ij}^{\alpha}(\lambda)| = \sup_{\lambda \in R_2} \left| \frac{m_{ij}^{\alpha}(\lambda)}{\rho(\lambda)^{\alpha}} \right| \leq (\epsilon_1 \rho)^{\alpha} \leq (\epsilon_2 \rho)^{\alpha}. \tag{21}
\]
Using Lemma 2, we have
\[
E[|\mathcal{M}_{ij}^{\alpha}(\lambda) - m_{ij}^{\alpha}(\lambda)|^\alpha] \leq C \sum_{l=0}^{n-1} E[[|\mathcal{M}_{ij}^{\alpha}(\lambda) | | \mathcal{F}_l+1] - E[|\mathcal{M}_{ij}^{\alpha}(\lambda) | | \mathcal{F}_l]^\alpha], \tag{22}
\]
and splitting at the \( l \)th generation shows that
\[
E[|\mathcal{M}_{ij}^{\alpha}(\lambda) | | \mathcal{F}_l] = \sum_k \left( \sum_x \exp(-\lambda^T z_{ik}) \right) m_{ij}^{n-l}(\lambda)
\]
\[
= \sum_k \left( \int_{\mathbb{R}^d} \exp(-\lambda^T z) Z_k^I(dz) \right) m_{ij}^{n-l}(\lambda). \tag{23}
\]
Following (9), let
\[
\mathcal{M}_{kj}^{n-l}(\lambda | l, s) = \int_{\mathbb{R}^d} e^{-\lambda^T x} Z_k^n(dx | l, s);
\]
then
\[
\int_{\mathbb{R}^d} \exp(-\lambda^T z) Z_k^{l+1}(dz) = \sum_h \sum_s \exp(-\lambda^T z_{ik}) \mathcal{M}_{hk}^{l}(\lambda | l, s). \tag{24}
\]
Let \( \Delta_l = E[|\mathcal{M}_{ij}^{\alpha}(\lambda) | | \mathcal{F}_l+1] - E[|\mathcal{M}_{ij}^{\alpha}(\lambda) | | \mathcal{F}_l] \). Then, using (23) and (24), we have
\[
\Delta_l = \sum_k \left( \sum_h \sum_s \exp(-\lambda^T z_{ik}) (\mathcal{M}_{hk}^{l}(\lambda | l, s) - m_{hk}(\lambda)) \right) m_{ij}^{n-l-1}(\lambda)
\]
and, given \( \mathcal{F}_l \), \( [\mathcal{M}_{hk}^{l}(\lambda | l, s) - m_{hk}(\lambda): s] \) is independent variables with zero mean. We need a bound on the moments of these variables. Let
\[
c_1 = 4 \left( \max_{i, j: \theta \in S_2} E[|\mathcal{M}_{ij}(\theta)|^\alpha] \right)^{1/\alpha},
\]
which is finite because \( S_2 \subset \Lambda_{2\alpha} \). Then, simple bounding gives
\[
\sup \{ E[|\mathcal{M}_{ij}(\lambda) - m_{ij}(\lambda)|^\alpha]: \lambda = \theta + i\eta, \theta \in S_2, i, j \} \leq (c_1)^\alpha.
\]
Let
\[ c_2 = \max_{i,j} \sup_{\theta \in S_2} \frac{v_j(\alpha \theta)}{v_i(\alpha \theta)}, \]
which is finite, by Theorem 1(ii), because \( S_2 \subset \Lambda \alpha \). Then, using the fact that
\[ m_{ij}^n(\alpha \theta) v_j(\alpha \theta) \leq \rho(\alpha \theta)^n v_i(\alpha \theta) \]
and (19), for \( \theta \in S_2 \) we have
\[ m_{ij}^n(\alpha \theta) \leq c_2(\epsilon \rho)^n. \]
Now, applying Lemma 2 again, (21), and the two bounds just given, for any \( \lambda \in R_2 \) we have
\[
\mathbb{E}[|\Delta_{l-1}|^\alpha] \leq C \sum_k \left( \sum_h \left[ m_{ih}^{l-1}(\alpha \theta) \mathbb{E}[|\mathcal{M}_{hk}^1(\lambda) - m_{hk}(\lambda)|^\alpha] \right] m_{kj}^{n-l}(\lambda)^\alpha \right)
\leq C p^2 c_2(\epsilon \rho)^{(l-1)\alpha} (c_1)\alpha (\epsilon \rho)^{(n-l)\alpha}
\leq c_3(\epsilon \rho)^{n\alpha},
\]
where \( c_3 \) is independent of \( \theta \in S_2 \) and \( l \). Hence, (22) gives
\[ \mathbb{E}[|\mathcal{M}_{ij}^n(\lambda) - m_{ij}^n(\lambda)|^\alpha] \leq n c_3(\epsilon \rho)^{n\alpha} \]
for any \( \lambda \in R_2 \).

