ASYMPTOTIC NORMALITY IN CRUMP–MODE–JAGERS PROCESSES: THE LATTICE CASE

SVANTE JANSON

Abstract. Consider a supercritical Crump–Mode–Jagers process such that all births are at integer times (the lattice case). We show that under a certain condition on the intensity of the offspring process, the second-order fluctuations of the age distribution are asymptotically normal; the condition is essential and not just a technicality. This extends to populations counted by a random characteristic.

1. Introduction

Consider a Crump–Mode–Jagers branching process, starting with a single individual born at time 0, where an individual has $N \leqslant \infty$ children born at the times when the parent has age $\xi_1 \leqslant \xi_2 \leqslant \dots$. Here N and $(\xi_i)_i$ are random, and different individuals have independent copies of these random variables. Technically, it is convenient to regard $\{\xi_i\}_1^N$ as a point process Ξ on $[0, \infty)$, and give each individual x an independent copy Ξ_x of Ξ. For further details, see e.g. Jagers [5].

We consider the supercritical case, when the population grows to infinity (at least with positive probability). As is well-known, under weak assumptions, the population grows exponentially, like $e^{\alpha t}$ for some constant $\alpha > 0$ known as the *Malthusian parameter*, see e.g. [5, Theorems (6.3.3) and (6.8.1)]; in particular, the population size properly normalized converges to some positive random variable, and the age distribution stabilizes. Our purpose is to study the second-order fluctuations of the age distribution, or more generally, of the population counted with a random characteristic.

We consider in this paper the lattice case; we thus assume that the ξ_i are integer-valued and thus all births occur at integer times a.s., but there is no $d > 1$ such that all birth times a.s. are divisible by d.

Our setting can, for example, be considered as a model for the (female) population of some animal that is fertile several years and gets one or several children once every year, with the numbers of children different years random and dependent.

Date: 15 November, 2017.

2010 *Mathematics Subject Classification.* 60J80; 60F05.

Key words and phrases. Crump–Mode–Jagers processes; age distribution.

Partly supported by the Knut and Alice Wallenberg Foundation.

Our main result (Theorem 2.1) shows that under the condition (A7) below on the intensity measure $\mathbb{E}\,\Xi$ of the offspring process, fluctuations are asymptotically normal, and with only a short-range dependence between different times. In a companion paper [8], we show that if (A7) does not hold, then the fluctuations behave differently.

Similar results are proved for multi-type Markov branching processes by Asmussen and Hering [1, Section VIII.3]. Their setting includes the single-type non-Markov case studied here, by taking the type of an individual to be its entire life history until present. However, the assumptions of [1] will in general not be satisfied by our processes.

Remark 1.1. Our setup includes the Galton–Watson case, where all births occur when the mother has age 1 (Example 2.3), but this case is much simpler than the general case and can be treated by simpler methods; see Jagers [5, Section 2.10], where results closely related to the ones below are given.

2. Assumptions and main result

Let $\mu := \mathbb{E}\,\Xi$ be the intensity measure of the offspring process; thus $\mu := \sum_{k=0}^{\infty} \mu_k \delta_k$, where μ_k is the expected number of children that an individual bears at age k (and δ_k is the Dirac delta, i.e., a point mass at k). Let $N_k := \Xi\{k\}$ be the number of children born to an individual at age k. Thus $N = \sum_{k=1}^{\infty} N_k$ and $\mu_k = \mathbb{E}\,N_k$.

We make the following standing assumptions, valid throughout the paper. The first assumption (supercriticality) is essential; otherwise there is no asymptotic behaviour to analyse. The assumptions (A2)–(A4) are simplifying and convenient but presumably not essential. ((A4) can be eliminated by using Theorem 6.1 to count only the living.)

(A1) The process is supercritical, i.e., $\mu([0,\infty]) = \sum_{k=0}^{\infty} \mu_i = \mathbb{E}\,N > 1$.
(A2) No children are born instantaneously, i.e., $\mu_0 = 0$.
(A3) $N \geqslant 1$ a.s. Thus the process a.s. survives.
(A4) There are no deaths.

Define, for all complex z such that either $z \geqslant 0$ or the sums or expectations below converge absolutely,

$$\widehat{\mu}(z) := \sum_{k=0}^{\infty} \mu_k z^k = \sum_{k=0}^{\infty} \mathbb{E}[N_k] z^k = \mathbb{E} \sum_{i=1}^{N} z^{\xi_i} \tag{2.1}$$

and the complex-valued random variable

$$\widehat{\Xi}(z) := \int_0^\infty z^x \,\mathrm{d}\Xi(x) = \sum_{i=1}^{N} z^{\xi_i} = \sum_{k=0}^{\infty} N_k z^k. \tag{2.2}$$

Thus $\widehat{\mu}(z) = \mathbb{E}\,\widehat{\Xi}(z)$.

We make two other standing assumptions:

(A5) $\widehat{\mu}\big(m^{-1}\big) = 1$ for some $m > 1$.

Thus $\alpha := \log m$ satisfies $\sum_{k=1}^{\infty} \mu_k e^{-k\alpha} = \widehat{\mu}(e^{-\alpha}) = 1$, so α is the Malthusian parameter, and the population grows roughly with a factor $e^{\alpha} = m$ for each generation.

(A6) $\mathbb{E}[\widehat{\Xi}(r)^2] < \infty$ for some $r > m^{-1/2}$.

We fix in the sequel some $r > m^{-1/2}$ satisfying (A6). We assume for convenience $r \leqslant 1$. Note that (A6) implies

$$\widehat{\mu}(r) = \mathbb{E}\,\widehat{\Xi}(r) < \infty. \tag{2.3}$$

Hence $\widehat{\mu}(z)$ and $\widehat{\Xi}(z)$ are defined, and analytic, at least for $|z| \leqslant r$. Since $\widehat{\mu}(z)$ is a strictly increasing function on $[0, \infty)$, m^{-1} in (A5) is the unique positive root of $\widehat{\mu}(z) = 1$. However, $\widehat{\mu}(z) = 1$ may have other complex roots. The crucial condition in the present paper is:

(A7) $\widehat{\mu}(z) \neq 1$ for every complex $z \neq m^{-1}$ with $|z| \leqslant m^{-1/2}$.

Let Z_n be the total number of individuals at time n. We define Z_n for all integers n by letting $Z_n := 0$ for $n < 0$. By assumption, $Z_0 = 1$. It is well-known that the number of individuals Z_n grows asymptotically like m^n as $n \to \infty$. For example, see e.g. [5, Theorem (6.3.3)] (and remember that we here consider the lattice case),

$$\mathbb{E}\,Z_n \sim c_1 m^n, \qquad \text{as } n \to \infty, \tag{2.4}$$

with some $c_1 > 0$. Moreover, since (A6) implies $\mathbb{E}[\widehat{\Xi}(m^{-1})^2] < \infty$,

$$Z_n/m^n \xrightarrow{\text{a.s.}} \mathcal{Z}, \qquad \text{as } n \to \infty, \tag{2.5}$$

for some random variable $\mathcal{Z} > 0$, see e.g. Nerman [9]. In particular, it follows that for any fixed $k \geqslant 1$

$$Z_{n-k}/Z_n \xrightarrow{\text{a.s.}} m^{-k}. \tag{2.6}$$

The number of individuals of age $\geqslant k$ at time n is Z_{n-k}. For large n, we expect this to be roughly $m^{-k} Z_n$, see (2.6), and to study the fluctuations, we define

$$X_{n,k} := Z_{n-k} - m^{-k} Z_n, \qquad k = 0, 1, \dots \tag{2.7}$$

Note that $X_{n,0} = 0$. Our main result (Theorem 2.1) yields asymptotic normality of $X_{n,k}$ when (A7) holds; this is extended to random characteristics in Theorem 6.1. For the case when (A7) does not hold, the asymptotic behaviour is different, see [8].

