The Annals of Probability
1989, Vol. 17, No. 4, 1466–1477

GEOMETRIC GROWTH IN NEAR-SUPERCRITICAL POPULATION SIZE DEPENDENT MULTITYPE GALTON–WATSON PROCESSES

BY FIMA C. KLEBANER

University of Melbourne

We consider a multitype population size dependent branching process in discrete time. A process is considered to be near-supercritical if the mean matrices of offspring distributions approach the mean matrix of a supercritical process as the population size increases. We show that if the convergence of the means to the supercritical mean is fast enough and the second moments of offspring distributions do not grow too fast as the population size increases, then the process grows geometrically fast. Similarly to the classical multitype Galton–Watson process, the process grows at the geometric rate determined by the largest eigenvalue of the limiting matrix in the direction of the corresponding left eigenvector.

Introduction. We consider a population size dependent multitype Galton–Watson branching process with d types. The process is a vector-valued Markov chain with state space N_0^d of all d-dimensional vectors $\mathbf{z}$ with nonnegative integer coordinates. It is defined recursively by

$$Z_{n+1}^j = \sum_{i=1}^{d} \sum_{\nu=1}^{Z_n^i} X_{ij\nu}^{(n)}(\mathbf{Z}_n), \qquad j = 1, \dots, d,$$

$n = 0, 1, \dots$, and $X_{ij\nu}^{(n)}(\mathbf{z})$ is the number of j-type offspring of an i-type parent particle when the process is in the state $\mathbf{z}$ at time n. The process starts at time $n = 0$ with $\mathbf{Z}_0$ particles. Given $\mathbf{Z}_n = \mathbf{z}$ the random variables $(X_{ij\nu}^{(n)}(\mathbf{z}))$ are i.i.d. distributed as $X_{ij1}^{(1)}(\mathbf{z})$. Moreover the vectors $(X_{ij1}^{(1)}(\mathbf{z}))_{j=1,\dots,d}$ are i.i.d. for any $i = 1, \dots, d$; with distributions depending on the state $\mathbf{z}$. Let

$$\mathbf{M}(\mathbf{z}) = \left(EX_{ij1}^{(1)}(\mathbf{z})\right)_{i,\, j=1,\dots,d}$$

and for $i = 1, \dots, d$,

$$\Gamma_i(\mathbf{z}) = \operatorname{Cov}\left(X_{ij1}^{(1)}(\mathbf{z}),\, X_{ik1}^{(1)}(\mathbf{z})\right)_{j,\, k=1,\dots,d}$$

denote the matrix of the means and the matrices of the covariances of offspring distributions when the population size is in the state $\mathbf{z}$; Γ_i gives the covariances between offspring of type j and k of a type-i parent.

Received March 1988; revised September 1988.
AMS 1980 *subject classifications*. Primary 60J80; secondary 60J05.
Key words and phrases. Multitype branching processes, multidimensional Markov chains, growth rates.

We shall also consider

$$\mathbf{V}(\mathbf{z}) = \sum_{i=1}^{d} z_i \Gamma_i(\mathbf{z}),$$

which is the conditional dispersion matrix of $\mathbf{Z}_{n+1}$ given $\mathbf{Z}_n = \mathbf{z}$, as will be seen later. When the offspring distributions do not depend on the state, then the process is the classical multitype branching process with mean matrix $\mathbf{M}$. Such a process is said to be critical or super(sub)critical if the largest in modulus eigenvalue of $\mathbf{M}$ is equal to 1 or greater (smaller) than 1. It is known that in the supercritical case with positively regular matrix $\mathbf{M}, \mathbf{Z}_n r^{-n} \to W\mathbf{v}$ a.s. and in L^2, where r is the largest eigenvalue of $\mathbf{M}$, $\mathbf{v}$, is the corresponding left eigenvector and W is a random variable [see Harris (1963)].

Let us call the process near-supercritical if the mean matrices $\mathbf{M}(\mathbf{z})$ approach the matrix of a supercritical process as $\|\mathbf{z}\| \to \infty$.

It is known that in the single-type near-supercritical processes, geometric growth occurs if

$$\sum |M_n - M|/n < \infty, \qquad \sum V_n/n^2 < \infty.$$

If these conditions hold, then Z_n/M^n converges a.s. and in L^2 to a nondegenerate limit. These conditions are also necessary for geometric growth for a class of processes [see Klebaner (1984), also Cohn and Klebaner (1986)]. The purpose of this paper is to generalize the results of geometric growth found in single-type near-supercritical processes to processes with many types. Our simplifying assumption will be that the moments of offspring distributions can be bounded by quantities that depend on the total population size. It turns out that if the convergence of the mean matrices to the limit is fast enough and the variances do not grow too fast, then the asymptotic behaviour of the process is similar to that in the classical case mentioned above. We obtain the result on the growth of the process $\mathbf{Z}_n$ by looking at projections of $\mathbf{Z}_n$ onto the suitable basis vectors. Choosing for the basis right eigenvectors and generalized eigenvectors of $\mathbf{M}$, we find that the rate of growth of the projection onto the vector corresponding to the largest eigenvalue is r^n, whereas the projections onto other vectors grow at the rate $o(r^n)$. The limit vector has nonrandom direction like that of the left eigenvector of $\mathbf{M}$ corresponding to the largest eigenvalue and random length. We also give necessary conditions for geometric growth from which it can be seen that there is a class of processes for which sufficient conditions for geometric growth are also necessary.