The region \( R \) can be covered with a finite number of open balls such that, when their radii are doubled, they still lie inside \( R_2 \). Hence, using (22), the bound just obtained, and Lemma 4, for a suitable constant \( c_4 \) we have
\[ \mathbb{E} \sup_{\lambda \in R} |\mathcal{M}_{ij}^n(\lambda) - m_{ij}^n(\lambda)| \leq c_4(\epsilon \rho)^{n\alpha} n^{1/\alpha}. \]
Combining this with (20) and (21) gives
\[ \mathbb{E} \sup_{\lambda \in R} |\mathcal{M}_{ij}^n(\lambda)| \leq (c_4 n^{1/\alpha} + c_1)(\epsilon \rho)^n \]
and, so, using (18), we have
\[ \mathbb{E} \sup_{\lambda \in R} \left| \frac{\mathcal{M}_{ij}^n(\lambda)}{\rho(\theta)^n} \right| \leq \frac{1}{\rho^n} \mathbb{E} \sup_{\lambda \in R} |\mathcal{M}_{ij}^n(\lambda)| \leq (c_4 n^{1/\alpha} + c_1) \epsilon. \]
Hence, (8) holds for any \( \epsilon \in (\epsilon_3, 1) \).

5. Approximation of measures

This section contains the preparatory work on saddlepoint approximations necessary for the proof of Theorem 7. The idea is to develop explicit estimates that apply to a particular measure through only a few of its attributes. Most treatments of these matters, made with a view to application to variables that are, or look like, sums of independent, identically distributed random variables, bring \( n \) into the picture sooner than we do. The treatment draws on ideas from Stone (1967), von Bahr (1967), Bhattacharya (1972), (1977), Chaganty and Sethuraman (1993), and Jensen (1995).
Recall that \( S(x, \epsilon) = \{ y \in \mathbb{R}^d : |x - y| < \epsilon \} \). Let \( \Gamma \) be a probability measure on \( \mathbb{R}^d \), with \( y := \Gamma(S(0, 1)) > \frac{1}{2} \), whose characteristic function vanishes outside \( \{ t : |t| \geq \zeta \} \) for some finite \( \zeta \). Let \( \Gamma_\epsilon \) be the measure given by \( \Gamma_\epsilon(A) = \Gamma(\{ |x| : \epsilon x \in A \}) \), so that its characteristic function vanishes outside \( \{ t : |t| \geq \epsilon^{-1} \zeta \} \). Let \( \nu \) and \( \mu \) be probability measures on \( \mathbb{R}^d \), let \( b_\epsilon \) be the supremum of the modulus of the density of \( (\nu - \mu) \ast \Gamma_\epsilon \), and let \( q \) be the supremum of the density of \( \mu \). For any \( f \), let

\[
\omega_f(x, \epsilon) = \sup \{|f(y) - f(z)| : y, z \in S(x, \epsilon)\}
\]

and

\[
\mathcal{I}(x) = \sup \{|f(y)| : y \in S(x, \epsilon)\}.
\]

The next result is contained in, and easily derived from, Lemma 2.2 of Bhattacharya (1972) and the remark following it.

**Lemma 9.** For any real-valued, bounded, Borel-measurable function \( f \), we have

\[
\sup_{x \in \mathbb{R}^d} \left| \int f(x)(v - \mu)(dx + u) \right| \leq \frac{1}{2\gamma - 1} \left[ b_\epsilon \int \mathcal{I}(x) \, dx + q \int \omega_f(x, 2\epsilon) \, dx \right].
\]

The idea now is to make this bound simpler by imposing suitable smoothness conditions on \( f \). The set \( \mathcal{G}(G) \) is made up of those \( f \in \mathcal{D} \) for which there is a bounded decreasing function \( g \) with \( \int_0^\infty g(r)(r + 1)^{d-1} \, dr = G \) and

\[
\sup \{|f(x)|, |\nabla f(x)| : |x| \geq (r - 1)\} \leq g(r).
\]

Note that \( \mathcal{G}(G) \), introduced to state Theorem 7, is defined in the same way, but with

\[
\int_0^\infty g(r)(r + 1)^{d+1} \, dr = G;
\]

therefore, \( \mathcal{G}(G) \subset \mathcal{G}'(G) \). Also, let \( \mathcal{C} \) denote the convex sets in \( \mathbb{R}^d \).

