BRANCHING IN A MARKOVIAN ENVIRONMENT

LILA GRECO AND LIONEL LEVINE

Abstract. A branching process in a Markovian environment consists of an irreducible Markov chain on a set of "environments" together with an offspring distribution for each environment. At each time step the chain transitions to a new random environment, and one individual is replaced by a random number of offspring whose distribution depends on the new environment. We give a first moment condition that determines whether this process survives forever with positive probability. On the event of survival we prove a law of large numbers and a central limit theorem for the population size. We also define a matrix-valued generating function for which the extinction matrix (whose entries are the probability of extinction in state j given that the initial state is i) is a fixed point, and we prove that iterates of the generating function starting with the zero matrix converge to the extinction matrix.

1. Introduction

Let S be a finite set of "environments", and let $(Q_t)_{t\in\mathbb{N}}$ be an irreducible Markov chain on the state space S. We call Q_t the **environment** at time t. To each $i \in S$ we associate a probability distribution ν_i on $\mathbb{N} = \{0, 1, \ldots\}$, which we call the **offspring distribution** of environment i. Consider a population of X_t identical individuals who reproduce one at a time: At each time step t, a single individual is replaced by a random number ξ_{t+1} of offspring, where ξ_{t+1} has the distribution $\nu_{Q_{t+1}}$, independent of the past. Thus, $X_{t+1} = X_t - 1 + \xi_{t+1}$. We refer to the pair (Q_t, X_t) as a **branching process in random environment (BPME)**.

Note that X_t denotes the population after t individuals have reproduced (*not* the population after t generations). The population after n generations is defined recursively by $Y_0 = X_0$ and

$$Y_n = X_{Y_0+Y_1+\ldots+Y_{n-1}} \tag{1}$$

for $n \geq 1$. Our choice of time index t (reproduction events) instead of n (generations) is a matter of convenience: the latter can be recovered from the former by equation (1).

Date: December 29, 2021.

2010 *Mathematics Subject Classification.* 60J80, 60J10, 60K37, 60F05, 15A24, 15B51 .

Key words and phrases. extinction matrix, martingale central limit theorem, matrix generating function, stochastic abelian network.

LL was partially supported by NSF grant DMS-1455272.

1.1. **A brief survey of related processes.** If there is only one environment ($\#S = 1$), then Y_n is easily seen to be a classical Galton-Watson branching process [Har02, AN72]. But if $\#S > 1$ then Y_n is distinct from the classical **multitype branching process** [KS66, KLPP97]. Our population is scalar-valued (all individuals are identical) and the offspring distribution depends on the global environment. By contrast, in multitype branching, the population is vector-valued, and the offspring distribution depends on the type of individual.

Branching processes have been generalized in various ways, but not, to our knowledge, in the manner of the present paper. Athreya and Karlin [AK71] and Jones [Jon04] consider branching in a stationary random environment, in which the environment changes *after each generation* instead of each reproduction event. Their analysis features the composition of a random sequence of univariate generating functions. As we will see (Theorem 1.4), our process can be analyzed by iteration of a single matrix-valued generating function.

A related process arises in queuing theory, the **batch Markovian arrival process** (BMAP) [Luc91]. The BMAP consists of population-environment pairs, but the offspring distribution does not depend on the environment; instead, the wait time between reproduction events depends on the environment, and the probability of transitioning from environment i to environment j depends not on i but on the number of offspring produced. A matrix generating function also figures prominently in the analysis of the BMAP.

The process X_t is an instance of **random walk in random environment** (RWRE), whose environment is constant in space and Markovian in time. There are quite a few central limit theorems for RWRE in a dynamic Markovian environment [BMP97, BZ06, DKL08], but those papers all assumed X_t has bounded increments. They also assumed an ellipticity condition which, in our setup, would amount to requiring that there is a probability measure ν on $\mathbb{N}$ and an $\epsilon > 0$ such that $\nu_i \geq \epsilon\nu$ for all i. Our situation is simplified by the environment being constant in space, which removes the substantial difficulty of proving ergodicity of the scenery process (the environment seen from the point of view of the random walker). So we are able to dispense with the boundedness and ellipticity assumptions: In our CLT (Theorem 1.3) we assume only that the offspring distributions ν_i have finite variance, and we allow ν_i, ν_j to have disjoint supports.

A further difference with RWRE is that we truncate the walk at zero — corresponding to extinction of the branching process — and our limit theorems are on the event of survival. Accordingly, the limiting random variable in Theorem 1.3 is $\chi \mathbb{1}_{\mathcal{S}}$, where χ is normally distributed, $\mathcal{S} = \{X_t > 0 \text{ for all } t\}$ is the event of survival, and χ is independent of $\mathcal{S}$. For RWRE it can be difficult to determine the variance of χ (or even to prove that its variance is not zero, see [RV13]). We give an explicit formula for the variance (Lemma 5.4) in terms of the stationary distribution of the environment chain, the offspring distribution means and variances, and a quantity $\phi(i)$ measuring the "excess fertility" of environment i. The vector $(\phi(i))_{i \in S}$ can be computed from the stationary distribution and the offspring means by solving a system of linear equations.

Branching in a Markovian environment is a special case of a **stochastic abelian network** (SAN) [BL16a]. Indeed, a (unary) SAN can be viewed as a multitype BPME, in which the types correspond to the processors in the network. Survival corresponds to a nonhalting network [CL21]. The original motivation for this paper was to generalize the halting criterion of [BL16b] to stochastic abelian networks. Theorem 1.1 accomplishes this in the special case of a SAN with a single unary abelian processor.

1.2. **Statement of main results.** Let

$$\mu := \sum_{i \in S} \pi_i \mu_i$$

where S is the set of possible environments, $\vec{\pi}$ is the unique stationary distribution of the environment Markov chain, and $\mu_i \leq \infty$ is the mean number of offspring produced by a single individual if the environment is in state i. Our first theorem generalizes the survival/extinction dichotomy for classical branching processes. Let $\mathcal{S} = \{X_t > 0 \text{ for all } t\}$ be the event of survival, Write $\mathbb{P}_{n.i}$ for the law of the BPME started from population n and environment i.

Theorem 1.1 (Survivial/Extinction)**.**

(1) *If $\mu < 1$, then the BPME goes extinct almost surely: $\mathbb{P}_{n.i}(\mathcal{S}) = 0$ for all $n \in \mathbb{N}$ and all $i \in S$.*

(2) *If $\mu = 1$, and the number of offspring produced before the first return to the starting state has positive or infinite variance, then the BPME goes extinct almost surely: $\mathbb{P}_{n.i}(\mathcal{S}) = 0$ for all $n \in \mathbb{N}$ and all $i \in S$.*

(3) *If $\mu > 1$, then the BPME with sufficiently large initial population survives forever with positive probability: For each environment i there exists n such that $\mathbb{P}_{n.i}(\mathcal{S}) > 0$.*

Our next result gives the asymptotic growth rate for the population on the event of survival. Note that X_t denotes the population after t individuals have reproduced (*not* the population after t generations), so the growth is linear rather than exponential in t.

Theorem 1.2 (Asymptotic Growth Rate)**.** *If each offspring distribution has finite variance, then for any initial population $n \geq 1$ and any initial environment i,*

$$\frac{X_t}{t} \to (\mu - 1)\mathbb{1}_{\mathcal{S}} \qquad \mathbb{P}_{n.i}\text{-almost surely as } t \to \infty$$

where $\mathcal{S} = \{X_t > 0 \text{ for all } t\}$ is the event of survival.

Our next result is a central limit theorem for the normalized population on the event of survival.

Theorem 1.3 (Central Limit Theorem)**.** *If each offspring distribution has finite variance, then for any initial population $n \geq 1$ and any initial environment,*

$$\frac{X_t - (\mu - 1)t\mathbb{1}_{\mathcal{S}}}{\sqrt{t}} \Rightarrow \chi \mathbb{1}_{\mathcal{S}} \qquad \text{as } t \to \infty$$

where $\chi \sim \mathcal{N}(0, \sigma_M^2)$ *is a normal random variable independent of the event of survival* $\mathcal{S} = \{X_t > 0 \text{ for all } t\}$*. The variance* σ_M^2 *is computed in Lemma 5.4.*

To state our next result, we define the **matrix generating function**

$$f(M) = \sum_{n \geq 0} P_n M^n$$

where M is an $S \times S$ substochastic matrix, and $(P_n)_{ij}$ is the probability that the environment transitions from state i to state j while producing n offspring. Here we interpret M^0 as the identity matrix. The **extinction matrix** E is the $S \times S$ matrix whose (i,j) entry is the probability that the BMPE started with population 1 in state i goes extinct in state j.