We shall adopt the standard notation for vectors and matrices. We shall agree that if a vector is written from the left, then it is a row vector, whereas the right vectors are to be understood as column vectors. The process $\mathbf{Z}$ is a row vector. $\mathbf{1}$ is a vector consisting of 1's. $\mathbf{u}^*$ denotes the complex conjugate of $\mathbf{u}$. The matrix norm $\|\cdot\|$ is the operator norm in which convergence of matrices is to be understood. Since all the matrix norms are equivalent to each other, it is not important which norm we choose. The L^2 norm of a random variable is denoted by $\|\cdot\|_2$. Constants are denoted by capital letters. C stands for an unspecified positive constant. $\mathbf{C}^d$ stands for the space of all d-dimensional complex vectors. $E_{\mathbf{i}}$ stands for the expectation operator when the process starts in the state $\mathbf{i}$.

Results.

Sufficient conditions for geometric growth.

THEOREM 1. *Let $(\mathbf{Z}_n)$ be a multitype population size dependent branching process, such that it is irreducible and satisfies*

$$\text{(1)} \qquad \mathbf{M}(\mathbf{z}) = \mathbf{M} + \mathbf{D}(\mathbf{z}),$$

where $\mathbf{M}$ is a mean matrix of a supercritical process. Suppose that there exists a sequence $(d(n))$ such that

$$\text{(2)} \qquad \|\mathbf{D}(\mathbf{z})\| \le d(\mathbf{z1}), \quad d(n) \text{ is nonincreasing and } \sum d(n)/n < \infty.$$

We shall assume that $\mathbf{M}$ is positively regular. Let $\mathbf{u}$ be a right eigenvector of $\mathbf{M}$ corresponding to the eigenvalue $r > 1$. Denote $W_n = \mathbf{Z}_n\mathbf{u}r^{-n}$. Then $\lim_{n\to\infty} E_{\mathbf{i}}W_n > 0$ exists for all $\mathbf{i}$. Moreover there exists a r.v. W such that $W_n \to W$ a.s.

REMARKS. 1. Instead of irreducibility it is enough to assume that for any initial value $\mathbf{i} \in N^d$ and C there is n such that $P_{\mathbf{i}}(\mathbf{Z}_n\mathbf{1} > C) > 0$.

2. It is known from the Perron–Frobenius theorem [see, e.g., Seneta (1981)] that a positive regular matrix has a maximal positive eigenvalue r, $r > |l|$, for any eigenvalue $l \ne r$; r is simple and has associated strictly positive right and left eigenvectors. In what follows, $\mathbf{u}$ and $\mathbf{v}$ are the right and left eigenvectors of $\mathbf{M}$ corresponding to r normalized so that $\mathbf{vu} = 1$.

THEOREM 2. *In addition to the conditions of Theorem 1 suppose that there exists a sequence $v^2(n)$ such that*

$$\text{(3)} \qquad \max_{1\le i\le d} \|\Gamma_i(\mathbf{z})\| \le v^2(\mathbf{z1}),$$

$$v^2(n)/n \text{ is nonincreasing and } \sum v^2(n)/n^2 < \infty.$$

Then W_n converges a.s. and in L^2 to a nondegenerate limit W.

THEOREM 3. *Suppose that the conditions of Theorems 1 and 2 hold. Suppose also that*

$$\text{(4)} \qquad \sum v(n)/n^{3/2} < \infty.$$

Then

$$\mathbf{Z}_n r^{-n} \to W\mathbf{v} \quad a.s. \text{ and } L^2.$$

The following theorems give necessary conditions for geometric growth. It is shown that for a class of processes sufficient conditions (2) and (3) are also necessary for geometric growth.

Necessary conditions for geometric growth.

THEOREM 4. *Suppose* $\mathbf{Z}_n r^{-n}$ *converges to a nondegenerate* $\mathbf{w}$ *in* L^1. *Let the components of* $\mathbf{w}$ *be* w_i *and suppose that the entries* $d_{ij}(\mathbf{z})$ *of* $\mathbf{D}(\mathbf{z})$ *have the same sign. Then*:

(i) *For all* $i = 1, \dots, d$, $P(w_i > 0) > 0$.
(ii)

$$\sum_{n=1}^{\infty} \|\mathbf{D}(\mathbf{Z}_n)\| < \infty \quad a.s. \text{ on } \{w_i > 0,\ i = 1, \dots, d\}.$$

(iii) *Suppose further that* $\|\mathbf{D}(\mathbf{z})\|$ *depends on* $\mathbf{z1}$ *and is monotone nonincreasing. Then*

$$\sum_{n=1}^{\infty} \|\mathbf{D}(n)\|/n < \infty.$$

THEOREM 5. *Suppose* $\mathbf{Z}_n r^{-n}$ *converges to a nondegenerate* $\mathbf{w}$ *in* L^2, *and that* $\mathbf{uV(z)u} + 2\mathbf{zuzD(z)u} + (\mathbf{zD(z)u})^2$ *has the same sign for all* $\mathbf{z}$. *Then*:

(i) $$\sum_{n=1}^{\infty} \mathbf{u}\Gamma_i(\mathbf{Z}_n)\mathbf{u}\mathbf{r}^{-n} < \infty \quad a.s. \text{ on } \{w_i > 0\}, \qquad i = 1, \dots, d.$$

(ii) *Suppose also that* $\mathbf{u}\Gamma_i(\mathbf{z})\mathbf{u}$ *depends on* $\mathbf{z1}$ *and* $\mathbf{u}\Gamma_i(n)\mathbf{u}/n$ *is monotone nonincreasing for* $i = 1, \dots, d$. *Then for each* i,

$$\sum_{n=1}^{\infty} \mathbf{u}\Gamma_i(n)\mathbf{u}/n^2 < \infty.$$

Moreover if all the entries of $\Gamma_i(n)$ *are positive, then*

$$\sum_{n=1}^{\infty} \|\Gamma_i(n)\|/n^2 < \infty.$$

Proofs. The following lemmas are the key tools in establishing sufficiency of the conditions for geometric growth.

LEMMA 1. *Let* $f(x)$ *be a positive nonincreasing function,* $0 < f(x) \searrow$. *Then for any* $m > 1$, $c > 0$,

$$\sum_{n=1}^{\infty} f(cm^n) < \infty \quad \text{if and only if} \quad \sum_{n=1}^{\infty} f(n)/n < \infty.$$

LEMMA 2. *Let*

$$0 < f(x) \searrow \quad \text{and} \quad \sum f(n)/n < \infty. \tag{5}$$

Then there exists $F(x)$ *with the following properties.* $F(x) \geq f(x)$, $F(x)$ *is nonincreasing,* $xF(x)$ *is nondecreasing,* $xF(x)$ *is concave on* R^+ *and* $\Sigma F(n)/n < \infty$.

REMARK. In what follows we shall replace (without change in notation) a function f that satisfies (5) by a dominating function constructed in Lemma 2, i.e., f satisfies (5) and in addition $xf(x)$ is nondecreasing and is *concave* on R^+.

LEMMA 3. *Let* $f(x)$ *satisfy* (5), $m > 1$, *and* (a_n) *be a sequence of positive numbers. If*

$$|a_{n+1} - a_n| < a_n f(a_n m^n),$$

then $\lim a_n = a$ *exists. Moreover* $a > 0$ *if* a_0 *is large enough.*

LEMMA 4. *Let* $0 < a_n$ *satisfy*

$$a_{n+1} < ca_n + b_n,$$

where $0 < c < 1$ *and* $0 < b_n$, $\Sigma b_n < \infty$. *Then* $\Sigma a_n < \infty$.

The proof of Lemma 1 follows from the Cauchy condensation principle; it is also given in Klebaner (1984), Lemma 1.

The proof of Lemma 2 is outlined in Klebaner (1985); we give it here in full since it is central for our arguments, also there is a change in notation.

PROOF OF LEMMA 2. Define

$$F(x) = \left(f(1)/x + 1/x \int_1^x f(t)\,dt \right) I(x \geq 1) + I(0 \leq x < 1) f(1).$$

Since f is nonincreasing $f(1) \geq f(t) \geq f(x)$ for $1 < t < x$. Hence $F(x) \geq f(x)$. Take $1 \leq x < y$. Since f is nonincreasing

$$x\int_1^y f(t)\,dt - y\int_1^x f(t)\,dt \leq (y - x)\left(xf(x) - \int_1^x f(t)\,dt \right)$$
$$\leq (y - x) f(x) \leq (y - x) f(1).$$

Dividing through by xy, we obtain F is nonincreasing.

Since $f > 0$, $xF(x)$ is nondecreasing.

$$\int_x^{(x+y)/2} f(t)\,dt \geq \int_{(x+y)/2}^y f(t)\,dt.$$

Hence

$$\int_1^{(x+y)/2} f(t)\,dt \geq \tfrac{1}{2}\int_1^x f(t)\,dt + \tfrac{1}{2}\int_1^y f(t)\,dt.$$

This gives the concavity of $xF(x)$ on $[1, +\infty)$. From the definition one can see that $xF(x)$ is concave on R^+.

By changing the order of integration in the double integral, we have

$$\int_1^\infty F(x)/x\,dx = f(1) + \int_1^\infty f(t)/t\,dt < \infty. \qquad \square$$

The proof of Lemma 3 is given in Klebaner (1984).