**Lemma 10.** For a constant \( \Delta \), depending only on the dimension, and \( \epsilon < 1 \), we have

\[
\sup \left\{ \left| \int_C f(x)(v - \mu)(dx + u) \right| : u \in \mathbb{R}^d, C \subset \mathcal{C}, f \in \mathcal{G}'(G) \right\} \leq G\Delta(b_\epsilon + \epsilon q).
\]

**Proof.** For a set \( A \subset \mathbb{R}^d \), let \( A^\epsilon = \{ y : |y - a| < \epsilon, a \in A \} \), and let \( \partial A \) be the boundary of \( A \). Let \( \mathcal{S} \) be the surface area of the unit ball. Then, for \( C \subset \mathcal{C} \), from von Bahr (1967, Equation (15)), we have

\[
\int 1_{A^\epsilon}(x)g(|x|) \, dx \leq 2\mathcal{S}G\epsilon.
\]

Take \( \epsilon < 1 \) and let \( h(x) = f(x)1_{C}(x) \), with \( f \in \mathcal{G}'(G) \). If \( x \in \partial A^\epsilon \) then

\[
\omega_h(x, \epsilon) \leq 2 \sup \{|f(y)| : y \in S(x, \epsilon)\} \leq 2g(|x|).
\]

If \( x \in S(x, \epsilon) \subset C \) then, using Taylor’s theorem, \( \omega_h(x, \epsilon) = \omega_f(x, \epsilon) \leq 2\epsilon g(|x|) \). Hence,

\[
\omega_h(x, \epsilon) \leq 2g(|x|)(1_{A^\epsilon}(x) + \epsilon 1_{C}(x))
\]
and so,
\[ \int \omega_\theta(x, \epsilon) \, dx \leq 2 \int (1_{\partial C^c}(x) + \epsilon \, 1_C(x)) g(|x|) \, dx \leq 6\epsilon \delta. \]

Also,
\[ \int \tilde{h}_\theta(x) \, dx \leq \int \tilde{f}_\epsilon(x) \, dx \leq \int g(|x|) \, dx = \delta. \]

Substituting these into the bound in Lemma (9) gives the result.

Recall that, for \( A \subset \mathbb{C}^d \), \( A_- \) is its intersection with \( \mathbb{R}^d \). Let \( \mathcal{K} \) be compact in \( \mathbb{R}^d \), let \( \tau < 1 \), and let
\[ \mathcal{K}_\tau = \{ \theta + \lambda \in \mathbb{C}^d : \theta \in \mathcal{K}, |\lambda| \leq \tau \}. \]

Now let \( B \) be an open set containing \( \mathcal{K}_\tau \), with \( B_- \) convex. For \( Z \) a measure on \( \mathbb{R}^d \), let
\[ \tilde{Z}(\lambda) = \int e^{-\lambda^\top x} \, Z(dx). \]

Suppose that \( \log \tilde{Z} \) is analytic on \( B \) and strictly convex on \( B_- \). Let \( -m(\theta) \) be the vector of its first derivatives and \( \Sigma(\theta) \) the matrix of its second derivatives on \( B_- \). Then \( \Sigma(\theta) \) is positive definite, because \( \log \tilde{Z} \) is strictly convex. Let \( u \) be a (finite) bound on the modulus of all derivatives of \( \log \tilde{Z} \) over \( \mathcal{K}_\tau \), up to and including order three. Let \( \epsilon \) be a lower bound on the smallest eigenvalue of \( \Sigma(\theta) \) as \( \theta \) varies through \( \mathcal{K} \), and let \( \tau = u/\epsilon \). Since \( \log \tilde{Z} \) is strictly convex and analytic on \( B_- \), we can take \( \epsilon > 0 \). Also, let
\[ \epsilon(a) = a^{-d} \sup \left\{ \left| \frac{\hat{Z}(\theta + it)}{Z(\theta)} \right| : \theta \in \mathcal{K}, a \leq |t| < \frac{1}{a} \right\}. \]