By the assumption (A6) and (2.2), $\mathbb{E}\,N_k^2 < \infty$ for every $k \geqslant 1$. Define, for $j, k \geqslant 1$,

$$\sigma_{jk} := \operatorname{Cov}(N_j, N_k) \tag{2.8}$$

and, at least for $|z| < r$,

$$\Sigma(z) := \sum_{i,j} \sigma_{ij} z^i \bar{z}^j = \operatorname{Cov}\Bigl(\sum_i N_i z^i, \sum_j N_j \bar{z}^j\Bigr) = \mathbb{E}\bigl|\widehat{\Xi}(z) - \widehat{\mu}(z)\bigr|^2. \tag{2.9}$$

Let, for $R > 0$, ℓ^2_R be the Hilbert space of infinite vectors

$$\ell^2_R := \Big\{(a_k)_{k=0}^\infty : \|(a_k)_0^\infty\|^2_{\ell^2_R} := \sum_{k=0}^\infty R^{2k}|a_k|^2 < \infty\Big\}. \tag{2.10}$$

Then the following holds. The proof is given in Section 5.

Theorem 2.1. *Assume* (A1)–(A7)*. Then, as $n \to \infty$,*

$$X_{n,k}/\sqrt{Z_n} \xrightarrow{\mathrm{d}} \zeta_k, \tag{2.11}$$

jointly for all $k \geqslant 0$, for some jointly normal random variables ζ_k with mean $\zeta_k = 0$ and covariance matrix given by, for any finite sequence $a_0, \dots, a_K$ of real numbers,

$$\operatorname{Var}\Big(\sum_k a_k\zeta_k\Big) = \frac{m-1}{m}\oint_{|z|=m^{-1/2}} \frac{\big|\sum_k a_k z^k - \sum_k a_k m^{-k}\big|^2}{|1-z|^2\,|1-\widehat{\mu}(z)|^2}\Sigma(z)\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}. \tag{2.12}$$

The convergence (2.11) *holds also in the stronger sense that $(Z_n^{-1/2}X_{n,k})_k \xrightarrow{\mathrm{d}} (\zeta_k)_k$ in the Hilbert space ℓ^2_R, for any $R < m^{1/2}$. The limit variables ζ_k are non-degenerate unless Ξ is deterministic, i.e., $N_k = \mu_k$ a.s. for each $k \geqslant 0$.*

Recall that joint convergence of an infinite number of variables means joint convergence of any finite set. (This is convergence in the product space $\mathbb{R}^\infty$, see [2].) Note that trivially $\zeta_0 = 0$ (included for completeness).

Remark 2.2. We consider above $X_{n,k}$ for $k \geqslant 0$, i.e., the age distribution of the population at time n. We can define $X_{n,k}$ by (2.7) also for $k < 0$; this means looking into the future and can be interpreted as predicting the future population. As shown in [8], Theorem 2.1 implies its own extension: (2.11)–(2.12) hold for all $k \in \mathbb{Z}$ (still jointly). This enables us, for example, to obtain (by standard linear algebra) the best linear predictor of Z_{n+1} based on the observed $Z_n, \dots, Z_{n-K}$ for any fixed K.

Example 2.3 (Galton–Watson)**.** The simplest example is a Galton–Watson process, where all children are born in a single litter at age 1 of the parent, so $N_k = 0$ for $k \geqslant 2$. (But all individuals live for ever in our setting. In the traditional setting, only the newborns are counted, i.e., $Z_n - Z_{n-1}$; the results are easily transferred to this version.) Then $N = N_1$, $m = \mu_1$ and $\widehat{\mu}(z) = mz$. Hence Assumption (A7) holds. We assume $\mathbb{E}\,N^2 < \infty$; then (A6) holds for any r; we also assume $N \geqslant 1$ a.s. and $\mathbb{P}(N > 1) > 0$; then (A1)–(A7) hold.

Thus Theorem 2.1 applies. The integral in (2.12) can easily be evaluated, and we obtain, for example, $\operatorname{Var}\big(\zeta_1\big) = \sigma_{11}^2 m^{-3}$. This can, of course, be shown in a much simpler and more straightforward way; see [5, Theorem (2.10.1)], which is essentially equivalent to our Theorem 2.1 in the Galton–Watson case (without assuming (A3)).

Example 2.4. Suppose that all children are born when the mother has age one or two, i.e., $N_k = 0$ for $k > 2$. Then $\widehat{\mu}(z) = \mu_1 z + \mu_2 z^2$, where by assumption $\mu_1 + \mu_2 > 1$ and $\mu_1 > 0$. (A5) yields $m^2 = \mu_1 m + \mu_2$, and thus

$$m = \frac{\mu_1 + \sqrt{\mu_1^2 + 4\mu_2}}{2}. \tag{2.13}$$

The equation $\widehat{\mu}(z) = 1$ has one other root, viz. γ_1 with

$$\gamma_1^{-1} = -\frac{\sqrt{\mu_1^2 + 4\mu_2} - \mu_1}{2}. \tag{2.14}$$

The condition (A7) is thus equivalent to $|\gamma_1| > m^{-1/2}$, or $\gamma_1^{-2} < m$, which after some elementary algebra is equivalent to, for example,

$$u_1^3 + 3u_1u_2 + u_2 - u_2^2 > 0. \tag{2.15}$$

Thus, Theorem 2.1 applies when (2.15) holds. See further [8].

2.1. **More notation.** For a random variable X in a Banach space $\mathcal{B}$, we define $\|X\|_{L^2(\mathcal{B})} := (\mathbb{E}\,\|X\|_{\mathcal{B}}^2)^{1/2}$, when $\mathcal{B} = \mathbb{R}$ or $\mathbb{C}$ abbreviated to $\|X\|_2$.

For infinite vectors $\vec{x} = (x_j)_{j=0}^\infty$ and $\vec{y} = (y_j)_{j=0}^\infty$, let $\langle \vec{x}, \vec{y} \rangle := \sum_{j=0}^\infty x_j y_j$, assuming that the sum converges absolutely.

C denotes different constants that may depend on the distribution of the branching process (i.e., on the distribution of N and (ξ_i)), but not on n and similar parameters; the constant may change from one occurrence to the next.

All unspecified limits are as $n \to \infty$.

3. Preliminaries

Let

$$B_n := Z_n - Z_{n-1} \tag{3.1}$$

be the number of individuals born at time n (with $B_0 = Z_0$). Thus,

$$Z_n = Z_{n-1} + B_n = \sum_{i=0}^n B_i, \qquad n \geqslant 0. \tag{3.2}$$

Let $B_{n,k}$ be the number of individuals born at time $n+k$ by parents that are themselves born at time n, and thus are of age k. Thus, recalling (A2),

$$B_n = \sum_{k=1}^n B_{n-k,k}, \qquad n \geqslant 1. \tag{3.3}$$

Let $\mathcal{F}_n$ be the σ-field generated by the life histories of all individuals born up to time n. (With $\mathcal{F}_{-1}$ trivial.) Then $B_{n,k}$ is $\mathcal{F}_n$-measurable, and B_n is $\mathcal{F}_{n-1}$-measurable by (3.3). Furthermore,

$$\mathbb{E}\big(B_{n,k} \mid \mathcal{F}_{n-1}\big) = \mu_k B_n, \qquad n \geqslant 0. \tag{3.4}$$