PROOF OF LEMMA 4. Iterations give

$$a_n < a_0 c^n + \sum_{i=0}^{n-1} c^{n-i} b_i.$$

The result follows by changing the order of summation in $\sum a_n$. □

PROOF OF THEOREM 1. Consider $E(\mathbf{Z}_{n+1}|\mathbf{Z}_n)$. Using (1) and (2), we have

$$E(\mathbf{Z}_{n+1}|\mathbf{Z}_n) = \mathbf{Z}_n\mathbf{M}(\mathbf{Z}_n) = \mathbf{Z}_n\mathbf{M} + \mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n). \tag{6}$$

Since $\mathbf{u}$ is an eigenvector corresponding to the eigenvalue r,

$$E(W_{n+1}|\mathbf{Z}_n) = W_n + r^{-(n+1)}\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}. \tag{7}$$

Hence

$$|EW_{n+1} - EW_n| \le r^{-(n+1)}|E\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}|. \tag{8}$$

By the Cauchy–Schwarz inequality for any $\mathbf{x}, \mathbf{y}$ and matrix $\mathbf{A}$,

$$|\mathbf{xAy}| \le \|\mathbf{x}\|\,\|\mathbf{A}\|\,\|\mathbf{y}\|,$$

and if $\mathbf{y}$ is fixed, then for any nonnegative $\mathbf{x}$,

$$\mathbf{xAy} \le C\mathbf{x1}\|\mathbf{A}\|. \tag{9}$$

Consider for a random nonnegative $\mathbf{X}$,

$$|E\mathbf{XD}(\mathbf{X})\mathbf{y}| \le E|\mathbf{XD}(\mathbf{X})\mathbf{y}| \le CE\mathbf{X1}\|\mathbf{D}(\mathbf{X})\|.$$

Using this with $\mathbf{X} = \mathbf{Z}_n$ and condition (2), we have for any fixed $\mathbf{y}$,

$$E|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}| \le CE\mathbf{Z}_n\mathbf{1}d(\mathbf{Z}_n\mathbf{1}).$$

Since $d(x)$ satisfies Lemma 2, we have by the concavity of $xd(x)$ for any fixed $\mathbf{y}$,

$$E|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}| \le CE\mathbf{Z}_n\mathbf{1}d(E\mathbf{Z}_n\mathbf{1}). \tag{10}$$

Since the right eigenvector $\mathbf{u}$ of $\mathbf{M}$ corresponding to r has all its coordinates strictly positive, $\mathbf{Z}_n\mathbf{1} < \mathrm{C}\mathbf{Z}_n\mathbf{u}$. Using this with (8) and (10), we obtain

$$|EW_{n+1} - EW_n| < CEW_n d(Cr^n EW_n).$$

Notice that from the irreducibility assumption or the assumption in Remark 1 it follows that $EW_n > 0$ for all n. Hence the conditions of Lemma 3 are fulfilled and there exists $\lim EW_n$. This limit is positive if $\mathbf{Z}_0\mathbf{u}$ is large enough. The rest of the proof consists of showing that $a_\mathbf{i} = \lim_{n\to\infty} E_\mathbf{i}W_n > 0$ is similar to that in the one-dimensional case. As in there [see Klebaner (1984)] by using

Chapman–Kolmogorov equations we obtain

$$a_{\mathbf{i}} = r^{-k} \sum_{\mathbf{j} \in \mathbf{N}^{\mathbf{d}}} P_{\mathbf{ij}}^{(k)} a_{\mathbf{j}},$$

where $P_{\mathbf{ij}}^{(k)}$ are the k-step transition probabilities of the Markov chain $(\mathbf{Z}_n)$. Since $P_{\mathbf{ij}}^{(k)} > 0$ for some k and $\mathbf{j}$ with $\mathbf{j1}$ large enough, we obtain that all $a_{\mathbf{i}}$ are positive. To see that a.s. convergence holds, define a martingale

$$(11) \qquad Y_n = W_n - T_n, \qquad T_n = \sum_{i=0}^{n-1} \left(E(W_{i+1}|F_i) - W_i\right) = \sum_{i=0}^{n-1} \mathbf{Z}_i \mathbf{D}(\mathbf{Z}_i)\mathbf{u} r^{-(i+1)},$$

where $F_n = \sigma(\mathbf{Z}_0, \dots, \mathbf{Z}_n)$. It is seen that Y_n is a martingale by using (7). We have from (10),

$$r^{-n} E|\mathbf{Z}_n \mathbf{D}(\mathbf{Z}_n)\mathbf{u}| < CEW_n d(Cr^n) < Cd(Cr^n).$$

The last bound is a term of a convergent series by (2) and Lemma 1. Hence it follows that $T_n \to T$ a.s. and L^1. Hence $\sup E|Y_n| < \infty$, and by the martingale convergence theorem $Y_n \to Y$ a.s. Hence $W_n \to W = Y + T$ a.s. □

PROOF OF THEOREM 2. The proof follows the line of argument of Klebaner (1984) [see also Cohn and Klebaner (1986), Section 4] and consists of showing that $\lim EW_n^2$ exists and is finite. The rest is obtained through the martingale convergence theorem applied to Y_n.