Fix \( \theta \in \mathcal{K} \). Let \( \nu \) be the probability measure given by
\[ \nu(A) = \int 1_A(x) e^{-\theta^\top x} \, Z(dx) \]
with mean \( m(\theta) \) and covariance matrix \( \Sigma(\theta) \). Let \( \hat{\nu}(\lambda) \) be the Laplace transform of \( \nu \), i.e. \( \hat{\nu}(\lambda) = \tilde{Z}(\theta + \lambda)/\tilde{Z}(\theta) \). Let \( \mathcal{N} \) be the normal distribution with mean \( m \) and covariance matrix \( \Sigma \). The idea is to use \( \mathcal{N} \) to approximate \( \nu \) and, hence, \( Z \), through Lemma 10, in a way that is suitably uniform. Note that \( \nu \) and \( \mathcal{N} \) both depend on \( \theta \); we use ‘\( \nu_\theta \)’ and ‘\( \mathcal{N}_\theta \)’ in the next definition to emphasise this. Let
\[ A(Z) = \sup \left\{ \int_C f(x)(\nu_\theta - \mathcal{N}_\theta)(dx + u) : u \in \mathbb{R}^d, f \in \mathcal{G}'(G), C \in \mathcal{E}, \theta \in \mathcal{K} \right\}. \]

Obviously, by bounding \( A(Z) \) we allow \( f \) to be shifted arbitrarily, \( f \) to vary within \( \mathcal{G}'(G) \), and integration over an arbitrary convex set.

**Theorem 9.** For \( \delta \leq \min\{ \tau, e^{-\delta^2/4} \} \), \( \epsilon \leq \min\{ 1, \xi/\delta, \xi \delta \} \), and a constant \( \Delta' \) depending only on the dimension, we have
\[ A(Z) \leq G \Delta' \left[ \frac{\tau}{e^{(d+1)/2}} + \epsilon(\epsilon/\xi) + \frac{e^{-\delta^2\epsilon^2/4} + \epsilon}{e^{d/2}} \right]. \]
Proof. Let $q$ be a bound on the density of $\mathfrak{M}$ and let $b_\epsilon$ be a bound on the modulus of the density of $(\nu - \mathfrak{M}) \ast \Gamma_\epsilon$. Note that $q \leq (2\pi c)^{-d/2}$, which is independent of $\vartheta$.

For $|\lambda| \leq \tau$, let

$$
\psi(\lambda) = \log \hat{\nu}(\lambda) - m^\top \lambda + \frac{1}{2} \lambda^\top \Sigma \lambda
$$

$$
= \log \hat{Z}(\vartheta + \lambda) - \log \hat{Z}(\vartheta) - m^\top \lambda + \frac{1}{2} \lambda^\top \Sigma \lambda.
$$

Then, by arrangement, $\psi$ is analytic on $|\lambda| \leq \tau$, with $\psi(0) = 0$ and all its first and second derivatives vanishing at $0$.

Using Taylor’s theorem and the analyticity of $\log \hat{Z}$ on $B$, for $|\lambda| \leq \tau$ we have

$$
|\psi(\lambda)| \leq \frac{d^3}{3!} u |\lambda|^3 \leq e^d u |\lambda|^3
$$

and, so, for $|\lambda| \leq \delta$,

$$
|\psi(\lambda)| \leq e^d u |\lambda|^3 \leq \left(4e^d u \frac{\epsilon}{\delta}\right) \frac{\xi}{4} |\lambda|^2 \leq \frac{\xi}{4} |\lambda|^2.
$$

Using these two inequalities gives

$$
|\exp(\psi(\lambda)) - 1| \leq e^d u |\lambda|^3 \exp\left(\frac{\xi}{4} |\lambda|^2\right) \text{ for } |\lambda| \leq \delta.
$$

The key point is that the right-hand side here does not depend on $\vartheta$. Using this bound, we have

$$
\int_{|t| < \delta} |\hat{\nu}(it) - e^{im^\top t} e^{-t^\top \Sigma t/2}| \, dt \leq \int_{|t| < \delta} e^{-t^\top \Sigma t/2} |e^{\psi(it)} - 1| \, dt
$$

$$
\leq e^d u \int_{|t| < \delta} |t|^3 e^{t^2/4} e^{-t^\top \Sigma t/2} \, dt
$$

$$
\leq e^d u \int_{|t| < \delta} |t|^3 e^{-t^2/4} \, dt
$$

$$
\leq e^d u \frac{1}{\epsilon^{(d+3)/2}} \int_0^\infty |z|^3 e^{-z^2/4} \, dz
$$

$$
\leq \frac{\Delta_1 \tau}{\epsilon^{(d+1)/2}},
$$

where $\Delta_1$ depends only on the dimension $d$. Thus,

$$
(2\pi)^d b_\epsilon \leq \int_{|t| < \xi/\epsilon} |\hat{\nu}(it) - e^{im^\top t} e^{-t^\top \Sigma t/2}| \, dt
$$

$$
\leq \int_{|t| < \delta} |\hat{\nu}(it) - e^{im^\top t} e^{-t^\top \Sigma t/2}| \, dt + \int_{\delta \leq |t| < \xi/\epsilon} |\hat{\nu}(it)| \, dt + \int_{\xi/\epsilon \leq |t|} e^{-t^\top \Sigma t/2} \, dt
$$

$$
\leq \frac{\Delta_1 \tau}{\epsilon^{(d+1)/2}} + s \left(\frac{\epsilon}{\xi}\right) \frac{\xi}{d} + e^{-\delta^2 \xi/4} \frac{(4\pi)^d/2}{\epsilon^{d/2}}.
$$