Theorem 1.4 (Extinction Matrix)**.** *The* (i,j) *entry of* E^n *is the probability that the BPME started with population* n *in state* i *goes extinct in state* j*. Moreover,* $f(E) = E$*, and*

$$\lim_{n\to\infty} f^n(M) = E$$

for any matrix M *satisfying* $0 \leq M \leq E$ *entrywise.*

Here f^n denotes the nth iterate of f.

1.3. **Plan of the paper.**

After the formal definitions in Section 2, we prove Theorems 1.1-1.4 in sections 3, 4, 5, and 6, respectively. We conclude with open questions and two conjectures in Section 7.

2. FORMAL DEFINITION AND AN EXAMPLE

2.1. **The Markovian environment.**

Let P be the transition matrix of an irreducible Markov chain on a finite state space S. The entry $P(i,j)$ is the probability of transitioning from state $i \in S$ to state $j \in S$. We call this Markov chain the **environment**. We associate to each state $i \in S$ an offspring distribution. These offspring distributions can be simultaneously described by a stochastic matrix $R : \mathbb{R}^S \to \mathbb{R}^{\mathbb{N}}$ called the **reproduction matrix**, where R_{in} is the probability that an individual has n offspring given that the current environment is i.

Let $(\xi_t^i)_{i \in S,\, t \in \mathbb{N}}$ be independent random variables such that $\mathbb{P}\left(\xi_t^i = n\right) = R_{in}$ for all $i \in S$ and $n \in \mathbb{N}$. We interpret ξ_t^i as the number of offspring produced at time t if the environment chain happens to be in state i at time t.

Definition 2.1 (Branching Process In A Markovian Environment (BPME))**.** A *branching process in a Markovian environment* is a sequence $(X_t, Q_t)_{t \geq 0}$ of population-state pairs that evolves as follows. We begin with initial population $X_0 \in \mathbb{N}_{\geq 1}$ and initial state $Q_0 \in S$. Given X_t and Q_t, we update:

$$Q_{t+1} \sim P(Q_t, \cdot) \text{ independent of } X_0, \dots, X_t, Q_0, \dots, Q_t, \text{ and } (\xi_t^i)_{t \in \mathbb{N}, i \in S}$$

$$X_{t+1} := \begin{cases} X_t - 1 + \xi_{t+1}, & X_t > 0 \\ 0, & X_t = 0 \end{cases} \tag{2}$$

where

$$\xi_{t+1} := \sum_{i \in S} \xi^i_{t+1} \mathbb{1}\{Q_{t+1} = i\}.$$

That is, the global state Q_t is updated to Q_{t+1} by taking one step in the environment Markov chain. If the population is positive, then one individual produces a random number of offspring ξ_{t+1} sampled from the offspring distribution of the new state Q_{t+1}. These new individuals are added to the population, while the reproducing individual is removed from the population. The integer X_t represents the total population after t such reproduction events.

In the branching process literature, time is usually indexed by the number of generations. But in our notation, the time index t is number of reproduction events so far. That is, between time t and time $t+1$, a single individual reproduces (assuming the population has not yet reached zero). This difference in indexing is convenient since the environment updates every time an individual reproduces.

The sequence $(X_t)_{t \geq 0}$ alone is not a Markov chain, but $((X_t, Q_t))_{t \geq 0}$ is a Markov chain. A state of the latter chain is an ordered pair (n, i) where $n \in \mathbb{N}$ and $i \in S$. We call such pairs **total states**. Following [BL16a] we adopt the notation $n.i$ for the total state (n, i).

Let

$$\mu_i := \sum_{n=0}^{\infty} n R_{in}$$

be the mean of the offspring distribution of state i. Let π_i be the stationary probability of environment i, and let

$$\mu := \sum_{i \in S} \mu_i \pi_i \tag{3}$$

be the mean number of offspring produced when a single individual reproduces in the stationary environment. This value μ will play a role analogous to the mean of the offspring distribution in the ordinary Galton-Watson branching process.

2.2. **An Example.** The following example illustrates why Theorem 1.1(3) requires a sufficiently large starting population for the BPME to have positive probability of surviving forever.

Let the environment chain be $S = \{a, b\}$ with $P_{ab} = P_{ba} = 1$. Its stationary distribution is $\pi_a = \pi_b = \frac{1}{2}$. Let the offspring distribution of state a be uniform on $\{0, 1, 2, 3, 4, 5\}$, and let the offspring distribution of state b be 0 with probability 1. Then $\mu = \mu_a \pi_a + \mu_b \pi_b = 5/4$. The matrix generating function is given by

$$f(M) = \begin{pmatrix} 0 & 1 \\ \frac{1}{6} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \frac{1}{6} & 0 \end{pmatrix} (M + M^2 + M^3 + M^4 + M^5).$$

The extinction matrix E solves the degree 5 polynomial equation $f(E) = E$. Using Theorem 1.4, we can estimate it by calculating $f^n(O)$ for large n:

$$E \approx \begin{pmatrix} 0 & 1 \\ 0.2459 & 0.3497 \end{pmatrix}.$$

The BPME started in state a with initial population $X_0 = 1$ goes extinct immediately (in state b), despite the fact that $\mu > 1$. However, the extinction matrix for initial population $X_0 = 2$ is given by:

$$E^2 \approx \begin{pmatrix} 0.2459 & 0.3497 \\ 0.0860 & 0.3681 \end{pmatrix}.$$

Since both row sums are < 1, this BPME with initial population 2 has positive probability to survive forever in either starting state.

In this example, Theorem 1.2 tells us that on the event of survival, this BPME population satisfies $X_t/t \to 1/4$ almost surely. Theorem 1.3 tells us that on the event of survival, the population X_t is asymptotically normal with mean $\frac{1}{4}t$ and variance $\frac{35}{24}t$.

2.3. **Review of Markov Chains.** We recall a few facts about Markov chains, which we will apply in later sections to the environment chain. Let $(Q_t)_{t\in\mathbb{N}}$ be an irreducible Markov chain with finite state space and transition matrix P. Denote by $\mathbb{P}_i$ and $\mathbb{E}_i$ the probability and expectation given $Q_0 = i$, and let $\tau_i := \min\{t \geq 1 \mid Q_t = i\}$.

Proposition 2.2. *(i) There is a unique probability row vector π such that $\pi P = \pi$. Moreover, $\pi_i > 0$ for all states i.*

(ii) For all $i, j \in S$

$$\mathbb{P}_i\left(\lim_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \mathbb{1}\{Q_s = j\} = \pi_j\right) = 1.$$

(iii) $\mathbb{E}_i[\tau_i] = \frac{1}{\pi_i} < \infty$.

(iv) If $N_j := \sum_{t=1}^{\tau_i} \mathbb{1}\{Q_t = j\}$ is the number of visits to j until hitting i, then $\mathbb{E}_i[N_j] = \pi_j \mathbb{E}_i[\tau_i]$. Note that (iii) is the special case obtained by setting $j = i$.

(v) There exist constants $C_0, C_1 > 0$ such that for all $t \geq 0$ and all states i, j,

$$\mathbb{P}_i(\tau_j > t) \leq C_0 e^{-C_1 t}.$$

Proof. (i) See [LP17], Proposition 1.14 (i) and Corollary 2.17.

(ii) This follows from the Birkhoff ergodic theorem.

(iii) See [LP17] Proposition 1.14 (ii).

(iv) This follows from [LP17] Lemma 10.5.

(v) See [AF02] Section 2.4.3

□

3. The $\mathbb{Z}$-Valued Process

The proof of Theorem 1.1 will proceed by comparing the BPME to a process whose population is allowed to become negative. This **$\mathbb{Z}$-valued BPME** is a sequence $(Y_t, Q_t)_{t\geq 0}$ of population-state pairs, with initial state $Q_0 \in S$, but now the initial population Y_0 is allowed to take values in $\mathbb{Z}$, and reproduction occurs regardless of whether the population is positive. Using the same definitions and notation as in Definition 2.1, the update rule is given by:

$$\begin{aligned} &Q_{t+1} \sim P(Q_t, \cdot) \text{ independent of } X_0, \ldots, X_t, Q_0, \ldots, Q_{t-1} \\ &Y_{t+1} := Y_t - 1 + \xi_{t+1} \end{aligned} \tag{4}$$

where

$$\xi_t := \sum_{i \in S} \xi_t^i \mathbb{1}\{Q_{t+1} = i\}.$$

Notice that we can recover the original BPME from the $\mathbb{Z}$-valued BPME with $Y_0 \geq 1$ by setting:

$$X_t = \begin{cases} Y_t & \text{if } Y_s > 0 \text{ for all } 0 \leq s < t \\ 0 & \text{else.} \end{cases}$$

Note that $X_t > 0$ for all t if and only if $Y_t > 0$ for all t.

3.1. Excursions of the environment.

The proof of Theorem 1.1 will proceed by considering excursions of the environment chain from its starting state in the $\mathbb{Z}$-valued BPME.