Consider for a generally complex $\mathbf{y}$,

$$(12) \qquad \mathbf{Z}_{n+1}\mathbf{y} = \sum_{j=1}^{d} y_j Z_{n+1}^j = \sum_{i=1}^{d} \sum_{\nu=1}^{Z_n^i} \sum_{j=1}^{d} y_j X_{ij\nu}^{(n)}(\mathbf{Z}_n) = \sum_{i=1}^{d} \sum_{\nu=1}^{Z_n^i} Y_{i\nu}^{(n)}(\mathbf{Z}_n),$$

with

$$Y_{i\nu}^{(n)}(\mathbf{Z}_n) = \sum_{j=1}^{d} y_j X_{ij\nu}^{(n)}(\mathbf{Z}_n).$$

Due to the conditional independence of $(X_{ij\nu}^{(n)}(\mathbf{Z}_n))$, given $\mathbf{Z}_n = \mathbf{z}$, for a fixed i, $(Y_{i\nu}^{(n)}(\mathbf{z}))$ are i.i.d. random variables with distributions depending on $\mathbf{z}$ and i. Moments of Y's are given by

$$E\left(Y_{i1}^{(n)}(\mathbf{z})\right) = \sum_{j=1}^{d} y_j m_{ij}(\mathbf{z}),$$

$$\sum_{i=1}^{d} z_i E\left(Y_{i1}^{(n)}(\mathbf{z})\right) = \mathbf{zM}(\mathbf{z})\mathbf{y},$$

$$\operatorname{Var}\left(Y_{i1}^{(n)}(\mathbf{z})\right) = \mathbf{y}\Gamma_i(\mathbf{z})\mathbf{y}^*,$$

$$(13) \qquad \sum_{i=1}^{d} z_i \operatorname{Var}\left(Y_{i1}^{(n)}(\mathbf{z})\right) = \mathbf{y} \sum_{i=1}^{d} z_i \Gamma_i(\mathbf{z})\mathbf{y}^* = \mathbf{yV}(\mathbf{z})\mathbf{y}^*.$$

Note that (13) shows that $\mathbf{V}(\mathbf{z})$ is the conditional dispersion matrix of $\mathbf{Z}_{n+1}$ given $\mathbf{Z}_n = \mathbf{z}$ as mentioned in the Introduction. Hence we obtain by using (1) for any $\mathbf{y}$,

$$(14)\qquad \begin{aligned} E\big(|\mathbf{Z}_{n+1}\mathbf{y}|^2|\mathbf{Z}_n = \mathbf{z}\big) &= |\mathbf{zM}(\mathbf{z})\mathbf{y}|^2 + \mathbf{yV}(\mathbf{z})\mathbf{y}^* \\ &= |\mathbf{zMy} + \mathbf{zD}(\mathbf{z})\mathbf{y}|^2 + \mathbf{yV}(\mathbf{z})\mathbf{y}^*. \end{aligned}$$

Taking $\mathbf{y}$ to be the right eigenvector corresponding to r, we have

$$(15)\qquad E\big((\mathbf{Z}_{n+1}\mathbf{u})^2|\mathbf{Z}_n = \mathbf{z}\big) = \big(\mathbf{zu}r + \mathbf{zD}(\mathbf{z})\mathbf{u}\big)^2 + \mathbf{uV}(\mathbf{z})\mathbf{u}^*.$$

Notice that for any $\mathbf{y}$ we have by (9) and the triangle inequality

$$(16)\qquad \mathbf{yV}(\mathbf{z})\mathbf{y}^* < C\|\mathbf{V}(\mathbf{z})\| < C\sum_{i=1}^{d} z_i\|\Gamma_i(\mathbf{z})\| < C(\mathbf{z1})v^2(\mathbf{z1}).$$

From (15) we have by using (9),

$$(17)\qquad \begin{aligned} |EW_{n+1}^2 - EW_n^2| &= r^{-2(n+1)}|E\mathbf{uV}(\mathbf{Z}_n)\mathbf{u}^* + 2E\mathbf{Z}_n\mathbf{u}\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u} + E(\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u})^2| \\ &< Cr^{-2n}E\big(\mathbf{Z}_n\mathbf{1}v^2(\mathbf{Z}_n\mathbf{1}) + (\mathbf{Z}_n\mathbf{1})^2d(\mathbf{Z}_n\mathbf{1}) + (\mathbf{Z}_n\mathbf{1})^2d^2(\mathbf{Z}_n\mathbf{1})\big) \\ &= Cr^{-2n}E(\mathbf{Z}_n\mathbf{1})^2\big[v^2(\mathbf{Z}_n\mathbf{1})/\mathbf{Z}_n\mathbf{1} + d(\mathbf{Z}_n\mathbf{1}) + d^2(\mathbf{Z}_n\mathbf{1})\big]. \end{aligned}$$

Denote by $g(x) = v^2(\sqrt{x})/\sqrt{x} + d(\sqrt{x}) + d^2(\sqrt{x})$. Then it is easily seen by (2) and (3) that g satisfies $\sum g(n)/n < \infty$ and g satisfies Lemma 2. Taking $h(x) = xg(x)$ and using its concavity, we have

$$(18)\qquad \begin{aligned} |EW_{n+1}^2 - EW_n^2| &< Cr^{-2n}Eh\big((\mathbf{Z}_n\mathbf{1})^2\big) < Cr^{-2n}h\big(E(\mathbf{Z}_n\mathbf{1})^2\big) \\ &< CEW_n^2g\big(Cr^{2n}EW_n^2\big). \end{aligned}$$

Now Lemma 3 gives $\lim EW_n^2$ exists and is finite.