Substituting these estimates for $b_\epsilon$ and $q$ into Lemma 10 gives the result.
For the main result, we also need to approximate the normal distribution $N$. To formulate the theorem, let

$$B(v) = \sup \left\{ \left| \int_C f(x) N(dx + m - v) - \int_C f(x) dx \right| : f \in \mathcal{G}(G), C \in \mathcal{C}, \vartheta \in \mathcal{K} \right\}.$$ 

Recall that $N$, $m$, and $\Sigma$ each depends on $\vartheta$.

**Theorem 10.** For a constant $\Delta''$, depending only on the dimension, we have

$$B(v) \leq \Delta'' \frac{1 + |v|^2}{c(d+2)/2} G.$$ 

**Proof.** Temporarily, let $h(x) = f(x)1_C(x)$ and, for a fixed $\vartheta \in \mathcal{K}$, let

$$B_* = \left| \int h(x + v)(\exp(-\frac{1}{2}x^\top \Sigma^{-1}x) - 1) \, dx \right|.$$ 

Then

$$B_* = \left| \int h(x + v)(\exp(-\frac{1}{2}x^\top \Sigma^{-1}x) - 1) \, dx \right| 
\leq \frac{1}{2} \int |f(x + v)||x^\top \Sigma^{-1}x| \, dx 
\leq \frac{1}{2c} \int |f(x)||x - v|^2 \, dx 
\leq \frac{1}{2c} \int g(|x|)|x|^2 \, dx + \frac{|v|^2}{2c} \int g(|x|) \, dx 
\leq \frac{1 + |v|^2}{2c} \mathcal{G}G$$

and, so,

$$B_* \leq \frac{B_*}{\sqrt{2\pi)^d det[\Sigma]}} \leq \frac{\mathcal{G}}{2\sqrt{2\pi)^d c^d}} \frac{1 + |v|^2}{c^{(d+2)/2}} G,$$

as required.

It is worth, very briefly, relating these results to those in Chaganty and Sethuraman (1993). Temporarily following the notation used there, let $T_n$ be a univariate random variable with moment-generating function $\exp(n\psi_n(z))$. Suppose that $\psi_n$ is analytic, bounded in $n$ and $z$ on $\Omega$, and there has a second derivative bounded below by $\alpha$. Let $\{\tau_n\}$ be a positive, bounded sequence inside $\Omega$ and suppose that, for any $a > 0$,

$$\epsilon_n(a) = \sup \left\{ \frac{\exp(n\psi_n(\tau_n + it))}{\exp(n\psi_n(\tau_n))} : a \leq |t| < \frac{1}{a} \right\} = o(a^{-1/2}).$$

These are the conditions in Chaganty and Sethuraman (1993, Theorem 3.3). Under these conditions we can take $\epsilon = na$ and $u = nU$, and $\vartheta$ is bounded.

Take $f$ to be a function of $x \in \mathbb{R}$ with a bounded derivative and with $f(x) = 0$ for $x \in (-\infty, -1]$ and $f(x) = e^{-\tau x}$ for $x \in [0, \infty)$. For $\tau \in (0, B]$, it is routine calculus to show that there exists a constant $C$ such that $f$ is in both $\mathcal{G}'(C/\tau)$ and $\mathcal{G}(C/\tau^2)$. By applying

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
Theorems 9 and 10 with this $f$, the convex set $C$ being $(0, \infty)$ and $K$ being the point set $\{\tau_n\}$ when estimating the $n$th distribution, we obtain

$$
P((T_n - n\psi_n'(\tau_n)) \in (0, \infty)) \approx \frac{1}{\tau_n \sqrt{2\pi n\psi_n''(\tau_n)}} = o\left(\frac{1}{\tau_n \sqrt{n}}\right),$$

which contains the conclusion of Chaganty and Sethuraman (1993, Theorem 3.3). Clearly, a multivariate version of this result would also follow directly from the discussion here, as would a result uniform in $\tau$ in compact subsets of $(0, \infty) \cap \Omega_-$. Also, when sums of independent, identically distributed variables are considered, when $\psi_n(z)$ does not depend on $n$, the result of Stone (1967, Theorem 3) is easily derived (for the nonlattice case).