Fix a starting environment $Q_0 = i$. Let $\tau_0 = 0$, and for $n \geq 1$ let

$$\tau_n := \inf\{t > \tau_{n-1} \mid Q_t = i\}.$$

be the time of nth return to state i.

Let

$$\Delta_n = \Delta_n^i = Y_{\tau_n} - Y_{\tau_{n-1}}.$$

be the net population change during the nth excursion from state i.

Lemma 3.1. *The sequence $(\Delta_n)_{n\geq 1}$ is independent and identically distributed (i.i.d.).*

Proof. We have

$$\Delta_n = \sum_{t=\tau_{n-1}+1}^{\tau_n} (\xi_t - 1) = \sum_{t=\tau_{n-1}+1}^{\tau_n} \sum_{j \in S} (\xi_t^j - 1)\mathbb{1}\{Q_t = j\}.$$

By the strong Markov property, the sequence $(Q_{\tau_{n-1}+1}, \ldots, Q_{\tau_n})$ has the same distribution for each n, and is independent of $(Q_0, \ldots, Q_{\tau_{n-1}})$. In addition, the stacks $(\xi_t^j)_{t\geq 0}$ are independent of the stopping times τ_n. Hence, for fixed j, the sequence $(\xi^j_{\tau_{n-1}+1}, \ldots, \xi^j_{\tau_n})$ has the same distribution for each n, and is independent of every other such sequence for varying n and j. It then follows from the expression for Δ_n above that $(\Delta_n)_{n\geq 1}$ is an i.i.d. sequence. ☐

Lemma 3.1 implies that the sequence $(Y_{\tau_n})_{n\geq 0}$ is a random walk with i.i.d. steps $\Delta_1, \Delta_2, \ldots$. The long-term behavior of such a walk is determined by the mean of the step distribution.

Proposition 3.2 ([Dur19, Exercise 5.4.1])**.** *Let $\Delta_1, \Delta_2, \ldots$ be i.i.d. real-valued random variables with finite mean, and $Y_n = Y_0 + \sum_{i=1}^n \Delta_i$. There are only four possibilities, one of which has probability one.*

(1) If $\mathbb{P}(\Delta_1 = 0) = 1$*, then* $Y_n = Y_0$ *for all* n*.*
(2) If $\mathbb{E}[\Delta_1] > 0$*, then* $Y_n \to \infty$*.*
(3) If $\mathbb{E}[\Delta_1] < 0$*, then* $Y_n \to -\infty$*.*
(4) If $\mathbb{E}[\Delta_1] = 0$ *and* $\mathbb{P}(\Delta_1 = 0) < 1$*, then* $-\infty = \liminf Y_n < \limsup Y_n = \infty$*.*

Therefore we need to calculate the expectation of Δ_1. Recall that $\mu = \sum_{i\in S} \mu_i \pi_i$ is the mean number of offspring produced when a single individual reproduces in the stationary environment, and τ_1 is the time of first return to the starting environment.

Lemma 3.3. *The mean number of offspring produced in one excursion from the starting state is* $\mathbb{E}[\Delta_1] = \mathbb{E}[\tau_1](\mu - 1)$*.*

Proof. Define $\mathcal{F}_t = \sigma(Y_0, \dots, Y_{t-1}, Q_0, \dots, Q_t)$. Note the inclusion of Q_t: this sigma field includes all information up to the time right before the tth individual reproduces. Then,

$$\begin{aligned}
\mathbb{E}[\Delta_n] &= \sum_{t=\tau_{n-1}+1}^{\tau_n} \sum_{j\in S} \mathbb{E}\left[\mathbb{E}\left[(\xi_t^j - 1)\mathbb{1}\{Q_t = j\} \mid \mathcal{F}_t\right]\right] \\
&= \sum_{t=\tau_{n-1}+1}^{\tau_n} \sum_{j\in S} \mathbb{E}\left[\mathbb{1}\{Q_t = j\}\mathbb{E}\left[(\xi_t^j - 1) \mid \mathcal{F}_t\right]\right] \\
&= \sum_{t=\tau_{n-1}+1}^{\tau_n} \sum_{j\in S} \mathbb{E}\left[\mathbb{1}\{Q_t = j\}(\mu_j - 1)\right] \\
&= \sum_{j\in S} (\mu_j - 1)\mathbb{E}\left[\sum_{t=\tau_{n-1}+1}^{\tau_n} \mathbb{1}\{Q_t = j\}\right].
\end{aligned}$$

Now by the strong Markov property and Proposition 2.2(iv), we have

$$\mathbb{E}\left[\sum_{t=\tau_{n-1}+1}^{\tau_n} \mathbb{1}\{Q_t = j\}\right] = \pi(j)\mathbb{E}[\tau_n - \tau_{n-1}].$$

By the strong Markov property we also have $\mathbb{E}[\tau_n - \tau_{n-1}] = \mathbb{E}[\tau_1]$. Thus,

$$\begin{aligned}
\mathbb{E}[\Delta_n] &= \sum_{j\in S} (\mu_j - 1)\pi(j)\mathbb{E}[\tau_1] \\
&= \mathbb{E}[\tau_1]\left(\sum_{j\in S} \mu_j \pi(j) - \sum_{j\in S} \pi(j)\right) \\
&= \mathbb{E}[\tau_1](\mu - 1).
\end{aligned}$$

□

3.2. **Extinction in the subcritical and critical cases.**

We now prove items (1) and (2) of Theorem 1.1.

Theorem 3.4. *Let* $(X_t, Q_t)_{t\geq 0}$ *be a BPME with any initial population* X_0 *and any initial environment* Q_0*.*

If $\mu < 1$, then $X_t \to 0$ almost surely.
If $\mu = 1$ and $0 < \operatorname{Var}(\Delta_1) \leq \infty$, then $X_t \to 0$ almost surely.

Proof. The $\mathbb{Z}$-valued process $Y_{\tau_n} = Y_0 + \sum_{k=1}^{n} \Delta_k$, at time of nth return to the initial environment Q_0, is a random walk with i.i.d. increments Δ_k. By Lemma 3.3, we have $\mathbb{E}[\Delta_k] = \mathbb{E}[\tau](\mu - 1)$. In the case $\mu < 1$, we are in case (3) of Proposition 3.2; in the case $\mu = 1$ and $\operatorname{Var}(\Delta_1) > 0$ we are in case (4). So in either case,

$$\liminf_{n\to\infty} Y_{\tau_n} = -\infty \qquad \text{a.s.}.$$

So with probability one there exists a time T such that $Y_T \leq 0$. Since Y is integer-valued and decreases by at most one at each time step, for the minimal such T we have $Y_T = 0$, and so $X_t = 0$ for all $t \geq T$. $\square$

3.3. **Survival in the supercritical case.** The proof of Theorem 1.1(3) will also proceed by studying Δ_n, the net population change during the nth excursion from the starting state in the $\mathbb{Z}$-valued BPME. If we apply Proposition 3.2 to the case $\mathbb{E}[\Delta_1] > 0$, we find that $Y_{\tau_n} \to \infty$ almost surely. However, this does not imply $X_t > 0$ for all t almost surely, or even with positive probability: We could have chosen a starting state that dooms the process to extinction in the first step, as in Example 2.2. To rule out this kind of scenario, we make the following definition, in which $\mathbb{P}_{m.i}$ denotes the law of the $\mathbb{Z}$-valued BPME started with $(Y_0, Q_0) = (m, i)$, and $\tau := \inf\{t \geq 1 \mid Q_t = i\}$ is the time of first return to state i.

Definition 3.5. Total state $m.i$ is **viable** if

$$\mathbb{P}_{m.i}\left(Y_\tau - Y_0 \geq 1 \text{ and } Y_t \geq 1 \text{ for all } t \in [0,\tau]\right) > 0.$$

In words, $m.i$ is viable if it is possible to start in state i with m individuals and return to state i with at least $m+1$ individuals, while keeping the population positive the whole time. Note that if $m.i$ is viable then $(m+1).i$ is viable.

Lemma 3.6. *Suppose $\mu > 1$. Then for every state $i \in S$, there exists m such that $m.i$ is viable.*

Proof. Fixing $i \in S$, let Cyc be the set of all tuples $(y_0.q_0, \ldots, y_n.q_n)$ (of any length) with each $y_t \in \mathbb{Z}$ and each $q_t \in S$, such that $y_0 = 0$ and $q_0 = q_n = i$ and $q_t \neq i$ for all $1 \leq t < n$. Let

$$\mathbb{P}\left((y_0.q_0, \ldots, y_n.q_n)\right) := \mathbb{P}_{y_0.q_0}\left(Y_t.Q_t = y_t.q_t \text{ for all } 1 \leq t \leq n\right).$$

The mean population change over one excursion from state i is

$$\mathbb{E}_{0.i}[\Delta_1] = \sum_{c \in \mathrm{Cyc}} y_n \mathbb{P}(c).$$

Since $\mu > 1$ we have $\mathbb{E}_{0.i}[\Delta_1] = \mathbb{E}[\tau](\mu - 1) > 0$ by Lemma 3.3, so at least one term on the right side is positive. Hence there exists $c = (0.q_0, \ldots, y_n.q_n) \in \mathrm{Cyc}$ such that $y_n \geq 1$ and $\mathbb{P}(c) > 0$. Since $y_0 = 0$ and the population can decrease by at most one per time step, each $y_i \geq -i$. So the excursion

$$c' = ((y_0 + n).q_0, \ldots, (y_n + n).q_n)$$

has $\mathbb{P}(c') = \mathbb{P}(c) > 0$ and all $y_i + n \geq 1$ and $y_n + n \geq n + 1$, so $n.i$ is viable. $\square$

Remark. A similar argument shows that if no state's offspring distribution is concentrated on 0, then $1.i$ is viable for all states i.