By the triangle inequality for Y_n and T_n given by (11),

$$\|Y_n\|_2 \le \|W_n\|_2 + \|T_n\|_2.$$

We show next that $\sup\|T_n\|_2 < \infty$. If $g_1(x) = d^2(\sqrt{x})$, then it is seen that it satisfies Lemma 2, so that $h_1(x) = xg_1(x)$ may be taken to be concave, and we have by (9) for any $\mathbf{y}$,

$$\begin{aligned} E(\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y})^2 &< CE(\mathbf{Z}_n\mathbf{1})^2d(\mathbf{Z}_n\mathbf{1})^2 = CEh_1\big((\mathbf{Z}_n\mathbf{1})^2\big) < Ch_1\big(E(\mathbf{Z}_n\mathbf{1})^2\big) \\ &< Ch_1\big(CE(\mathbf{Z}_n\mathbf{u})^2\big) < Ch_1(Cr^{2n}) < Cr^{2n}d^2(Cr^n). \end{aligned}$$

Hence

$$(19)\qquad \|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}r^{-n}\|_2 < Cd(Cr^n)$$

and by Lemma 1,

$$(20)\qquad \sum_{n=0}^{\infty} \|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}\|_2 r^{-n} < C\sum d(Cr^n) < \infty.$$

Thus $\sup\|T_n\|_2 < \infty$ and T_n converges in L^2 and a.s. Since $\sup\|Y_n\|_2 < \infty$, the martingale convergence theorem yields the L^2 and a.s. convergence of (Y_n).

Hence W_n converges a.s. and in L^2 to W. Moreover $EW = \lim EW_n > 0$, and the proof is completed. □

PROOF OF THEOREM 3. Let $\mathbf{y}$ be an eigenvector or a generalized eigenvector of $\mathbf{M}$ corresponding to an eigenvalue l, $|l| < r$, where r is the largest in modulus eigenvalue. Then we show

$$|\mathbf{Z}_n\mathbf{y}|r^{-n} \to 0 \quad \text{a.s. and } L^2. \tag{21}$$

Let first $\mathbf{y}$ be an eigenvector of $\mathbf{M}$ corresponding to the eigenvalue l. Using (14), we have

$$E|\mathbf{Z}_{n+1}\mathbf{y}|^2 = E|\mathbf{Z}_n\mathbf{y}l + \mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}|^2 + E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^*. \tag{22}$$

Hence

$$\begin{aligned} \|\mathbf{Z}_{n+1}\mathbf{y}\|_2 &\le \|\mathbf{Z}_n\mathbf{y}l + \mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}\|_2 + \left(E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^*\right)^{1/2} \\ &\le \|\mathbf{Z}_n\mathbf{y}\|_2\,|l| + \|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}\|_2 + \left(E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^*\right)^{1/2}. \end{aligned} \tag{23}$$

The function $v^2(\sqrt{x})/\sqrt{x}$ satisfies Lemma 2, so that $h_2(x) = \sqrt{x}\,v^2(\sqrt{x})$ may be taken to be concave, and we have for any $\mathbf{y}$, using (16) and Theorem 2,

$$\begin{aligned} E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^* &< CE\|\mathbf{V}(\mathbf{Z}_n)\| < CE(\mathbf{Z}_n\mathbf{1})v^2(\mathbf{Z}_n\mathbf{1}) = CEh_2\left((\mathbf{Z}_n\mathbf{1})^2\right) \\ &\le Ch_2\left(E(\mathbf{Z}_n\mathbf{1})^2\right) < Ch_2(Cr^{2n}) < Cr^nv^2(Cr^n). \end{aligned} \tag{24}$$

Hence

$$\left(E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^*\right)^{1/2}r^{-n} < Cr^{-n/2}v(Cr^n). \tag{25}$$

Equations (20) and (25) and the condition (4) together with Lemma 1 imply that

$$\sum\left(\|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}\|_2 + \left(E\mathbf{y}\mathbf{V}(\mathbf{Z}_n)\mathbf{y}^*\right)^{1/2}\right)r^{-n} < \infty. \tag{26}$$

Let $a_n = \|\mathbf{Z}_n\mathbf{Y}\|_2 r^{-n}$ and b_n be $r^{-1}\cdot$ (nth term) of the series in (26). Equations (23) and (26) establish that a_n, b_n satisfy Lemma 4. Hence

$$\sum\|\mathbf{Z}_n\mathbf{y}\|_2 r^{-n} < \infty. \tag{27}$$

This implies (21).

We remark here that we can show (21) for any eigenvector of $\mathbf{M}$ under weaker assumption (3) by considering second moments rather than the L^2 norms. However this approach does not seem to work for generalized eigenvectors.