### 6. Proof of Theorem 7 and Corollary 2

**Proof of Theorem 7.** Note first that $P(W_i(\theta) = 0)$ is a fixed point of the multivariate generating function of the underlying Galton–Watson process. Hence, when the martingale converges in mean, so that $E W_i(\theta) = 1$, these probabilities must be less than 1 and, so, must equal the extinction probabilities from that starting type. Thus, $W_i(\theta) > 0$ agrees with the survival set almost surely for $\theta \in \Lambda_-$. The continuity of $W_i(\theta)$ now means that the null set can be chosen independently of $\theta$. For the rest of the calculation, we deal with sample paths in $\delta$. Then there exists an $N$ such that $W^n_i(\theta) > 0$ for all $n \geq N$.

We take $n \geq N$. Let $K$ be a compact subset of $\Lambda_-$ and fix $j$. Let the function $w^n$ be defined by

$$w^n(\lambda) = \log\left(\frac{W^n_i(\lambda)v_j(\lambda)}{v_j(\lambda)}\right).$$

Using Theorems 1(ii), 3, and 5, there exists an open $B$ (with $\Lambda_- \subset B \subset \Lambda$) such that, for each $n$, $w^n$ is analytic in $\lambda \in B$, $\{w^n\}$ converges as $n \to \infty$ to an analytic function on $B$, and, for some $\tau$,

$$K_\tau = \{\theta + \lambda \in \mathbb{C}^d : \theta \in K, |\lambda| \leq \tau\} \subset B.$$  

Analyticity of $w^n$ and its limit on $B$ mean that all its derivatives are uniformly bounded on $K_\tau$.

To make the connection with the previous section, for fixed $j$ and $n$ let the measure $Z$ be $Z^n_{ij}$ and, for fixed $\theta$, let

$$v(dx) = \frac{e^{-\theta^\top x}Z(dx)}{\int e^{-\theta^\top y}Z(dy)}.$$  

Note that, with $f(x) = e^{\theta^\top x}h(x)$ and $v_1 \equiv v_1(\theta)$,

$$e^{\theta^\top x} \int_C h(x)Z^n_{ij}(dx + n\nu_1(\theta)) = \frac{e^{-\theta^\top \nu_1(\theta)}}{v(\theta)^n} \int_C f(x)e^{-\theta^\top x}Z^n_{ij}(dx + n\nu_1(\theta))$$

$$= \frac{\int e^{-\theta^\top x}Z^n_{ij}(dx)}{\rho(\theta)^n} \int_C f(x)v(dx + n\nu_1)$$

$$= \frac{v_j(\theta)W^n_{ij}(\theta)}{v_j(\theta)} \int_C f(x)v(dx + n\nu_1).$$

Now, by Theorem 5, $v_j(\theta)W^n_{ij}(\theta)/v_j(\theta)$ converges to $v_1(\theta)uj_j(\theta)W_i(\theta)$ uniformly in $\theta \in K$. By assumption, $f \in \mathcal{F}(G)$, and, so, the results of the previous section can be applied to
consider the convergence of \( \int_C f(x) v(dx + n u_1) \), once we show that \( Z \) satisfies the appropriate conditions.

For \( \lambda \in B \), substitution shows that

\[
\log(\tilde{Z}(\lambda)) - n \log(\rho(\lambda)) = w^n(\lambda).
\]

It follows from the uniform boundedness of the derivatives of \( w^n \) that \( \{m - nu_1 : n\} \) and \( \{\Sigma - nu_2 : n\} \) are bounded uniformly in \( \theta \in K \). Furthermore, there are constants \( l > 0 \) and \( L < \infty \) such that, for all sufficiently large \( n \), we have \( v/n \leq 1 \) and \( u/n \leq L \) and, thus, \( r \leq L/l \).

For any \( \epsilon > 0 \),

\[
\epsilon(\epsilon) = e^{-d} \sup_{\theta \in K} \left\{ \frac{\tilde{Z}(\theta + i\eta)}{Z(\theta)} : \epsilon \leq |\eta| < \frac{1}{\epsilon} \right\}
\]

\[
= e^{-d} \sup_{\theta \in K} \left\{ \frac{v_j(\theta)}{v_i(\theta) W_{ij}(\theta)} \left| \frac{M^n_{ij}(\theta + i\eta)}{\rho(\theta)^n} \right| : \epsilon \leq |\eta| < \frac{1}{\epsilon} \right\}
\]

\[
\to 0
\]

geometrically quickly as \( n \to \infty \), using Theorems 5 and 6. Hence, using Theorem 9, \( n^{d/2} A(Z^n) \to 0 \). Furthermore, since \( \{m - nu_1 : n\} \) is uniformly bounded, Theorem 10 gives \( n^{d+2}/B(m - nu_1) \to 0 \). Hence,

\[
n^{d/2} \int_C f(x)v(dx + nu_1) - \frac{1}{\sqrt{(2\pi)^d \det[u_2]}} \int_C f(x) dx \to 0
\]

uniformly in \( \theta \in K \) and \( C \in \mathbb{C} \).