For $k \geqslant 1$, let

$$W_{n,k} := B_{n,k} - \mathbb{E}\big(B_{n,k} \mid \mathcal{F}_{n-1}\big) = B_{n,k} - \mu_k B_n. \tag{3.5}$$

(Thus $W_{n,k} = 0$ if $n < 0$.) Then $W_{n,k}$ is $\mathcal{F}_n$-measurable with

$$\mathbb{E}\big(W_{n,k} \mid \mathcal{F}_{n-1}\big) = 0. \tag{3.6}$$

Let further

$$W_n := B_n - \sum_{k=1}^{n} \mu_k B_{n-k} = B_n - \sum_{k=1}^{\infty} \mu_k B_{n-k}. \tag{3.7}$$

Thus $W_0 = B_0 = Z_0$, and for $n \geqslant 1$, by (3.7), (3.3) and (3.5),

$$W_n = \sum_{k=1}^{n} W_{n-k,k}. \tag{3.8}$$

Lemma 3.1. *Assume* (A1)–(A6)*. Then, for all $n \geqslant 1$ and $k \geqslant 1$, $\mathbb{E}[W_{n,k}^2] \leqslant Cr^{-2k} m^n$ and $\mathbb{E}[W_n^2] \leqslant Cm^n$.*

Proof. Recall that N_k is the number of children born at age k of an individual, and that $\mathbb{E}\, N_k = \mu_k$. Furthermore, by (2.2), $\widehat{\Xi}(r) \geqslant N_k r^k$ and thus

$$\operatorname{Var} N_k \leqslant \mathbb{E}\, N_k^2 \leqslant r^{-2k}\, \mathbb{E}[\widehat{\Xi}(r)^2] = Cr^{-2k}. \tag{3.9}$$

Let $n \geqslant 0$ and $k \geqslant 1$. Given $\mathcal{F}_{n-1}$, $B_{n,k}$ is the sum of B_n independent copies of N_k, and thus, see (3.5), (3.4) and (3.9),

$$\mathbb{E}\big(W_{n,k}^2 \mid \mathcal{F}_{n-1}\big) = B_n \operatorname{Var}(N_k) \leqslant Cr^{-2k} B_n. \tag{3.10}$$

Taking the expectation and using (2.4) we find

$$\mathbb{E}[W_{n,k}^2] \leqslant Cr^{-2k}\, \mathbb{E}\, B_n \leqslant Cr^{-2k}\, \mathbb{E}\, Z_n \leqslant Cr^{-2k} m^n, \tag{3.11}$$

as asserted. Consequently $\|W_{n,k}\|_2 \leqslant Cr^{-k} m^{n/2}$ and, by (3.8) and Minkowski's inequality, using $rm^{1/2} > 1$,

$$\|W_n\|_2 \leqslant \sum_{k=1}^{n} \|W_{n-k,k}\|_2 \leqslant Cm^{n/2} \sum_{k=1}^{\infty} (rm^{1/2})^{-k} \leqslant Cm^{n/2}. \tag{3.12}$$

$\square$

For $n \geqslant 0$ and $k \geqslant 1$, by (2.7),

$$\begin{aligned} X_{n+1,k} &= Z_{n+1-k} - m^{-k} Z_{n+1} = X_{n,k-1} + m^{1-k} Z_n - m^{-k} Z_{n+1} \\ &= X_{n,k-1} + m^{-k}(mZ_n - Z_{n+1}). \end{aligned} \tag{3.13}$$

Furthermore, by (3.1) and (2.7), we have, for $k \geqslant 0$,

$$B_{n-k} = Z_{n-k} - Z_{n-k-1} = X_{n,k} - X_{n,k+1} + (m-1) m^{-k-1} Z_n. \tag{3.14}$$

By (3.2), (3.7) and (3.14), recalling that $X_{n,0} = 0$ by (2.7) and $\widehat{\mu}(m^{-1}) = 1$ by (A5), for $n \geqslant 0$,

$$\begin{aligned}
mZ_n - Z_{n+1} &= (m-1)Z_n - B_{n+1} = (m-1)Z_n - \sum_{k=1}^{\infty} \mu_k B_{n+1-k} - W_{n+1} \\
&= (m-1)Z_n - \sum_{k=1}^{\infty} \mu_k \bigl(X_{n,k-1} - X_{n,k} + (m-1)m^{-k} Z_n\bigr) - W_{n+1} \\
&= (m-1)Z_n - \sum_{k=1}^{\infty} \mu_k \bigl(X_{n,k-1} - X_{n,k}\bigr) - (m-1)\widehat{\mu}(m^{-1}) Z_n - W_{n+1} \\
&= \sum_{k=1}^{\infty} \mu_k \bigl(X_{n,k} - X_{n,k-1}\bigr) - W_{n+1}.
\end{aligned} \tag{3.15}$$

Consequently, (3.13) yields, for $n \geqslant 0$ and $k \geqslant 1$,

$$X_{n+1,k} = X_{n,k-1} + m^{-k} \Biggl(\sum_{k=1}^{\infty} \mu_k \bigl(X_{n,k} - X_{n,k-1}\bigr) - W_{n+1} \Biggr). \tag{3.16}$$

We write this in vector form. Let $\vec{X}_n := (X_{n,k})_{k=0}^{\infty}$. Furthermore, let

$$\vec{v} := (0, m^{-1}, m^{-2}, \dots) = \bigl(m^{-k} \mathbf{1}\{k > 0\}\bigr)_{k=0}^{\infty} \tag{3.17}$$

and let

$$\Psi\bigl((y_k)_0^{\infty}\bigr) := \sum_{k=1}^{\infty} \mu_k (y_k - y_{k-1}), \tag{3.18}$$

for vectors $(y_k)_0^{\infty}$ such that the sum converges; finally, let S be the shift operator $S\bigl((y_k)_0^{\infty}\bigr) := (y_{k-1})_0^{\infty}$ with $y_{-1} := 0$.

Then (3.16) can be written, again recalling $X_{n,0} = 0$,

$$\vec{X}_{n+1} = S(\vec{X}_n) + \bigl(\Psi(\vec{X}_n) - W_{n+1}\bigr)\vec{v} = T(\vec{X}_n) - W_{n+1}\vec{v}, \tag{3.19}$$

where T is the linear operator

$$T(\vec{y}) := S(\vec{y}) + \Psi(\vec{y})\vec{v}. \tag{3.20}$$

The recursion (3.19) leads to the following formula.

Lemma 3.2. *For every $n \geqslant 0$,*

$$\vec{X}_n = -\sum_{k=0}^{n} W_{n-k} T^k(\vec{v}). \tag{3.21}$$

Proof. For the initial value $\vec{X}_0$, we have by (2.7) $X_{0,k} = -m^{-k} Z_0$ for $k \geqslant 1$, and thus by (3.17) $\vec{X}_0 = -Z_0 \vec{v} = -W_0 \vec{v}$, recalling that $W_0 = B_0 = Z_0$. This verifies (3.21) for $n = 0$. The general case follows by (3.19) and induction. $\square$

We shall consider T defined in (3.20) as an operator on the complex Hilbert space ℓ^2_R defined in (2.10) for a suitable $R > 0$. Recall that the *spectrum* $\sigma(T)$ of a linear operator in a complex Hilbert (or Banach) space is the set of complex numbers λ such that $\lambda - T$ is *not* invertible; see e.g. [3, Section VII.3].