We show next that (27) and hence (21) hold for generalized eigenvectors. Now let l be an eigenvalue of $\mathbf{M}$ that has a corresponding Jordan block of size s. Let $\mathbf{u}_1$ be the corresponding eigenvector and $\mathbf{u}_2, \ldots, \mathbf{u}_s$ be the corresponding generalized eigenvectors, i.e.,

$$\mathbf{M}\mathbf{u}_t = l\mathbf{u}_t + \mathbf{u}_{t-1}, \qquad 1 < t \le s.$$

The proof of (27) is by induction on t. We know that (27) holds for $t = 1$. Suppose (27) holds for all indices up to $t - 1$. From (14) we obtain

$$E|\mathbf{Z}_{n+1}\mathbf{u}_t|^2 = E|\mathbf{Z}_n\mathbf{u}_t l + \mathbf{Z}_n\mathbf{u}_{t-1} + \mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}_t|^2 + E\mathbf{u}_t\mathbf{V}(\mathbf{Z}_n)\mathbf{u}_t^*.$$

Hence

$$\|\mathbf{Z}_{n+1}\mathbf{u}_t\|_2 \le \|\mathbf{Z}_n\mathbf{u}_t\|_2\,|l| + \|\mathbf{Z}_n\mathbf{u}_{t-1}\|_2 + \|\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}_t\|_2 + \left(E\mathbf{u}_t\mathbf{V}(\mathbf{Z}_n)\mathbf{u}_t^*\right)^{1/2}. \tag{28}$$

Using (20), (26) and the induction assumption (27), we can see that the series of the last three terms in (28) times r^{-n} converges. Now Lemma 4 applies to yield (27) for $\mathbf{u}_t$. Now let $\mathbf{u}_1, \dots, \mathbf{u}_d$ be a basis for $\mathbf{C}^d$ consisting of right eigenvectors and generalized eigenvectors of $\mathbf{M}$. Let $\mathbf{u}_1$ correspond to r and $\mathbf{v}$ be the left eigenvector corresponding to r normalized so that $\mathbf{v}\mathbf{u}_1 = 1$. It is easily seen that $\mathbf{v}$ is orthogonal to $\mathbf{u}_i$, $i \ge 2$. Let $\mathbf{a}$ be arbitrary. Then

$$\mathbf{a} = \sum_{i=1}^{d} a_i\mathbf{u}_i \quad \text{and} \quad \mathbf{v}\mathbf{a} = \sum_{i=1}^{d} a_i\mathbf{v}\mathbf{u}_i = a_1. \tag{29}$$

Hence by (21) and Theorem 2,

$$\mathbf{Z}_n\mathbf{a}r^{-n} = \sum_{i=1}^{d} a_i\mathbf{Z}_n\mathbf{u}_i r^{-n} \to a_1W = \mathbf{v}\mathbf{a}W \quad \text{a.s. and } L^2. \tag{30}$$

This implies the statement of the theorem. □

PROOF OF THEOREM 4. Let $\mathbf{w}_n = \mathbf{Z}_n r^{-n}$ and $\mathbf{w} = \lim \mathbf{w}_n$ in L^1. We show first that all components of $\mathbf{w}$ are positive. From (7),

$$EW_n = EW_0 + r^{-1}\sum_{i=0}^{n-1} E\mathbf{w}_i\mathbf{D}(\mathbf{Z}_i)\mathbf{u}. \tag{31}$$

L^1 convergence implies convergence of moments EW_n which in turn implies

$$\left|\sum E\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}\right| < \infty. \tag{32}$$

Since $\mathbf{D}(\mathbf{z})$ are matrices with elements of the same sign for all $\mathbf{z}$ and $\mathbf{u}$ has positive components

$$|\mathbf{w}_n D(\mathbf{z})\mathbf{u}| = \mathbf{w}_n|\mathbf{D}(\mathbf{z})|\mathbf{u} \tag{33}$$

and by (32),

$$\left|\sum E\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}\right| = E\sum \mathbf{w}_n|\mathbf{D}(\mathbf{Z}_n)|\mathbf{u} < \infty. \tag{34}$$

Equation (33) implies for any $\mathbf{y}$,

$$E|\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}| < CE\mathbf{w}_n|\mathbf{D}(\mathbf{Z}_n)|\mathbf{1} < CE\mathbf{w}_n|\mathbf{D}(\mathbf{Z}_n)|\mathbf{u}. \tag{35}$$

Equations (34) and (35) establish that for any $\mathbf{y}$,

$$\sum E|\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}| < \infty. \tag{36}$$

We show next that for any eigenvector or a generalized eigenvector of $\mathbf{M}$, $\mathbf{y}$ corresponding to an eigenvalue $l \ne r$,

$$\sum |E\mathbf{w}_n\mathbf{y}| < \infty. \tag{37}$$

As in the proof of Theorem 3 we establish (37) for eigenvectors first, then use induction to establish (37) for generalized eigenvectors. If $\mathbf{y}$ is an eigenvector corresponding to l, then by (6),

$$EZ_{n+1}\mathbf{y} = lEZ_n\mathbf{y} + EZ_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}. \tag{38}$$

Hence

$$|E\mathbf{w}_{n+1}\mathbf{y}| \le |l|r^{-1}|E\mathbf{w}_n\mathbf{y}| + r^{-1}E|\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{y}|. \tag{39}$$

In view of $|l| < r$ and (36), (39) shows that $(|E\mathbf{w}_n\mathbf{y}|)$ satisfy the conditions of Lemma 4. Thus (37) holds for eigenvectors. Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s$ be, respectively, an eigenvector and generalized eigenvectors corresponding to l. Suppose (37) holds for $\mathbf{u}_1, \dots, \mathbf{u}_{t-1}$. Then it follows from (6) that

$$|E\mathbf{w}_{n+1}\mathbf{u}_t| < |l|r^{-1}|E\mathbf{w}_n\mathbf{u}_t| + r^{-1}|E\mathbf{w}_n\mathbf{u}_{t-1}| + r^{-1}|E\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}_t|.$$