**Proof of Corollary 2.** In Theorem 7, take \( f \in \mathcal{D} \) to be 1 on \( |x| \leq b \) and 0 on \( |x| > b + 1 \), with bounded derivatives. The corollary then follows.

### 7. Extended Perron–Frobenius

A matrix \( M = \{m_{ij}\}_{p \times p} \) has \( n \)th power \( M^n \) with entries denoted by \( m^n_{ij} \). The eigenvalues of \( M \) are the zeros of the characteristic polynomial \( q(z) = \det[zI - M] \), which is of degree \( p \). Denote the roots of \( q(z) \) by \( \rho_1, \ldots, \rho_p \), with the roots listed in order so that

\[
|\rho_1| \geq |\rho_2| \geq \cdots \geq |\rho_p|.
\]

The eigenvalue \( \rho_1 \) of \( M \) is the maximum-modulus eigenvalue if \( |\rho_1| > |\rho_2| \), and is then denoted by \( \rho \). Let the entries \( m_{ij}(\lambda) \) of the matrix \( M(\lambda) = \{m_{ij}(\lambda)\}_{p \times p} \) be functions of \( \lambda \in L \). Clearly, the eigenvalues and their multiplicities all depend on \( \lambda \).

**Proof of Theorem 1.** Parts (i) and (ii) of Theorem 1 arise from routine applications of the implicit function theorem; details can be found in Biggins and Rahimzadeh Sani (2004). Similar results have been obtained before, particularly for matrices of Laplace transforms; see, for example, Miller (1961, Theorem 1(a)), Lancaster and Tismenetsky (1985, Theorem 11.5.1), Ney and Numelin (1987, Theorem 4.1), and Kontoyiannis and Meyn (2003, Proposition 4.8(iii)). Now, by part (i), there is an open set containing \( \tilde{L} \) on which \( \rho(\lambda) = |\rho_1(\lambda)| > |\rho_2(\lambda)| \), where \( \rho = \rho_1 \) (the eigenvalues here depend on \( \lambda \) but this is left implicit in the notation). For the rest of the proof, \( \lambda \) will be confined to this open set.

https://doi.org/10.1239/aap/1127483742 Published online by Cambridge University Press
The resolvent of $M(\lambda)$ is defined by

$$R(z) = \{r_{ij}(z)\}_{p \times p} = (I - zM(\lambda))^{-1},$$

which, for all $i$ and $j$, has the expansion $r_{ij}(z) = \sum_{n=0}^{\infty} z^n m^n_{ij}(\lambda)$ when $z < \|M(\lambda)\|^{-1}$, where $\| \cdot \|$ is a matrix norm (see Lancaster and Tismenetsky (1985, Theorem 11.1.1)). For a fixed $\lambda \in \tilde{\mathbb{L}}$, let $B = B(\tilde{\lambda}, \tilde{\delta})$ and $M_1$ be such that, for all $\lambda \in B$, $\|M(\lambda)\| < M_1$. Now we take $|z| < 1/M_1$ and $\lambda \in B$, and suppress $\lambda$ in the notation.

Let

$$h(z) = \det[I - z\lambda] = (1 - z\rho) \prod_{k=2}^{p}(1 - z\rho_k).$$

Inverting $I - z\lambda$ and rewriting using partial fractions, we obtain

$$r_{ij}(z) = \frac{d_{ij}(z)}{(1 - z\rho) \prod_{k=2}^{p}(1 - z\rho_k)} = \frac{a_{ij}}{1 - z\rho} + \frac{\sum_{k=0}^{p-2} b_{k,ij} z^k}{\prod_{k=2}^{p}(1 - z\rho_k)},$$

where $d_{ij}(z)$ are polynomials in $z$ of degree at most $(p - 1)$, given by the appropriate entry in the adjoint matrix of $I - z\lambda$, and $a_{ij}$ and $b_{k,ij}$ are defined in such a way that (25) holds. For fixed $\lambda$, $|b_{k,ij}/\rho^k|$ are bounded over $i$, $j$, and $k$ (with $k \leq p - 2$) by $C$, say. Then, expansion of (25) gives

$$m^n_{ij} = a_{ij} \rho^n + \sum_{h=0}^{p-2} b_{h,ij} \sum_{k_2+\cdots+k_p=n-h} \rho_2^{k_2} \cdots \rho_p^{k_p} z^n$$

and so,

$$m^n_{ij} = \sum_{h=0}^{p-2} b_{h,ij} \sum_{k_2+\cdots+k_p=n-h} \rho_2^{k_2} \cdots \rho_p^{k_p}.\tag{26}$$