We are now ready to prove the main result of this section.

Theorem 3.7. *Let X_t be a BPME with $\mu > 1$. Then for all viable $m.i$ we have*

$$\mathbb{P}_{m.i}\left(X_t > 0 \text{ for all } t\right) > 0.$$

Proof. Let $m.i$ be viable. Write $\mathbb{P} = \mathbb{P}_{m.i}$ for the law of the $\mathbb{Z}$-valued the BPME $(Y_t, Q_t)_{t\geq 0}$, started with initial population m and initial state i. Since $X_t > 0$ for all t if and only if $Y_t > 0$ for all t, it suffices to prove $\mathbb{P}\left(Y_t > 0 \text{ for all } t\right) > 0$.

Since $m.i$ is viable, there exists δ such that

$$\mathbb{P}\left(Y_\tau \geq m+1 \text{ and } Y_t \geq 1 \text{ for all } t \in (0,\tau]\right) \geq \delta > 0.$$

By the strong Markov property and induction on n it follows that

$$\mathbb{P}\left(Y_{\tau_n} \geq m+n \text{ and } Y_t \geq 1 \text{ for all } t \in (0,\tau_n]\right) \geq \delta^n$$

for all $n \geq 1$. Write $\mathcal{E}_n = \{Y_{\tau_n} \geq m+n\}$. By the strong Markov property at time τ_n,

$$\begin{aligned}\mathbb{P}\left(Y_t > 0 \text{ for all } t\right) &\geq \mathbb{P}\left(Y_t > 0 \text{ for all } t > \tau_n \mid \mathcal{E}_n \text{ and } Y_t \geq 1 \text{ for all } t \in (0,\tau_n]\right) \cdot \delta^n \\ &= \mathbb{P}\left(Y_t > 0 \text{ for all } t > \tau_n \mid \mathcal{E}_n\right) \cdot \delta^n.\end{aligned}$$

So it suffices to find n such that the right side is strictly positive, or equivalently,

$$\mathbb{P}\left(Y_t = 0 \text{ for some } t > \tau_n \mid \mathcal{E}_n\right) < 1.$$

Consider the following events for $k \geq 0$:

$$\begin{aligned} A_k &:= \{Y_t = 0 \text{ for some } t \in (\tau_k, \tau_{k+1}]\} \\ B_k &:= \left\{Y_{\tau_k} \leq \frac{\mu-1}{2}\mathbb{E}\left[\tau\right] k\right\} \\ C_k &:= \left\{\tau_{k+1} - \tau_k \geq \frac{\mu-1}{2}\mathbb{E}\left[\tau\right] k\right\}\end{aligned}$$

where $\tau = \tau_1$. Observe that

$$\{Y_t = 0 \text{ for some } t > \tau_n\} = \bigcup_{k=n}^{\infty} A_k$$

and for each k,

$$A_k \subseteq B_k \cup C_k.$$

The latter holds because Y_t decreases by at most one at each time step; hence if $Y_{\tau_k} > \frac{\mu-1}{2}\mathbb{E}\left[\tau\right] k$, then Y_t cannot reach 0 before time τ_{k+1} unless $\tau_{k+1} - \tau_k \geq \frac{\mu-1}{2}\mathbb{E}\left[\tau\right] k$. We wish to show

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k \mid \mathcal{E}_n\right) < 1.$$

Recall from Lemmas 3.1 and 3.3 that $Y_{\tau_k} = Y_0 + \sum_{m=1}^{k} \Delta_m$, where the Δ_m are i.i.d. with mean $\mathbb{E}\left[\tau\right](\mu - 1)$. By the strong law of large numbers, $\frac{1}{k}Y_{\tau_k} \to$

$\mathbb{E}[\tau](\mu-1)$ almost surely as $k\to\infty$. In particular, this implies $\mathbb{P}(B_k \text{ i.o.})=0$, so, $\mathbb{P}(\bigcup_{k=n}^{\infty} B_k)\to 0$ as $n\to\infty$. By the FKG inequality, since B_k is a decreasing event and $\mathcal{E}_n$ is an increasing event with respect to the offspring random variables $(\xi_t^j)_{j\in S,\, t\in\mathbb{N}}$, we have

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} B_k \mid \mathcal{E}_n\right) \le \mathbb{P}\left(\bigcup_{k=n}^{\infty} B_k\right).$$

Since each $\tau_{k+1}-\tau_k$ has the same distribution as τ, we have

$$\sum_{k\ge 0}\mathbb{P}(C_k)<\infty.$$

In addition, C_k is independent of $\mathcal{E}_n$ for all $k\ge n$ (since $\tau_{k+1}-\tau_k$ is independent of Y_{τ_n} by the strong Markov property).

Now choose n large enough so that $\mathbb{P}(\bigcup_{k=n}^{\infty} B_k)<\frac{1}{2}$ and $\sum_{k=n}^{\infty}\mathbb{P}(C_k)<\frac{1}{2}$. Then

$$\begin{aligned}\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k \mid \mathcal{E}_n\right) &\le \mathbb{P}\left(\bigcup_{k=n}^{\infty} B_k \mid \mathcal{E}_n\right)+\mathbb{P}\left(\bigcup_{k=n}^{\infty} C_k \mid \mathcal{E}_n\right)\\ &\le \mathbb{P}\left(\bigcup_{k=n}^{\infty} B_k\right)+\mathbb{P}\left(\bigcup_{k=n}^{\infty} C_k\right)\\ &< \frac{1}{2}+\frac{1}{2}\end{aligned}$$

where we have used the FKG inequality to remove the conditioning on the B term, and independence to remove the conditioning on the C term. □

Theorem 1.1(3) follows immediately from Theorem 3.7 together with Lemma 3.6.

4. Laws of large numbers

In this section we prove strong laws of large numbers for Y_t and X_t. As above, the environment chain $(Q_t)_{t\ge 0}$ is irreducible with stationary distribution π, and for each state i the offspring distribution has finite mean μ_i, and we let $\mu=\sum_{i\in S}\pi_i\mu_i$.

Theorem 4.1. *For any initial populations $Y_0\in\mathbb{Z}$ and $X_0\in\mathbb{N}_{\ge 1}$, it holds almost surely as $t\to\infty$*

$$\frac{Y_t}{t}\to\mu-1$$

and

$$\frac{X_t}{t}\to(\mu-1)\mathbb{1}_{\mathcal{S}}$$

where $\mathcal{S}=\{X_t\ge 1 \text{ for all } t\}$ is the event of survival.

Proof. We start by observing that

$$Y_t-Y_0=\sum_{s=1}^{t}(\xi_s-1)=\sum_{s=1}^{t}\sum_{i\in S}\mathbb{1}\{Q_s=i\}(\xi_s^i-1).$$

Switching the order of summation, the right side can be written as

$$\sum_{i\in S} Z^i_{L_t}$$

where Z^i_k denotes a sum of k independent copies of $\xi^i_1 - 1$, and $L_t = \sum_{s=1}^t \mathbb{1}\{Q_s = i\}$ is the local time of state i.

By the strong law of large numbers for i.i.d. sums, $Z^i_k/k \to \mu_i - 1$ almost surely as $k \to \infty$. It follows from Proposition 2.2(ii) that

$$\frac{Z^i_{L_t}}{t} = \frac{Z^i_{L_t}}{L_t}\frac{L_t}{t} \to (\mu_i - 1)\pi_i$$

almost surely as $t \to \infty$. Summing over i yields $Y_t/t \to \mu - 1$ almost surely.

The strong law for X_t now follows by observing that $X_t = Y_t$ on $\mathcal{S}$ and $X_t/t \to 0$ almost surely on $\mathcal{S}^c$. $\square$

5. Central Limit Theorems

In this section we prove a central limit theorem for Y_t and X_t under a second moment assumption on the offspring distributions. We will proceed by defining a martingale involving Y_t, calculating its quadratic variation, and invoking the martingale central limit theorem. Then to pass from Y_t to X_t, we show that the limiting normal random variable is independent of the event of survival.