Induction assumption and (36) allows involving Lemma 4 to yield (37) for $\mathbf{u}_t$. Let eigenvectors and generalized eigenvectors of $\mathbf{M}$ form a basis in C^d with $\mathbf{u}_1$ corresponding to r. Equation (37) implies

$$\lim_{n\to\infty} |E\mathbf{w}_n\mathbf{u}_i| = 0, \tag{40}$$

for any $\mathbf{u}_i$ with $i \ge 2$. Since $\mathbf{w}_n\mathbf{u} \to \mathbf{w}\mathbf{u}$ in L^1 implies $E\mathbf{w}_n\mathbf{u} \to E\mathbf{w}\mathbf{u}$, it follows that $E\mathbf{w}\mathbf{u}_i = 0$, for $i \ge 2$. Since it is assumed that $\mathbf{w} \ne \mathbf{0}$ a.s., this implies $E\mathbf{w} = C\mathbf{v}$. Since all $v_i > 0$ we obtain for all i, $Ew_i > 0$, which in turn implies (i).

Equation (32) together with the assumption of the constant sign of $\mathbf{D}(\mathbf{z})$ implies by (31),

$$\sum_{n=1}^{\infty} \mathbf{w}_n|\mathbf{D}(\mathbf{Z}_n)|\mathbf{u} < \infty \quad \text{a.s.} \tag{41}$$

Since $\mathbf{u}$ has positive coordinates this implies for all j, $1 \le j \le d$,

$$\sum_{i=1}^{d}\sum_{n=1}^{\infty} \mathbf{w}_n^i|d_{ij}(\mathbf{Z}_n)| < \infty.$$

Now $\liminf w_n^i > 0$ on the set $\{w_i > 0\}$, hence

$$\sum_{n=1}^{\infty} |d_{ij}(\mathbf{Z}_n)| < \infty \quad \text{a.s.,} \tag{42}$$

on this set. Since (42) holds for all i, j we obtain statement (ii) of the theorem. Statement (iii) follows from (ii) by Lemma 1. □

PROOF OF THEOREM 5. Consideration of the second moments of W_n gives from (15),

$$EW_{n+1}^2 = EW_n^2 + r^{-2(n+1)}E\Big(\mathbf{u}\mathbf{V}(\mathbf{Z}_n)\mathbf{u} + 2\mathbf{Z}_n\mathbf{u}\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u} + \big(\mathbf{Z}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}\big)^2\Big).$$

This together with L^2 convergence implies

$$\left| \sum_{n=1}^{\infty} E\left(\mathbf{u}\mathbf{V}(\mathbf{Z}_n)\mathbf{u} + 2\mathbf{w}_n\mathbf{u}\mathbf{w}_n D(\mathbf{Z}_n)\mathbf{u} + \left(\mathbf{w}_n\mathbf{D}(\mathbf{Z}_n)\mathbf{u}\right)^2\right) r^{-2n} \right| < \infty.$$

Assumption of constant sign of the summands together with (41) implies

$$\sum_{n=1}^{\infty} \mathbf{u}\mathbf{V}(\mathbf{Z}_n)\mathbf{u} r^{-2n} < \infty \quad \text{a.s.}$$

This implies that for any $i = 1, \dots, d$,

$$\sum_{n=1}^{\infty} \mathbf{u} Z_n^i \Gamma_i(\mathbf{Z}_n)\mathbf{u} r^{-2n} < \infty \quad \text{a.s.}$$

Thus on the set $\{w_i > 0\}$,

$$\sum_{n=1}^{\infty} \mathbf{u}\Gamma_i(\mathbf{Z}_n)\mathbf{u} r^{-n} < \infty \quad \text{a.s.}$$

The rest of the proof follows by Lemma 1. Notice that if all the entries of covariance matrices are positive, then $\sum_{n=1}^{\infty} \mathbf{u}\Gamma_i(\mathbf{Z}_n)\mathbf{u} r^{-n} = \mathbf{u}(\sum_{n=1}^{\infty} \Gamma_i(\mathbf{Z}_n) r^{-n})\mathbf{u}$, from which the convergence of matrix series can be deduced to complete the proof. □

REFERENCES

COHN, H. and KLEBANER, F. C. (1986). Geometric rate of growth in Markov chains with application to population size dependent models with dependent offspring. *Stochastic Anal. Appl.* 4 283–307.

HARRIS, T. E. (1963). *Branching Processes*. Springer, New York.

KLEBANER, F. C. (1984). Geometric rate of growth in population size dependent branching processes. *J. Appl. Probab.* **21** 40–49.

KLEBANER, F. C. (1985). A limit theorem for population size dependent branching processes. *J. Appl. Probab.* **22** 48–57.

SENETA, E. (1981). *Nonnegative Matrices and Markov Chains*. Springer, New York.

DEPARTMENT OF STATISTICS
RICHARD BERRY BUILDING
UNIVERSITY OF MELBOURNE
PARKVILLE, VICTORIA
AUSTRALIA 3052