Lemma 3.3. *Suppose that $1 \leqslant R < m$ and that $\widehat{\mu}(R^{-1}) < \infty$. Then $\vec{v} \in \ell^2_R$, Ψ is a bounded linear functional on ℓ^2_R and T is a bounded linear operator on ℓ^2_R. Furthermore, if $\lambda \in \mathbb{C}$ with $|\lambda| > R$, then $\lambda \in \sigma(T)$ if and only if $\lambda \neq m$ and $\widehat{\mu}(\lambda^{-1}) = 1$.*

Proof. We have, by (3.17) and (2.10),

$$\|\vec{v}\|^2_{\ell^2_R} = \sum_{k=1}^{\infty} R^{2k} m^{-2k} < \infty, \tag{3.22}$$

because $R < m$. Next, it is clear from (2.10) that the shift operator S is bounded on ℓ^2_R (with norm R). Furthermore, by (2.1) and assumption,

$$\sum_{k=1}^{\infty} R^{-2k} \mu_k^2 \leqslant \widehat{\mu}(R^{-1})^2 < \infty \tag{3.23}$$

and it follows by the Cauchy–Schwarz inequality that $\Psi_1\big((a_k)_0^\infty\big) := \sum_{k=1}^\infty \mu_k a_k$ defines a bounded linear functional Ψ_1 on ℓ^2_R. Since Ψ can be written $\Psi = \Psi_1 - \Psi_1 S$, Ψ too is bounded. It now follows from (3.20) that T is a bounded linear operator on ℓ^2_R.

For the final statement we note that the mapping $(a_k)_0^\infty \mapsto \sum_{k=0}^\infty a_k z^k$ is an isometry of ℓ^2_R onto the Hardy space H^2_R consisting of all analytic functions $f(z)$ in the disc $\{z : |z| < R\}$ such that

$$\|f\|^2_{H^2_R} := \sup_{r<R} \frac{1}{2\pi} \int_0^{2\pi} |f\big(re^{i\theta}\big)|^2 \, d\theta < \infty. \tag{3.24}$$

(See e.g. [4].) In particular, $\vec{v}$ corresponds to the function

$$v(z) := \sum_{k=1}^{\infty} m^{-k} z^k = \frac{z/m}{1 - z/m} = \frac{z}{m - z}. \tag{3.25}$$

We use the same notations Ψ, S and T for the corresponding linear functional and operators on H^2_R, and note that the shift operator S on ℓ^2_R corresponds to the multiplication operator $Sf(z) = zf(z)$ on H^2_R. The definition (3.20) thus translates to

$$Tf(z) = zf(z) + \Psi(f)v(z). \tag{3.26}$$

Consequently, for any $h \in H^2_R$, the equation $(\lambda - T)f = h$ is equivalent to

$$(\lambda - z)f(z) - \Psi(f)v(z) = h(z). \tag{3.27}$$

Any solution to (3.27) has to be of the form

$$f(z) = c\frac{v(z)}{\lambda - z} + \frac{h(z)}{\lambda - z}, \tag{3.28}$$

where

$$c = \Psi(f) = c\Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) + \Psi\Bigl(\frac{h(z)}{\lambda - z}\Bigr). \tag{3.29}$$

Suppose $|\lambda| > R$; then $1/(\lambda - z)$ is a bounded analytic function on the domain $\{|z| < R\}$, so it follows from (3.24) and $v, h \in H^2_R$ that $v(z)/(\lambda - z) \in H^2_R$ and $h(z)/(\lambda - z) \in H^2_R$. If $\Psi\bigl(v(z)/(\lambda - z)\bigr) \neq 1$, then (3.29) has a unique solution c for any $h \in H^2_R$, and thus (3.27) has a unique solution $f \in H^2_R$, given by (3.28). In other words, then $\lambda - T$ is invertible on H^2_R and $\lambda \notin \sigma(T)$. (Continuity of $(\lambda - T)^{-1}$ is automatic, by the closed graph theorem.) Conversely, if $\Psi\bigl(v(z)/(\lambda - z)\bigr) = 1$, then (3.27) has either no solution or infinitely many solutions f for any given $h \in H^2_R$, and thus $\lambda \in \sigma(T)$.

We have shown that for $|\lambda| > R$,

$$\lambda \in \sigma(T) \iff \Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) = 1. \tag{3.30}$$

We analyse the condition in (3.30) further. If $|\lambda| > R$ and $\lambda \neq m$, then, by (3.25),

$$\frac{v(z)}{\lambda - z} = \frac{z}{(\lambda - z)(m - z)} = \frac{1}{m - \lambda}\Bigl(\frac{\lambda}{\lambda - z} - \frac{m}{m - z}\Bigr). \tag{3.31}$$

Furthermore, $\lambda/(\lambda - z) = \sum_{k=0}^{\infty} \lambda^{-k} z^k$ and thus by (3.18) and (2.1),

$$\Psi\Bigl(\frac{\lambda}{\lambda - z}\Bigr) = \sum_{k=1}^{\infty} \mu_k \lambda^{-k}(1 - \lambda) = (1 - \lambda)\widehat{\mu}\bigl(\lambda^{-1}\bigr). \tag{3.32}$$

Hence, (3.31) yields, recalling $\widehat{\mu}(m^{-1}) = 1$ by (A5),

$$\begin{aligned}
\Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) &= \frac{1}{m - \lambda}\left(\Psi\Bigl(\frac{\lambda}{\lambda - z}\Bigr) - \Psi\Bigl(\frac{m}{m - z}\Bigr)\right) \\
&= \frac{1}{m - \lambda}\bigl((1 - \lambda)\widehat{\mu}(\lambda^{-1}) - (1 - m)\widehat{\mu}(m^{-1})\bigr) \\
&= \frac{1}{m - \lambda}\bigl((1 - \lambda)\widehat{\mu}(\lambda^{-1}) + m - 1\bigr).
\end{aligned} \tag{3.33}$$

Consequently, for $|\lambda| > R$ with $\lambda \neq m$, by (3.30) and (3.33),

$$\begin{aligned}
\lambda \in \sigma(T) &\iff \Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) = 1 \\
&\iff (1 - \lambda)\widehat{\mu}(\lambda^{-1}) + m - 1 = m - \lambda \\
&\iff (1 - \lambda)\widehat{\mu}(\lambda^{-1}) = 1 - \lambda \\
&\iff \widehat{\mu}(\lambda^{-1}) = 1.
\end{aligned} \tag{3.34}$$

In the special case $\lambda = m$, we find by continuity, letting $\lambda \to m$ in (3.33),

$$\begin{aligned}
\Psi\Bigl(\frac{v(z)}{m - z}\Bigr) &= \lim_{\lambda \to m} \Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) = -\frac{d}{d\lambda}\bigl((1 - \lambda)\widehat{\mu}(\lambda^{-1})\bigr)\big|_{\lambda = m} \\
&= \widehat{\mu}(m^{-1}) - (m - 1)m^{-2}\widehat{\mu}'(m^{-1}) < \widehat{\mu}(m^{-1}) = 1
\end{aligned} \tag{3.35}$$

since $\widehat{\mu}'(x) > 0$ for $x > 0$. Hence $m \notin \sigma(T)$. $\square$

Remark 3.4. It is easily seen that $\lambda \in \sigma(T)$ for every λ with $|\lambda| \leqslant R$, e.g. by taking $h = v$ in (3.27)–(3.28) and noting that $v(z)/(\lambda - z) \notin H^2_R$. Thus we have a complete description of the spectrum $\sigma(T)$ on ℓ^2_R.