5.1. The martingale.

As above, suppose the environment chain $(Q_t)_{t\geq 0}$ is irreducible with transition matrix P and stationary distribution π, and for each state i the offspring distribution has finite mean μ_i (We will impose finite second moment in the next section, but it is not needed yet). Write $\vec{\pi}$ and $\vec{\mu}$ for the column vectors with coordinates π_i and μ_i respectively. As above, let

$$\mu = \langle \vec{\pi}, \vec{\mu}\rangle := \sum_{i\in S} \pi_i \mu_i.$$

Lemma 5.1. *There is a unique vector $\vec{\varphi} \in \mathbb{R}^S$ satisfying $\langle \vec{\pi}, \vec{\varphi}\rangle = 0$ and*

$$\varphi_i - \sum_{j\in S} P_{ij}\varphi_j = \mu_i - \mu$$

for all $i \in S$.

Proof. Since $\vec{\pi}$ is the unique stationary distribution of the environment chain, the left null space of $I-P$ is spanned by $\vec{\pi}$. The column space of $I-P$ is the orthogonal complement of the left null space,

$$\mathrm{Im}(I-P) = \{\vec{v} \in \mathbb{R}^S : \langle \vec{\pi}, \vec{v}\rangle = 0\}.$$

We check that $\mu\vec{1} - \vec{\mu}$ is orthogonal to $\vec{\pi}$, and thus is in the image of $I - P$:

$$\begin{aligned}\sum_{i\in S}\pi_i(\mu - \mu_i) &= \mu\sum_{i\in S}\pi_i - \sum_{i\in S}\pi_i\mu_i \\ &= \mu - \mu \\ &= 0.\end{aligned}$$

Therefore there exists a vector $\vec{\varphi}$ such that $(I-P)\vec{\varphi} = \vec{\mu} - \mu\vec{1}$ as desired.

The right null space of $I-P$ is spanned by $\vec{1}$, so $\vec{\varphi}$ is unique up to adding scalar multiples of $\vec{1}$. Therefore there is a unique such vector satisfying $\langle \vec{\pi}, \vec{\varphi} \rangle = 0$. ☐

We will see that the vector $\vec{\varphi}$ has a natural interpretation in terms of the BPME: Its coordinate $\varphi(i)$ represents the **long-term excess fertility** of environment i, in the sense of Corollary 5.3 below.

We now define a martingale for the $\mathbb{Z}$-valued branching process $(Y_t, Q_t)_{t\geq 0}$, adapted to the filtration

$$\mathcal{F}_t := \sigma((\xi^i_s)_{i\in S, 0\leq s\leq t}, (Q_s)_{0\leq s\leq t+1}).$$

Note the inclusion of Q_{t+1}: This sigma-algebra tells us what state we will transition to next, but not how many offspring will be produced.

Lemma 5.2. *Write $\varphi(i)$ for the i^{th} coordinate of the vector $\vec{\varphi}$ of Lemma 5.1. Then*

$$M_t := Y_t - (\mu - 1)t + \varphi(Q_{t+1})$$

is a martingale adapted to $\mathcal{F}_t$.

Note that $Y_t - \sum_{s=1}^t (\mu_{Q_s} - 1)$ is also a martingale, but we will find M_t much more useful!

Proof. Recall that $Y_t = Y_{t-1} - 1 + \xi_t$ where $\xi_t = \sum_{i\in S} \xi^i_t \mathbb{1}\{Q_t = i\}$ and ξ^i_t is sampled from the offspring distribution of state i. Hence

$$\mathbb{E}\left[M_t \mid \mathcal{F}_{t-1}\right] = Y_{t-1} - 1 - (\mu-1)t + \mathbb{E}\left[\xi_t + \varphi(Q_{t+1}) \mid \mathcal{F}_{t-1}\right].$$

Now

$$\begin{aligned}
\mathbb{E}\left[\xi_t + \varphi(Q_{t+1}) \mid \mathcal{F}_{t-1}\right] &= \sum_{i\in S} \mathbb{1}\{Q_t = i\}\mathbb{E}\left[\xi^i_t + \sum_{j\in S} \mathbb{1}\{Q_{t+1} = j\}\varphi(j) \mid \mathcal{F}_{t-1}\right] \\
&= \sum_{i\in S} \mathbb{1}\{Q_t = i\}\left(\mu_i + \sum_{j\in S} P(i,j)\varphi(j)\right) \\
&= \sum_{i\in S} \mathbb{1}\{Q_t = i\}(\mu + \varphi(i)) \\
&= \mu + \varphi(Q_t)
\end{aligned}$$

where we have used the fact that $\sum_{j\in S} P(i,j)\varphi(j) - \varphi(i) = \mu - \mu_i$. Combining this with the above, we have

$$\begin{aligned}
\mathbb{E}\left[M_t \mid \mathcal{F}_{t-1}\right] &= Y_{t-1} - 1 - (\mu-1)t + \mu + \varphi(Q_t) \\
&= M_{t-1}.
\end{aligned}$$

☐

Corollary 5.3. *If the environment chain is aperiodic, then*

$$\varphi(i) = \lim_{t\to\infty} \left(\mathbb{E}_{0.i}\left[Y_t\right] - (\mu-1)t\right).$$

Proof. Let $Y_0 = 0$ and $Q_0 = i$. Writing $\mathbb{E} = \mathbb{E}_{0.i}$, equating $M_0 = \mathbb{E}\left[M_t\right]$ yields

$$\varphi(i) = \mathbb{E}\left[Y_t\right] - (\mu - 1)t - \mathbb{E}\left[\varphi(Q_{t+1})\right].$$

Since the environment is aperiodic, $\mathbb{P}\left(Q_t = j\right) \to \pi_j$ as $t \to \infty$, and hence

$$\mathbb{E}\left[\varphi(Q_{t+1})\right] \to \langle \vec{\pi}, \vec{\varphi} \rangle = 0. \qquad \square$$

5.2. **Quadratic Variation.** Assume now that each offspring distribution has mean μ_i and variance $\sigma_i^2 < \infty$. In this case $EM_t^2 < \infty$; to see this, note that φ is bounded (since the state space S is finite) and

$$Y_t = Y_0 + \sum_{s=1}^{t} \sum_{i \in S} \mathbb{1}\{Q_s = i\}(\xi_s^i - 1)$$

and each ξ_s^i is square-integrable, so M_t is a finite sum of square-integrable random variables. The **quadratic variation** associated with M_t is

$$V_t := \sum_{s=1}^{t} \mathbb{E}\left[(M_t - M_{t-1})^2 \mid \mathcal{F}_{t-1}\right].$$

Lemma 5.4. *$V_t/t \to \sigma_M^2$ almost surely as $t \to \infty$, where*

$$\sigma_M^2 := \sum_{i \in S} \pi_i(\sigma_i^2 - (\mu - \mu_i)^2 + 2\mu_i\varphi_i) \tag{5}$$

where $\vec{\varphi}$ is given by Lemma 5.1.

Proof. The increments of M are given by

$$M_s - M_{s-1} = \xi_s - \mu + \varphi(Q_{s+1}) - \varphi(Q_s).$$

Squaring and taking conditional expectation, we break the result into three terms:

$$\begin{aligned}\mathbb{E}\left[(M_s - M_{s-1})^2 \mid \mathcal{F}_{s-1}\right] = {} & \underbrace{\mathbb{E}\left[(\xi_s - \mu)^2 \mid \mathcal{F}_{s-1}\right]}_{(1)} + \underbrace{\mathbb{E}\left[(\varphi(Q_{s+1}) - \varphi(Q_s))^2 \mid \mathcal{F}_{s-1}\right]}_{(2)} \\ & + \underbrace{2\mathbb{E}\left[(\xi_s - \mu)(\varphi(Q_{s+1}) - \varphi(Q_s)) \mid \mathcal{F}_{s-1}\right]}_{(3)}.\end{aligned}$$

We consider these terms one at a time. First, since Q_s is $\mathcal{F}_{s-1}$-measurable,

$$(1) = \sum_{i \in S} \mathbb{1}\{Q_s = i\}\mathbb{E}\left[(\xi_s^i - \mu)^2 \mid \mathcal{F}_{s-1}\right].$$

The conditioning on the right side can be dropped by independence. Adding and subtracting μ_i we obtain

$$\sum_{i \in S} \mathbb{1}\{Q_s = i\}\mathbb{E}((\xi_s^i - \mu_i) - (\mu - \mu_i))^2$$