Hence, for $\alpha \in \{|\rho_2|/|\rho|, 1\}$, we have

$$|\rho^{-n} m^n_{ij} - a_{ij}| \leq C(p - 1)(n + 1)^{p - 1} \omega^{n-p}.\tag{27}$$

Let $A$ be the matrix with entries $[a_{ij}]$. Then $\rho^{-n} M^n \to A$ and, so, $M A = \rho A$. Parts (i) and (ii) of the theorem now give $A = cuv^T$ for some scalar $c$; however,

$$v = \rho^{-n} M^n v \to A v = \rho v,$$

using $u^T v = 1$, and, so, $c = 1$.

To translate (26) with $a_{ij} = v_i u_j$ into an asymptotic estimate of $M(\lambda)^n$, the boundedness of $b_{k,ij}$ as $\lambda$ varies is now needed. Let $c_1(B)$ denote the continuously differentiable functions on $B$. The function $h(z)$ is a polynomial in $z$ with coefficients in $c_1(B)$ and $1/\rho$ is a simple root of this polynomial. Hence, $h_1(z) = h(z)/(1 - \rho z)$ is a polynomial in $z$ with coefficients in $c_1(B)$; the same is true of $d_{ij}(z)$. From (25) and $A = v u^T$, we have

$$d_{ij}(z) = v_i u_j h_1(z) + \sum_{k=0}^{p-2} b_{k,ij} z^k (1 - z\rho).$$
Equating powers of $z$ here shows that the $b_{k,ij}$ are in $c_1(B)$. Hence, the supremum of $|b_{k,ij}(\lambda)/\rho(\lambda)^j|$ over $i$, $j$, $k$, and $\lambda$ (with $k \leq p - 2$) in the closed ball of radius $\frac{1}{2}\delta$ centred at $\lambda$ will be finite; denote this finite supremum by $C$.

Let $3\epsilon$ be less than $|\rho(\tilde{\lambda})| - |\rho_2(\tilde{\lambda})|$ and small enough that balls of radius $3\epsilon$ centred on the distinct eigenvalues of $M(\lambda)$ are disjoint. Then any point within a distance $\epsilon$ of one of $\rho_2(\tilde{\lambda}), \ldots, \rho_p(\tilde{\lambda})$ is smaller in magnitude than every point in the $\epsilon$-ball centred on $\rho(\lambda)$. Let $\delta < \frac{1}{4}\delta$ be small enough to ensure that the maximum distance between the eigenvalues at $\lambda$ and those at $\lambda \in B(\tilde{\lambda}, \delta)$, after a suitable pairing, is less than $\epsilon$. Then, for $\lambda \in B(\tilde{\lambda}, \delta)$ and $j = 2, 3, \ldots, p$,

$$\frac{|\rho_j(\lambda)|}{\rho(\lambda)} \leq \frac{|\rho_j(\tilde{\lambda})| + \epsilon}{\rho(\tilde{\lambda}) - \epsilon} \leq \frac{\rho(\tilde{\lambda}) - 2\epsilon}{\rho(\tilde{\lambda}) - \epsilon} = \alpha < 1.$$ 

With these definitions, the estimate (27) holds uniformly on $B(\tilde{\lambda}, \delta)$, and this implies (2).

Although we do not need it here, it is worth noting the following result, which applies to suitable matrices of Fourier transforms. Its proof requires only obvious modifications of the discussion already given.

**Theorem 11.** Theorem 1 remains true when ‘analytic’ is replaced by ‘$q$-times differentiable ($q \geq 1$)’ and ‘$\mathbb{C}^d$’ is replaced by ‘$\mathbb{R}^d$’.

### Acknowledgements

Part of this work was carried out when ARS was on sabbatical at the University of Sheffield in 2003. ARS would like to thank the University of Sheffield for hosting him and the Mosheh Institute of Mathematics, Teacher Training University of Tehran, for their financial support under grant no. S83-116. Both authors are grateful to Peter Hall for pointing us in the direction of Bhattacharya’s work.

### References