Lemma 3.5. *Suppose that $1 \leqslant R < m$ and that $\widehat{\mu}(R^{-1}) < \infty$. Suppose furthermore that $\widehat{\mu}(z) \neq 1$ for every complex $z \neq m^{-1}$ with $|z| < R^{-1}$. Then, for every $R_1 > R$, there exists $C = C(R_1)$ such that*

$$\|T^n\|_{\ell^2_R} \leqslant C R_1^n, \qquad n \geqslant 0. \tag{3.36}$$

Proof. By Lemma 3.3, T is a bounded linear operator on ℓ^2_R and if $\lambda \in \sigma(T)$ with $|\lambda| > R$, then $\widehat{\mu}(\lambda^{-1}) = 1$ and $\lambda^{-1} \neq m^{-1}$. By assumption, there is no such λ, and thus $\sigma(T) \subseteq \{\lambda : |\lambda| \leqslant R\}$. (Actually, equality holds by Remark 3.4.) In other words, the spectral radius

$$r(T) := \sup_{\lambda \in \sigma(T)} |\lambda| \leqslant R. \tag{3.37}$$

By the spectral radius formula [3, Lemma VII.3.4], $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$ and thus (3.37) implies that, for any $R_1 > R$, $\|T^n\|^{1/n} < R_1$ for large n, which yields (3.36). $\square$

4. A first normal convergence result

Let $\vec{\eta} := (\eta_0, \eta_1, \eta_2, \dots)$, where $(\eta_k)_0^\infty$ are jointly normal random variables with means $\mathbb{E}\,\eta_k = 0$ and covariances

$$\operatorname{Cov}(\eta_j, \eta_k) = \sigma_{jk} = \operatorname{Cov}(N_j, N_k), \tag{4.1}$$

see (2.8). Note that $\eta_0 = 0$ since $N_0 = 0$.

Lemma 4.1. *Assume* (A1)–(A6), *and let $\vec{\eta}^{(k)} = (\eta_j^{(k)})_{j=0}^\infty$, $k = 1, 2, \dots$, be independent copies of the random vector η. Then, as $n \to \infty$,*

$$Z_n^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{d}} (1 - 1/m)^{1/2} m^{-k/2} \eta_j^{(k)}, \tag{4.2}$$

jointly for all (j, k) with $j \geqslant 0$ and $k \geqslant 0$.

Proof. Consider first a fixed $k \geqslant 0$. Given B_{n-k}, the vector $\vec{B}_{n-k} := (B_{n-k,j})_{j=0}^\infty$ is the sum of B_{n-k} independent copies of the random vector $\vec{N}$, and by (3.5), the vector $\vec{W}_{n-k} := (W_{n-k,j})_{j=0}^\infty$ is the sum of B_{n-k} independent copies of the centered random vector $\vec{N} - \mathbb{E}\,\vec{N}$. By (3.1) and (2.6),

$$\frac{B_n}{Z_n} = 1 - \frac{Z_{n-1}}{Z_n} \xrightarrow{\text{a.s.}} 1 - m^{-1} > 0. \tag{4.3}$$

In particular, $B_n \to \infty$ a.s., and thus $B_{n-k} \to \infty$. Consequently, by the central limit theorem for i.i.d. finite-dimensional vector-valued random variables, and the definition of η_j,

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{d}} \eta_j \overset{\mathrm{d}}{=} \eta_j^{(k)}, \tag{4.4}$$

jointly for any finite set of $j \geqslant 0$.

Moreover, by (4.3) and (2.6),

$$B_{n-k}/Z_n \xrightarrow{\text{a.s.}} (1-1/m)m^{-k}, \tag{4.5}$$

and thus (4.2) for a fixed k follows from (4.3) and (4.4).

To extend this to several k, the problem is that $W_{n-k,j}$ for different k are, in general, dependent. (For example, conditioned on Z_{n-1} and B_{n-1}, $W_{n-1,1}$ determines $B_{n-1,1}$ which contributes to B_n, and thus influences $W_{n,j}$.) We therefore approximate $W_{n-k,j}$ as follows.

We may assume that for each k, we have an infinite sequence $(\vec{N}^{(k,i)})_{i\geqslant 1}$ of independent copies of $\vec{N}$, such that $\vec{W}_{n-k}$ is the sum $\sum_{i=1}^{B_{n-k}} \vec{N}^{(k,i)}$ of the first B_{n-k} vectors; furthermore, these sequences for different k are independent.

Fix $J, K \geqslant 1$ and consider only $j \leqslant J$ and $k \leqslant K$. Let, for $0 \leqslant k \leqslant K$,

$$\overline{B}_{n-k} := \lfloor m^{K-k} B_{n-K} \rfloor \tag{4.6}$$

and let

$$\overline{W}_{n-k,j} := \sum_{i=1}^{\overline{B}_{n-k}} \vec{N}_j^{(k,i)}. \tag{4.7}$$

Then by the central limit theorem, exactly as for (4.4),

$$\overline{B}_{n-k}^{-1/2} \overline{W}_{n-k,j} \xrightarrow{\text{d}} \eta_j^{(k)}, \tag{4.8}$$

jointly for all $j \leqslant J$ and $k \leqslant K$; note that now, if we condition on B_{n-K}, the left-hand sides for different k are independent. Furthermore, by (4.3) and (2.6), $\overline{B}_{n-k}/B_{n-k} \xrightarrow{\text{a.s.}} 1$ for every k. Hence (4.8) yields, jointly,

$$B_{n-k}^{-1/2} \overline{W}_{n-k,j} \xrightarrow{\text{d}} \eta_j^{(k)}. \tag{4.9}$$

Moreover, using (4.7),

$$\mathbb{E}\big((\overline{W}_{n-k,j} - W_{n-k,j})^2 \mid B_{n-k}, \overline{B}_{n-k}\big) = |B_{n-k} - \overline{B}_{n-k}| \operatorname{Var} N_j \tag{4.10}$$

and, consequently, for every fixed $j \geqslant 0$, $k \geqslant 0$ and $\varepsilon > 0$,

$$\mathbb{P}\big(|\overline{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2} \mid B_{n-k}, \overline{B}_{n-k}\big) \leqslant |1 - \overline{B}_{n-k}/B_{n-k}| \sigma_{jj} \varepsilon^{-2} \xrightarrow{\text{a.s.}} 0.$$

Taking the expectation, we obtain by dominated convergence that for every j and k, $\mathbb{P}\big(|\overline{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2}\big) \to 0$ for every $\varepsilon > 0$, and thus

$$B_{n-k}^{-1/2} \overline{W}_{n-k,j} - B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\text{p}} 0. \tag{4.11}$$

Combining (4.9) and (4.11) yields

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\text{d}} \eta_j^{(k)}, \tag{4.12}$$

still jointly for all $j \leqslant J$ and $k \leqslant K$. The result follows by this and (4.5), since J and K are arbitrary. $\square$

5. Proof of Theorem 2.1

In this section we assume (A1)–(A7). Note that (2.3) implies that $\widehat{\mu}(z)$ is analytic in the disc $D_r := \{|z| < r\}$, and thus the points z there with $\widehat{\mu}(z) = 1$ form a discrete set. By (A7), they all satisfy $|z| > m^{-1/2}$ except the root $z = m^{-1}$. Hence we may decrease r so that the disc D_r contains no roots of $\widehat{\mu}(z) = 1$ except m^{-1}, and still $r > m^{-1/2}$. Thus, assuming (A1)–(A6), and with $R := 1/r$, (A7) is equivalent to

(A7′) There exists R with $1 \leqslant R < m^{1/2}$ such that $\widehat{\mu}(R^{-1}) < \infty$ and, furthermore, $\widehat{\mu}(z) \neq 1$ for every complex $z \neq m^{-1}$ with $|z| < R^{-1}$.

We fix an R such that (A7′) holds, and (A6) holds with $r = 1/R$. Note that R may be chosen arbitrarily close to $m^{1/2}$. Furthermore, we fix R_1 with $R < R_1 < m^{1/2}$. Then (A7′) and Lemma 3.5 show that (3.36) holds, i.e., $\|T^n\|_{\ell^2_R} = O\big(R_1^n\big)$.