Taking the time average, since $\frac{1}{t}\sum_{s=1}^{t}\mathbb{1}\{Q_s = i\} \to \pi_i$ almost surely as $t \to \infty$, we have

$$\frac{1}{t}\sum_{s=1}^{t}\sum_{i\in S}\mathbb{1}\{Q_s = i\}(\sigma_i^2 + (\mu - \mu_i)^2) \to \sigma^2 + \tau^2 \tag{6}$$

where

$$\sigma^2 := \sum_{i\in S}\pi_i\sigma_i^2$$

and

$$\tau^2 := \sum_{i\in S}\pi_i(\mu - \mu_i)^2.$$

For the second term,

$$\begin{aligned}(2) &= \mathbb{E}\left[(\varphi(Q_{s+1}) - \varphi(Q_s))^2 \mid \mathcal{F}_{s-1}\right]\\ &= \sum_{i\in S}\mathbb{1}\{Q_s = i\}\sum_{j\in S}P(i,j)(\varphi(j) - \varphi(i))^2\end{aligned}$$

Expanding the square and using the definition of φ, this becomes

$$\begin{aligned}&= \sum_{i\in S}\mathbb{1}\{Q_s = i\}\left[\sum_{j\in S}P(i,j)\varphi(j)^2 - 2\varphi(i)\underbrace{\sum_{j\in S}P(i,j)\varphi(j)}_{\mu-\mu_i+\varphi(i)} + \varphi(i)^2\right]\\ &= \sum_{i\in S}\mathbb{1}\{Q_s = i\}\left[\sum_{j\in S}P(i,j)\varphi(j)^2 - 2\varphi(i)(\mu - \mu_i) - \varphi(i)^2\right].\end{aligned}$$

Now taking the time average, and writing $\vec{\varphi^2}$ for the vector whose ith entry is $\varphi(i)^2$ and $\vec{\mu}\vec{\varphi}$ for the vector whose ith entry is $\mu_i\varphi(i)$, we obtain

$$\begin{aligned}&\frac{1}{t}\sum_{s=1}^{t}\sum_{i\in S}\mathbb{1}\{Q_s = i\}\left[\sum_{j\in S}P(i,j)\varphi(j)^2 - 2\mu\varphi(i) + 2\mu_i\varphi(i) - \varphi(i)^2\right]\\ &\to \sum_{i\in S}\pi_i\left[\sum_{j\in S}P(i,j)\varphi(j)^2 - 2\mu\varphi(i) + 2\mu_i\varphi(i) - \varphi(i)^2\right]\\ &= \sum_{j\in S}\varphi(j)^2\underbrace{\sum_{i\in S}\pi_iP(i,j)}_{\pi_j} - 2\mu\underbrace{\langle\vec{\pi},\vec{\varphi}\rangle}_{0} + 2\langle\vec{\pi},\vec{\mu}\vec{\varphi}\rangle - \left\langle\vec{\pi},\vec{\varphi^2}\right\rangle\\ &= 2\langle\vec{\pi},\vec{\mu}\vec{\varphi}\rangle.\end{aligned} \tag{7}$$

For the third term, since ξ_s^i is independent of $\sigma(\mathcal{F}_{s-1}, Q_{s+1})$, we have

$$\begin{aligned}
&2\mathbb{E}\left[(\xi_s-\mu)(\varphi(Q_{s+1})-\varphi(Q_s)) \mid \mathcal{F}_{s-1}\right] \\
&\quad = 2\sum_{i\in S}\mathbb{1}\{Q_s=i\}\mathbb{E}\left[(\xi_s^i-\mu)\left(\sum_{j\in S}\mathbb{1}\{Q_{s+1}=j\}\varphi(j)-\varphi(i)\right) \mid \mathcal{F}_{s-1}\right] \\
&\quad = 2\sum_{i\in S}\mathbb{1}\{Q_s=i\}(\mu_i-\mu)\left(\sum_{j\in S}P(i,j)\varphi(j)-\varphi(i)\right) \\
&\quad = 2\sum_{i\in S}\mathbb{1}\{Q_s=i\}(\mu_i-\mu)(\mu-\mu_i).
\end{aligned}$$

The limit of the time average is

$$-\frac{2}{t}\sum_{s=1}^{t}\sum_{i\in S}\mathbb{1}\{Q_s=i\}(\mu_i-\mu)^2 \to -2\tau^2. \tag{8}$$

Adding (6), (7), and (8), we conclude that

$$\frac{V_t}{t} \to \sigma^2-\tau^2+2\langle\vec{\pi},\vec{\mu}\vec{\varphi}\rangle = \sigma_M^2. \qquad \square$$

Remark. $\sigma_M^2 \geq 0$ since it is a limit of nonnegative random variables. In particular, taking all ξ_t^i deterministic so that the first term $\sigma^2=0$, we obtain the inequality

$$\tau^2 \leq 2\langle\vec{\pi},\vec{\mu}\vec{\varphi}\rangle. \tag{9}$$

It would be interesting to give a more direct proof of this algebraic fact. Note that if at least one of the offspring distributions has positive variance, then $\sigma^2>0$ and hence $\sigma_M^2>0$ by (9). On the other hand, if all offspring distributions are deterministic, then a necessary and sufficient condition for $\sigma_M^2=0$ is that all excursions from a fixed state have the same net number of offspring. To avoid trivialities, we assume from now on that $\sigma_M^2>0$.

5.3. Applying the martingale CLT.

Our goal in this section is to prove the following central limit theorems for Y_t and X_t.

Theorem 5.5. *Assume that $\sigma_M^2>0$. Then we have convergence in distribution*

$$\frac{Y_t-(\mu-1)t}{\sqrt{t}} \Rightarrow \chi$$

and

$$\frac{X_t-(\mu-1)t\mathbb{1}_{\mathcal{S}}}{\sqrt{t}} \Rightarrow \chi\mathbb{1}_{\mathcal{S}}$$

where χ is a normal random variable with mean 0 and variance σ_M^2, and χ is independent of $\mathcal{S}$, the event of survival.

To prove these, we will use the following version of the martingale central limit theorem. As above, let V_t be the quadratic variation associated to the martingale M_t. Write $K_t=M_t-M_{t-1}$.

Proposition 5.6 (Martingale CLT, see [Dur19, Theorem 8.2.8])**.** *Suppose that as* $t \to \infty$

- *(i)* $\frac{V_t}{t} \to \sigma_M^2 > 0$ *in probability, and*
- *(ii)* $\frac{1}{t}\sum_{s=1}^{t} \mathbb{E}\left[K_s^2 \mathbb{1}\{|K_s| > \epsilon\sqrt{t}\}\right] \to 0$ *for all* $\epsilon > 0$.

Then $M_t/\sqrt{t} \Rightarrow \mathcal{N}(0, \sigma_M^2)$ *as* $t \to \infty$.

We start by verifying the above conditions (i) and (ii) for the martingale defined in Lemma 5.2. Condition (i) follows from Lemma 5.4.

To check the Lindeberg condition (ii), let J be a constant such that $|\varphi(i)| \le J$ for all $i \in S$. Since $\xi_s \ge 0$,

$$|K_s| = |\xi_s - \mu + \varphi(Q_{s+1}) - \varphi(Q_s)| \le \xi_s + \mu + 2J.$$

If t is sufficiently large so that $\frac{\epsilon}{2}\sqrt{t} > \mu + 2J$, then:

$$\{|K_s| > \epsilon\sqrt{t}\} \subset \{\xi_s > \frac{\epsilon}{2}\sqrt{t}\}.$$

In addition, on the event $\xi_s > \frac{\epsilon}{2}\sqrt{t}$, with t large enough that $\frac{\epsilon}{2}\sqrt{t} > \mu + 2J$, we have:

$$K_s^2 \le (\xi_s + \mu + 2J)^2 < (2\xi_s)^2.$$

Hence we have:

$$\mathbb{E}\left[K_s^2 \mathbb{1}\{|K_s| > \epsilon\sqrt{t}\}\right] \le \mathbb{E}\left[K_s^2 \mathbb{1}\{\xi_s > \frac{\epsilon}{2}\sqrt{t}\}\right] \le \mathbb{E}\left[4\xi_s^2 \mathbb{1}\{\xi_s > \frac{\epsilon}{2}\sqrt{t}\}\right].$$

We will show that this quantity goes to 0 uniformly in s as $t \to \infty$. We have:

$$\begin{aligned}
\mathbb{E}\left[\xi_s^2 \mathbb{1}\{\xi_s > \frac{\epsilon}{2}\sqrt{t}\}\right] &= \sum_{i\in S} \mathbb{E}\left[\mathbb{E}\left[(\xi_s^i)^2 \mathbb{1}\{\xi_s^i > \frac{\epsilon}{2}\sqrt{t}, Q_s = i\} \mid \mathcal{F}_{s-1}\right]\right] \\
&= \sum_{i\in S} \mathbb{E}\left[\mathbb{1}\{Q_s = i\}\mathbb{E}\left[(\xi_s^i)^2 \mathbb{1}\{\xi_s^i > \frac{\epsilon}{2}\sqrt{t}\}\right]\right] \\
&= \sum_{i\in S} \mathbb{P}\left(Q_s = i\right) \mathbb{E}\left[(\xi_s^i)^2 \mathbb{1}\{\xi_s^i > \frac{\epsilon}{2}\sqrt{t}\}\right] \\
&\le \max_{i\in S} \mathbb{E}\left[(\xi_s^i)^2 \mathbb{1}\{\xi_s^i > \frac{\epsilon}{2}\sqrt{t}\}\right].
\end{aligned}$$