Lemma 5.1. *Assume* (A1)–(A7)*. If* $R < m^{1/2}$*, then*

$$\mathbb{E}\,\|\vec{X}_n\|^2_{\ell^2_R} \leqslant Cm^n \tag{5.1}$$

and thus

$$\mathbb{E}\,X^2_{n,k} \leqslant CR^{-2k}m^n \tag{5.2}$$

for all $n, k \geqslant 0$*.*

Proof. By (3.21), Lemma 3.1, (3.36) and Minkowski's inequality,

$$\begin{aligned}\|\vec{X}_n\|_{L^2(\ell^2_R)} &\leqslant \sum_{k=0}^{n} \|W_{n-k}\|_{L^2} \|T^k(\vec{v})\|_{\ell^2_R} \leqslant C\sum_{k=0}^{n} m^{(n-k)/2}R_1^k \\ &= Cm^{n/2}\sum_{k=0}^{\infty}(R_1/m^{1/2})^k = Cm^{n/2}.\end{aligned} \tag{5.3}$$

This yields (5.1), and (5.2) follows by (2.10). ∎

Define for convenience $W_{n,j}$ also for $n < 0$ by $W_{-1,1} := W_0$ and $W_{n,j} = 0$ for $n \leqslant -1$ and $j \geqslant 1$ with $(n,j) \neq (-1,1)$. Then (3.8) holds also for $n \leqslant 0$, provided the sum is extended to ∞, and (3.21) can be written

$$\vec{X}_n = -\sum_{k=0}^{\infty}\sum_{j=1}^{\infty} W_{n-k-j,j}T^k(\vec{v}). \tag{5.4}$$

For each finite M define also the truncated sum

$$\vec{X}_{n,M} := -\sum_{k=0}^{M}\sum_{j=1}^{M} W_{n-k-j,j}T^k(\vec{v}). \tag{5.5}$$

Lemma 4.1 implies that for any fixed M, as $n \to \infty$,

$$Z_n^{-1/2}\vec{X}_{n,M} \xrightarrow{\mathrm{d}} -\sum_{k=0}^{M}\sum_{j=1}^{M}(1-m^{-1})^{1/2}m^{-(k+j)/2}\eta_j^{(k+j)}T^k(\vec{v}) \tag{5.6}$$

in ℓ^2_R. Furthermore, by (5.4)–(5.5), Minkowski's inequality, Lemma 3.1 and (3.36), regarding $\vec{X}_n$ and $\vec{X}_{n,M}$ as elements of $L^2(\ell^2_R)$, the space of ℓ^2_R-valued random variables with square integrable norm,

$$\begin{aligned}\|\vec{X}_n - \vec{X}_{n,M}\|_{L^2(\ell^2_R)} &\leqslant \sum_{k>M \text{ or } j>M} \|W_{n-k-j,j}\|_{L^2}\|T^k(\vec{v})\|_{\ell^2_R} \\ &\leqslant C \sum_{k>M \text{ or } j>M} r^{-j} m^{(n-k-j)/2} R_1^k \\ &= Cm^{n/2} \sum_{k>M \text{ or } j>M} (R/m^{1/2})^j (R_1/m^{1/2})^k. \end{aligned} \tag{5.7}$$

Since the sum on the right-hand side of (5.7) converges, it tends to 0 as $M \to \infty$, and thus $m^{-n/2}\big(\vec{X}_n - \vec{X}_{n,M}\big) \to 0$ in $L^2(\ell^2_R)$, and thus in probability, uniformly in n. Since $Z_n/m^n \overset{\text{a.s.}}{\longrightarrow} \mathcal{Z} > 0$, see (2.5), $\sup_n m^n/Z_n$ is an a.s. finite random variable; hence also

$$Z_n^{-1/2}\big(\vec{X}_n - \vec{X}_{n,M}\big) = \Big(\frac{m^n}{Z_n}\Big)^{1/2} m^{-n/2}\big(\vec{X}_n - \vec{X}_{n,M}\big) \overset{\text{p}}{\longrightarrow} 0 \tag{5.8}$$

as $M \to \infty$, uniformly in n.

Moreover, the right-hand side of (5.6) converges as $M \to \infty$ in $L^2(\ell^2_R)$, and thus in distribution, since by (3.9)

$$\mathbb{E}[(\eta_j^{(k)})^2] = \operatorname{Var} N_j \leqslant Cr^{-2j} = CR^{2j}, \tag{5.9}$$

and thus, using also (3.36),

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{j=1}^{\infty} m^{-(k+j)/2}\|\eta_j^{(k+j)}T^k(\vec{v})\|_{L^2(\ell^2_R)} &= \sum_{k=0}^{\infty}\sum_{j=1}^{\infty} m^{-(k+j)/2}\|\eta_j^{(k+j)}\|_{L^2}\|T^k(\vec{v})\|_{\ell^2_R} \\ &\leqslant C\sum_{k=0}^{\infty}\sum_{j=1}^{\infty} m^{-(k+j)/2}R^jR_1^k < \infty. \end{aligned} \tag{5.10}$$

It follows, see [2, Theorem 4.2], that (5.6) extends to $M = \infty$, i.e.,

$$Z_n^{-1/2}\vec{X}_n \overset{\text{d}}{\longrightarrow} -\sum_{k=0}^{\infty}\sum_{j=1}^{\infty}(1-m^{-1})^{1/2}m^{-(k+j)/2}\eta_j^{(k+j)}T^k(\vec{v}) \tag{5.11}$$

in ℓ^2_R as $n \to \infty$. The right-hand side is obviously a Gaussian random vector in ℓ^2_R, which we write as $\vec{\zeta} = (\zeta_0, \zeta_1, \dots)$. Then (5.11) yields (2.11).

It remains to calculate the covariances of ζ_k. Let $\vec{a} = (a_0, a_1, \dots)$ be a (real) vector with only finitely many non-zero elements. Then, by (5.11),

$$\sum_{\ell=0}^{\infty} a_\ell\zeta_\ell = \langle \vec{a}, \vec{\zeta}\rangle = -(1-m^{-1})^{1/2}\sum_{k=0}^{\infty}\sum_{j=1}^{\infty} m^{-(k+j)/2}\eta_j^{(k+j)}\langle T^k(\vec{v}), \vec{a}\rangle \tag{5.12}$$

with the sum converging absolutely in L^2 by (5.10).

By the definition of $\eta_j^{(k)}$ in (4.1) and Lemma 4.1,

$$\operatorname{Cov}\Bigl(m^{-k/2}\eta_i^{(k)}, m^{-\ell/2}\eta_j^{(\ell)}\Bigr) = m^{-(k+\ell)/2}\delta_{k,\ell}\sigma_{ij} = \oint_{|w|=m^{-1/2}} \sigma_{ij} w^k \bar{w}^\ell \frac{|\mathrm{d}w|}{2\pi m^{-1/2}}. \tag{5.13}$$

Hence, (5.12) yields

$$\begin{aligned}
&(1-m^{-1})^{-1} \operatorname{Var}\bigl(\langle \vec{a}, \vec{\zeta}\rangle\bigr) \\
&\quad= \sum_{k=0}^\infty \sum_{\ell=0}^\infty \sum_{i=1}^\infty \sum_{j=1}^\infty \langle T^k(\vec{v}), \vec{a}\rangle \langle T^\ell(\vec{v}), \vec{a}\rangle \oint_{|w|=m^{-1/2}} \sigma_{ij} w^{k+i} \bar{w}^{\ell+j} \frac{|\mathrm{d}w|}{2\pi m^{-1/2}} \\
&\quad= \oint_{|w|=m^{-1/2}} \sum_{i=1}^\infty \sum_{j=1}^\infty \sigma_{ij} w^i \bar{w}^j \left| \sum_{k=0}^\infty w^k \langle T^k(\vec{v}), \vec{a}\rangle \right|^2 \frac{|\mathrm{d}w|}{2\pi m^{-1/2}}.
\end{aligned} \tag{5.14}$$