Now since $\mathbb{E}\left[(\xi_s^i)^2\right] < \infty$, we have $\mathbb{E}\left[(\xi_s^i)^2 \mathbb{1}\{\xi_s^i > \frac{\epsilon}{2}\sqrt{t}\}\right] \to 0$ as $t \to \infty$. Moreover, this rate is uniform in s since $(\xi_s^i)_{s\ge 0}$ are i.i.d. samples from the offspring distribution of state i. This verifies condition (ii) of the martingale CLT and hence we have shown

$$M_t/\sqrt{t} \Rightarrow \chi \sim \mathcal{N}(0, \sigma_M^2). \tag{10}$$

Proof of Theorem 5.5. Since $(Y_t - (\mu - 1)t) - M_t$ is bounded, it follows from (10) that

$$Z_t := \frac{Y_t - (\mu - 1)t}{\sqrt{t}} \Rightarrow \chi.$$

To prove the CLT for X_t, recall that $X_t = 0$ eventually on $\mathcal{S}^c$, and that $X_t = Y_t$ for all t on $\mathcal{S}$. Applying the CLT for Y_t, we find that

$$\frac{X_t - (\mu - 1)t\mathbb{1}_{\mathcal{S}}}{\sqrt{t}}$$

converges in distribution to 0 on $\mathcal{S}^c$, and to χ on $\mathcal{S}$. It remains to show that χ is independent of $\mathcal{S}$. To this end, fix an environment state i and let

$$s = \inf\{u > t^{1/4} \,:\, Q_u = i\}.$$

(The choice of $t^{1/4}$ is unimportant; any function tending to ∞ slower than $\sqrt{t}$ will do.) By the strong Markov property, the random variable

$$Z'_t := \frac{Y_t - Y_s - (\mu - 1)t}{\sqrt{t}}.$$

is independent of the event

$$\mathcal{S}_s := \{Y_u > 0 \text{ for all } u < s\}$$

of survival up to time s.

Note that $s = t^{1/4} + \tau$ for a random variable τ satisfying $P(\tau > a) \leq \max_j P(\tau_{ji} > a)$ for all a, where τ_{ji} is the first hitting time of state i starting from state j. By Proposition 2.2(v) the hitting times τ_{ji} have exponential tails, so by Borel-Cantelli we have $s/\sqrt{t} \to 0$ almost surely. By Theorem 4.1 we have $Y_s/s \to \mu - 1$ almost surely, so

$$\frac{Y_s}{\sqrt{t}} = \frac{Y_s}{s}\frac{s}{\sqrt{t}} \to 0$$

almost surely, and hence $Z_t - Z'_t \to 0$ almost surely. Since $\mathcal{S}_s \downarrow \mathcal{S}$, and Z'_t is independent of $\mathcal{S}_s$, for any fixed $\epsilon > 0$ and $a \in \mathbb{R}$ we have for large enough t

$$\begin{aligned} P(Z_t > a, \mathcal{S}) &\leq P(Z_t > a, \mathcal{S}_s) \\ &\leq P(Z'_t > a - \epsilon, \mathcal{S}_s) \\ &= P(Z'_t > a - \epsilon)P(\mathcal{S}_s) \\ &\to P(\chi > a - \epsilon)P(\mathcal{S}) \end{aligned}$$

as $t \to \infty$. Likewise,

$$\begin{aligned} P(Z_t > a, \mathcal{S}) &\geq P(Z_t > a, \mathcal{S}_s) - \epsilon \\ &\geq P(Z'_t > a + \epsilon, \mathcal{S}_s) - \epsilon \\ &\to P(\chi > a + \epsilon)P(\mathcal{S}) - \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that

$$P(Z_t > a, \mathcal{S}) \to P(\chi > a)P(\mathcal{S}). \qquad \square$$

Remark. Under the same hypotheses as Proposition 5.6, the martingale convergence theorem gives the stronger conclusion that $(M_{tu}/\sqrt{t})_{u\in[0,1]}$ converges weakly on $C[0,1]$ to $(\sigma_M B_u)_{u\in[0,1]}$ where B is a standard Brownian motion. This yields a corresponding strengthening of Theorem 5.5, namely

$$\left(\frac{X_{tu} - (\mu-1)tu\mathbb{1}_{\mathcal{S}}}{\sqrt{t}}\right)_{u\in[0,1]} \Rightarrow (\sigma_M B_u \mathbb{1}_{\mathcal{S}})_{u\in[0,1]}$$

as $t \to \infty$, where B is a standard Brownian motion independent of $\mathcal{S}$, the event of survival.

6. Matrix Generating Function

6.1. **Extinction matrix.** In Theorem 1.1 we obtained qualitative results about the survival of BPME. In this section we introduce a matrix generating function to obtain quantitative estimates of the extinction probabilities.

In the ordinary Galton-Watson branching process with offspring distribution $(p_0, p_1, p_2, \dots)$, the generating function for the offspring distribution is given by:

$$f(x) := \sum_{k=0}^{\infty} p_k x^k, \quad |x| \le 1.$$

Many elementary branching process results can be obtained by analyzing the generating function [Har02, AN72]. For instance,

- If q is the extinction probability of the branching process, then q is the smallest fixed point of f in $[0,1]$.
- If $\mu \le 1$, then $q = 1$.
- If $\mu > 1$, then $q \in [0,1)$ and q is the unique fixed point of f in $[0,1)$.
- For every $t \in [0,1)$, we have $\lim_{n\to\infty} f^n(t) = q$, where f^n refers to the nth iterate of f.

We will prove some analogous results for BPME, namely that the extinction matrix is a fixed point of the generating function, and that iterates of the generating function starting at any matrix (entrywise) between the zero matrix and the extinction matrix converge to the extinction matrix.

Recall that P denotes the environment chain transition matrix and R denotes the reproduction matrix: R_{jn} is the probability of producing n offspring if the environment state is j. For $n \in \mathbb{N}$ we define the $S \times S$ matrix P_n by

$$(P_n)_{ij} := P_{ij} R_{jn}$$

the probability that environment i transitions to environment j and n offspring are produced.

A nonnegative matrix M with real entries is called **stochastic** if all of its row sums are 1, and **substochastic** if all of its row sums are ≤ 1. Note that P_n is substochastic for each n. We define the **matrix generating function**

$$f(M) = \sum_{n=0}^{\infty} P_n M^n \tag{11}$$

where we interpret $M^0 = I$ (the $S \times S$ identity matrix).

We make the following observations about $f(M)$.

Lemma 6.1. *Let f be the matrix generating function of a BPME. Then*

- *$f(M)$ converges for all substochastic matrices M.*
- *If M is substochastic, then $f(M)$ is substochastic.*
- *If M is stochastic, then $f(M)$ is stochastic.*

Proof. Note that $\sum_{n=0}^{\infty} P_n = P$, the transition matrix of the environment chain. Writing $\mathbf{1}$ for the all 1's vector and $\leq$ for coordinatewise inequality of vectors, a matrix M with nonnegative entries is substochastic if and only if $M\mathbf{1} \leq \mathbf{1}$, and equality holds if and only if M is stochastic.

Let M be substochastic. Since all entries of $M, P_0, P_1, \ldots$ are nonnegative, we have

$$f(M)\mathbf{1} = \sum_{n=0}^{\infty} P_n M^n \mathbf{1} \leq \sum_{n=0}^{\infty} P_n \mathbf{1} = P\mathbf{1} = \mathbf{1}$$

and if M is stochastic then equality holds. □

For integers x and y, and environments i and j, denote by $\{x.i \to y.j\}$ the event that the total state transitions from $x.i$ to $y.j$ in one time step; that is, state i transitions to state j and $y - x + 1$ offspring are produced. This event has probability

$$\mathbb{P}\left(x.i \to y.j\right) = \begin{cases} P_{ij} R_{j,y-x+1} & \text{if } y \geq x - 1 \\ 0 & \text{else.} \end{cases}$$

We now introduce a matrix of extinction probabilities. Recall that $n.i$ denotes the total state with population n and environment i. We say that the initial total state $X_0.Q_0 = n.i$ *halts in* $0.j$ if $T := \inf\{t : X_t = 0\}$ is finite and satisfies $Q_T = j$. The **extinction matrix** is the $S \times S$ matrix E with entries

$$E_{ij} = \mathbb{P}\left(1.i \text{ halts in } 0.j\right).$$

Note that E is substochastic, since for all $i \in S$

$$\sum_{j \in S} E_{ij} = \mathbb{P}\left(1.i \text{ halts in } 0.j \text{ for some } j\right) \leq 1.$$

Let E^n be the nth power of the extinction matrix, and let f be the matrix generating function (11).