Furthermore, if $|w| = m^{-1/2}$, then $\sum_{k=0}^\infty \|w^k T^k(\vec{v})\|_{\ell^2_R} < \infty$ by (3.36), and thus

$$\sum_{k=0}^\infty w^k T^k(\vec{v}) = (1-wT)^{-1}(\vec{v}). \tag{5.15}$$

Let $\lambda := w^{-1}$, so $|\lambda| = m^{1/2} > R$. We use as in the proof of Lemma 3.3 the standard isometry $\ell^2_R \to H^2_R$, and let $f(z) \in H^2_R$ be the function corresponding to $(1-wT)^{-1}(\vec{v}) = \lambda(\lambda - T)^{-1}(\vec{v})$. Thus, see (3.26)–(3.27),

$$(\lambda - z)f(z) - \Psi(f)v(z) = (\lambda - T)f(z) = \lambda v(z) \tag{5.16}$$

and thus, cf. (3.27)–(3.29),

$$f(z) = b\frac{v(z)}{\lambda - z} \tag{5.17}$$

for a constant b such that $b = \Psi(f) + \lambda$. This yields by (3.33)

$$b - \lambda = \Psi(f) = \frac{b}{m-\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1}) + m - 1\bigr) \tag{5.18}$$

with the solution

$$b = \frac{\lambda(m-\lambda)}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}. \tag{5.19}$$

Hence, using (3.31), for $|z| \leqslant R$,

$$\begin{aligned}
f(z) = b\frac{v(z)}{\lambda - z} &= \frac{\lambda}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\Bigl(\frac{\lambda}{\lambda - z} - \frac{m}{m-z}\Bigr) \\
&= \frac{\lambda}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\sum_{\ell=0}^\infty (\lambda^{-\ell} - m^{-\ell})z^\ell. \\
&= \frac{1}{(w-1)(1-\widehat{\mu}(w))}\sum_{\ell=0}^\infty (w^\ell - m^{-\ell})z^\ell.
\end{aligned} \tag{5.20}$$

Thus, $(1-wT)^{-1}(\vec{v}) = \bigl(((w-1)(1-\widehat{\mu}(w)))^{-1}(w^\ell - m^{-\ell})\bigr)_\ell$ and, using (5.15),

$$\sum_{k=0}^{\infty} w^k \langle T^k(\vec{v}), \vec{a}\rangle = \langle (1-wT)^{-1}(\vec{v}), \vec{a}\rangle = \frac{1}{(w-1)(1-\widehat{\mu}(w))} \sum_{\ell=0}^{\infty} a_\ell (w^\ell - m^{-\ell}). \tag{5.21}$$

Hence (2.12) follows from (5.14).

Finally, by (2.12), the variable ζ_k is degenerate only if $\Sigma(z) = 0$ for every z with $|z| = m^{-1/2}$, and thus, by (2.9), $\widehat{\Xi}(z) = \widehat{\mu}(z)$ a.s. for every such z, which by (2.1)–(2.2) implies $N_k = \mu_k$ a.s. for every k. $\square$

6. Random characteristics

A random characteristic is a random function $\chi(t) : [0,\infty) \to \mathbb{R}$ defined on the same probability space as the prototype offspring process Ξ; we assume that each individual x has an independent copy (Ξ_x, χ_x) of (Ξ, χ), and interpret $\chi_x(t)$ as the characteristic of x at age t. We consider as above the lattice case, and define, denoting the birth time of x by τ_x,

$$Z_n^\chi := \sum_{x:\tau_x \leqslant n} \chi_x(n - \tau_x), \tag{6.1}$$

the total characteristic of all individuals at time n. See further Jagers [5]. We assume:

(A8) There exists $R_2 < m^{1/2}$ such that $\mathbb{E}[\chi(k)^2] \leqslant C R_2^{2k}$ for some $C < \infty$ and all $k \geqslant 0$.

We define

$$\lambda_k^\chi := \mathbb{E}\,\chi(k), \tag{6.2}$$

$$\Lambda^\chi(z) := \sum_{k=0}^{\infty} \lambda_k^\chi z^k, \tag{6.3}$$

$$\lambda^\chi := \bigl(1 - m^{-1}\bigr)\Lambda^\chi\bigl(m^{-1}\bigr) = \sum_{k=0}^{\infty} \bigl(m^{-k} - m^{-k-1}\bigr)\lambda_k^\chi, \tag{6.4}$$

$$\kappa_{j,k} := \operatorname{Cov}\bigl(\chi(j), N_k\bigr). \tag{6.5}$$

Note that (A8) implies

$$|\lambda_k^\chi| = |\,\mathbb{E}\,\chi(k)| \leqslant C R_2^k. \tag{6.6}$$

Hence, the sum in (6.3) converges absolutely at least for $|z| \leqslant m^{-1/2}$.

We extend Theorem 2.1 (which is the deterministic case $\chi(k) = \sum_{j\leqslant k} a_j$).

Theorem 6.1. *Assume* (A1)–(A8)*. Then, as* $n \to \infty$*,*

$$Z_n^{-1/2}\bigl(Z^\chi - \lambda^\chi Z_n\bigr) \xrightarrow{\mathrm{d}} \zeta^\chi, \tag{6.7}$$

for some normal random variable ζ^χ with mean $\zeta^\chi = 0$ and variance

$$\begin{aligned}\operatorname{Var}(\zeta^\chi) = \frac{m-1}{m}\Biggl(&\sum_{k=0}^{\infty} m^{-k}\operatorname{Var}\bigl(\chi(k)\bigr)\\ &- 2\oint_{|z|=m^{-1/2}} \frac{(1-z)\Lambda^\chi(z)-\lambda^\chi}{(z-1)(1-\widehat{\mu}(z))}\sum_{k=0}^{\infty}\sum_{j=1}^{\infty}\kappa_{kj} z^j\bar{z}^k \frac{|\mathrm{d}z|}{2\pi m^{-1/2}}\\ &+\oint_{|z|=m^{-1/2}} \frac{|(1-z)\Lambda^\chi(z)-\lambda^\chi|^2}{|1-z|^2\,|1-\widehat{\mu}(z)|^2}\sum_{i,j}\sigma_{ij} z^i\bar{z}^j\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}.\Biggr)\end{aligned} \tag{6.8}$$

Joint asymptotic normality for several characteristics, with a corresponding formula for asymptotic covariances, follow by the proof, or by the Cramér–Wold device.

Proof. We use results from Section 5, and assume as we may that R is chosen with $R_2 < R < m^{1/2}$. We define

$$V^\chi_{n,k} := \sum_{x:\tau_x=n}\bigl(\chi_x(k)-\lambda^\chi_k\bigr) = \sum_{x:\tau_x=n}\chi_x(k)-\lambda^\chi_k B_n. \tag{6.9}$$

Then, (6.1) implies

$$Z^\chi_n = \sum_{k=0}^{n}\bigl(V^\chi_{n-k,k}+\lambda^\chi_k B_{n-k}\bigr) = \sum_{k=0}^{\infty}\bigl(V^\chi_{n-k,k}+\lambda^\chi_k B_{n-k}\bigr) \tag{6.10}$$

and, recalling (6.4), (3.1) and (2.7),

$$\begin{aligned} Z^\chi_n - \lambda^\chi Z_n &= \sum_{k=0}^{\infty}\Bigl(V^\chi_{n-k,k}+\lambda^\chi_k\bigl(B_{n-k}-(m^{-k}-m^{-k-1})Z_n\bigr)\Bigr)\\ &= \sum_{k=0}^{\infty}\Bigl(V^\chi_{n-k,k}+\lambda^\chi_k\bigl(X_{n,k}-X_{n,k+1}\bigr)\Bigr)\\ &= \sum_{k=0}^{\infty}V^\chi_{n-k,k}+\langle\vec{X}_n,\Delta\vec{\lambda}^\chi\rangle,\end{aligned} \tag{6.11}$$

where $\Delta\vec{\lambda}^\chi$ is the vector $\bigl(\lambda^\chi_k-\lambda^\chi_{k-1}\bigr)_{k=0}^{\infty}$ (with $\lambda^\chi_{-1} := 0$).

Given B_{n-k}, $V^\chi_{n-k,k}$ is the sum of B_{n-k} independent copies of $\chi(k)-\lambda^\chi_k = \chi(k)-\mathbb{E}\,\chi(k)$. Hence, using (A8), (2.4) and $B_{n-k}\leqslant Z_{n-k}$,

$$\mathbb{E}\bigl(V^\chi_{n-k,k}\bigr)^2 = \mathbb{E}\bigl(\mathbb{E}\bigl(V^\chi_{n-k,k}\bigr)^2 \mid B_{n-k}\bigr) = \operatorname{Var}\bigl(\chi(k)\bigr)\,\mathbb{E}\,B_{n-k}\leqslant Cm^{n-k}R_2^{2k} \tag{6.12}$$

and, using (6.6) and Lemma 5.1,

$$\mathbb{E}\bigl(\lambda^\chi_k(X_{n,k}-X_{n,k+1})\bigr)^2\leqslant CR_2^{2k}\bigl(\mathbb{E}\,X^2_{n,k}+\mathbb{E}\,X^2_{n,k+1}\bigr)\leqslant Cm^n(R_2/R)^{2k}. \tag{6.13}$$

Since we assume $R_2 < R < m^{1/2}$, it follows by standard arguments that if we replace χ by the truncated characteristic $\chi_K(k) := \chi(k)\mathbf{1}\{k\leqslant K\}$, then the error $Z_n^{-1/2}\bigl(Z_n^\chi - \lambda^\chi Z_n - (Z_n^{\chi_K} - \lambda^{\chi_K} Z_n)\bigr)$ tends to 0 in probability as $K \to \infty$, uniformly in n, and as a consequence, see [2, Theorem 4.2], it suffices to prove Theorem 6.1 for the truncated characteristic χ_K. Hence we may in the sequel assume (changing notation) that $\chi(k) = 0$ for $k > K$, for some $K < \infty$.

Let $\vec{\vartheta} = (\vartheta_0, \vartheta_1, \dots)$ be a random vector such that $(\vec{\vartheta}, \vec{\eta})$ is jointly normal with mean 0 and covariances given by (4.1) and

$$\operatorname{Cov}(\vartheta_j, \vartheta_k) = \operatorname{Cov}\bigl(\chi(j), \chi(k)\bigr), \tag{6.14}$$

$$\operatorname{Cov}(\vartheta_j, \eta_k) = \kappa_{j,k} := \operatorname{Cov}\bigl(\chi(j), N_k\bigr). \tag{6.15}$$

Let $\bigl(\vec{\vartheta}^{(k)}, \vec{\eta}^{(k)}\bigr)$ be independent copies of $(\vec{\vartheta}, \vec{\eta})$.

The proof of Lemma 4.1 extends to show that (4.2) holds jointly with

$$Z_n^{-1/2} V^\chi_{n-k,k} \xrightarrow{\mathrm{d}} \bigl(1 - m^{-1}\bigr)^{1/2} m^{-k/2} \vartheta_k^{(k)}, \qquad k \geqslant 0. \tag{6.16}$$

Hence, by the proof in Section 5, (5.11) holds jointly with (6.16) for all k. Consequently, by (6.11) (where we only have to sum for $k \leqslant K$),

$$\begin{aligned}\bigl(1 - m^{-1}\bigr)^{-1/2} & Z_n^{-1/2}\bigl(Z_n^\chi - \lambda^\chi Z_n\bigr) \\ & \xrightarrow{\mathrm{d}} \sum_{k=0}^\infty m^{-k/2} \vartheta_k^{(k)} - \sum_{k=0}^\infty \sum_{j=1}^\infty m^{-(k+j)/2} \eta_j^{(k+j)} \langle T^k(\vec{v}), \Delta\vec{\lambda}^\chi\rangle. \end{aligned} \tag{6.17}$$

Write the right-hand side as $A_1 - A_2$, and note that A_1 and A_2 are jointly normal with means 0. It remains to calculate $\operatorname{Var}(A_1 - A_2)$.

$\operatorname{Var}(A_2)$ was calculated in Section 5, see (5.14) and (2.12), which yields the last term in (6.8), using $\sum_k (\lambda_k^\chi - \lambda_{k-1}^\chi) z^k = (1-z)\Lambda^\chi(z)$ and (6.4).

Since the terms in the sum A_1 are independent,

$$\operatorname{Var}(A_1) = \sum_{k=0}^\infty m^{-k} \operatorname{Var}\bigl(\vartheta_k\bigr) = \sum_{k=0}^\infty m^{-k} \operatorname{Var}\bigl(\chi(k)\bigr). \tag{6.18}$$

Finally, using (6.15) and (5.21),

$$\begin{aligned}\operatorname{Cov}(A_1, A_2) &= \sum_{k=0}^\infty \sum_{j=1}^\infty m^{-(k+j)} \kappa_{k+j,j} \langle T^k(\vec{v}), \Delta\vec{\lambda}^\chi\rangle \\ &= \sum_{k=0}^\infty \sum_{j=1}^\infty \langle T^k(\vec{v}), \Delta\vec{\lambda}^\chi\rangle \oint_{|z|=m^{-1/2}} z^{k+j} \sum_{\ell=0}^\infty \bar{z}^\ell \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \langle (1 - zT)^{-1}(\vec{v}), \Delta\vec{\lambda}^\chi\rangle \sum_{\ell=0}^\infty \sum_{j=1}^\infty z^j \bar{z}^\ell \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \frac{(1-z)\Lambda^\chi(z) - (1-m^{-1})\Lambda^\chi(m^{-1})}{(z-1)(1-\widehat{\mu}(z))} \sum_{\ell=0}^\infty \sum_{j=1}^\infty z^j \bar{z}^\ell \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \end{aligned} \tag{6.19}$$

The result (6.8) follows by combining (6.18), (6.19) and (2.12), recalling (6.4). □

Acknowledgement. I thank Peter Jagers and Olle Nerman for helpful comments.

References

[1] Søren Asmussen and Heinrich Hering, *Branching Processes.* Birkhäuser, Boston, 1983.

[2] Patrick Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[3] Nelson Dunford and Jacob T. Schwartz, *Linear Operators. I. General Theory.* Interscience Publishers, New York, 1958.

[4] Peter L. Duren, *Theory of H^p Spaces.* Academic Press, New York, 1970.

[5] Peter Jagers, *Branching Processes with Biological Applications.* John Wiley & Sons, London, 1975.

[6] Svante Janson, *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, UK, 1997.

[7] Svante Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110** (2004), 177–245.

[8] Svante Janson, Fluctuations in Crump–Mode–Jagers processes: the lattice case. In preparation.

[9] Olle Nerman, On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete* **57** (1981), no. 3, 365–395.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`

URL: `http://www2.math.uu.se/~svante/`