Lemma 6.2. *$(E^n)_{ij} = \mathbb{P}\left(n.i \text{ halts in } 0.j\right)$, and $f(E) = E$.*

Proof. We prove the first part by induction on n. If it holds for E^n, then

$$(E^{n+1})_{ij} = \sum_{k \in S} \mathbb{P}\left(n.i \text{ halts in } 0.k\right) \mathbb{P}\left(1.k \text{ halts in } 0.j\right).$$

The population must reach 1 before reaching 0, as it decreases by at most one per time step. Now $\mathbb{P}\left(n.i \text{ halts in } 0.k\right)$ is also the probability that the BPME started at $(n+1).i$ eventually reaches a population of 1 individual, and the first time it does so it is in environment k. Hence, the above sum is equal to $\mathbb{P}\left((n+1).i \text{ halts in } 0.j\right)$, completing the induction.

Now observe that

$$(P_n E^n)_{ij} = \sum_{k\in S} \mathbb{P}\left(1.i \to n.k\right) \mathbb{P}\left(n.k \text{ halts in } 0.j\right).$$

Summing over n, the (i,j) entry of $\sum_{n=0}^{\infty} P_n E^n$ equals $\mathbb{P}\left(1.i \text{ halts in } 0.j\right)$, which is E_{ij}. Thus $f(E) = E$. □

Now we state the main goal of this section.

Theorem 6.3. *Let O be the $S \times S$ zero matrix. Then $\lim_{n\to\infty} f^n(O) = E$.*

Write $M \le N$ if $M_{ij} \le N_{ij}$ for all $i,j \in S$. Note that if M, N are substochastic and $M \le N$ then

$$f(M) \le f(N). \tag{12}$$

Corollary 6.4. *If $O \le M \le E$, then $\lim_{n\to\infty} f^n(M) = E$.*

Proof. By (12), $O \le M \le E$ implies $f^n(O) \le f^n(M) \le f^n(E)$ for all $n \ge 1$. Taking the limit, we find

$$E = \lim_{n\to\infty} f^n(O) \le \lim_{n\to\infty} f^n(M) \le \lim_{n\to\infty} f^n(E) = E.$$ □

6.2. **Extinction in m generations.** To prepare for the proof of Theorem 6.3, we develop some notation describing the multi-step transitions of the BPME.

Denote by $\{x.i \xrightarrow{n} y.j\}$ the event that the total state transitions from $x.i$ to $y.j$ in n time steps; that is, there exist total states $x_1.k_1, \dots, x_{n-1}.k_{n-1}$ such that

$$x.i \to x_1.k_1 \to \cdots \to x_{n-1}.k_{n-1} \to y.j.$$

Next we define events $\rightsquigarrow_m$ and $\overset{n}{\rightsquigarrow}_m$ to describe how the population can decrease over longer time periods. For example, the event $\{x.i \rightsquigarrow_2 (x-1).j\}$ means that one individual produces any number of offspring, but all offspring of the first individual fail to produce any offspring, and the environment transitions from i to j in the process. Likewise, $\{x.i \rightsquigarrow_m (x-1).j\}$ can be interpreted as a single individual's family tree going extinct in at most m generations. Finally, $\overset{n}{\rightsquigarrow}_m$ means that each of n individuals' family trees go extinct in at most m generations.

Formally, these events are defined as follows. We first define $x.i \overset{n}{\rightsquigarrow}_1 y.j$ if and only if $y = x - n$ and $x.i \xrightarrow{n} y.j$. We define $x.i \overset{0}{\rightsquigarrow}_m y.j$ if and only if $x = y$ and $i = j$. For $m, n \ge 2$ we recursively define

$$\{x.i \overset{1}{\rightsquigarrow}_m (x-1).j\} := \bigcup_{n\in\mathbb{N}} \bigcup_{k\in S} \{x.i \to (x-1+n).k \overset{n}{\rightsquigarrow}_{m-1} (x-1).j\}$$

$$\{x.i \overset{n}{\rightsquigarrow}_m (x-n).j\} := \bigcup_{k\in S} \{x.i \overset{1}{\rightsquigarrow}_m (x-1).k \overset{n-1}{\rightsquigarrow}_m (x-n).j\}.$$

The union defining $\overset{1}{\rightsquigarrow}_m$ includes $n = 0$, and the corresponding event is $\{x.i \to (x-1).j\}$. We write $\rightsquigarrow_m$ to mean $\overset{1}{\rightsquigarrow}_m$.

We make a few observations:

- $\overset{n}{\rightsquigarrow}_m$ results in population decrease of exactly n. Moreover, the population at the end is strictly smaller than the population at any previous time.

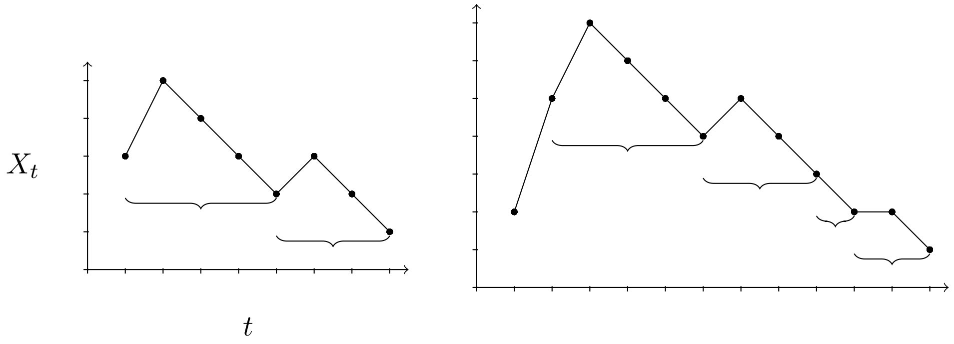

FIGURE 1. Left: An example of the event $\overset{2}{\rightsquigarrow}_2$. Right: An example of the event $\rightsquigarrow_3$. Each $\rightsquigarrow_2$ event is marked with a brace.

- Since extinction in at most m generations implies extinction in at most $m+1$ generations,
$$\{1.i \rightsquigarrow_m 0.j\} \subseteq \{1.i \rightsquigarrow_{(m+1)} 0.j\} \tag{13}$$
as can be verified from the formal definition by induction on m.
- Extinction in t time steps implies extinction in at most t generations. Conversely, extinction in m generations implies extinction in a finite number of time steps. Hence
$$\bigcup_{m\geq 1} \{1.i \rightsquigarrow_m 0.j\} = \bigcup_{t\geq 1} \{1.i \xrightarrow{t} 0.j\}. \tag{14}$$
- For all $x, y \in \mathbb{Z}$, we have $\mathbb{P}\left(x.i \overset{n}{\rightsquigarrow}_m (x-n).j\right) = \mathbb{P}\left(y.i \overset{n}{\rightsquigarrow}_m (y-n).j\right)$.

Some illustrations of these events are shown in Figure 1. Only the population size is depicted, not the state of the environment.

The next lemma gives an interpretation of the entries of the nth power of the mth iterate of f applied to the zero matrix.

Lemma 6.5. *For all $n \geq 0$, $m \geq 1$, we have*
$$(f^m(O)^n)_{ij} = \mathbb{P}\left(n.i \overset{n}{\rightsquigarrow}_m 0.j\right).$$

Proof. First if $n = 0$, then by our convention $M^0 = I$ (the identity matrix), we have $(f^m(O)^0)_{ij} = I_{ij} = \mathbb{P}\left(0.i \overset{0}{\rightsquigarrow}_m 0.j\right)$ for all m, as desired. Next, if $m = n = 1$, then $(f^1(O)^1)_{ij} = (P_0)_{ij} = \mathbb{P}\left(1.i \overset{1}{\rightsquigarrow}_1 0.j\right)$, as desired.

Now we proceed by induction on the pair (m, n) in lexicographic order. Supposing the lemma holds for the pairs $(m, 1)$ and (m, n), we check that it holds for the pair $(m, n+1)$:

$$\begin{aligned} f^m(O)^{n+1}_{ij} &= \sum_{k\in S} f^m(O)_{ik} f^m(O)^n_{kj} \\ &= \sum_{k\in S} \mathbb{P}\left(1.i \overset{1}{\rightsquigarrow}_m 0.k\right)\mathbb{P}\left(n.k \overset{n}{\rightsquigarrow}_m 0.j\right) \\ &= \sum_{k\in S} \mathbb{P}\left((n+1).i \overset{1}{\rightsquigarrow}_m n.k\right)\mathbb{P}\left(n.k \overset{n}{\rightsquigarrow}_m 0.j\right) \\ &= \mathbb{P}\left((n+1).i \overset{(n+1)}{\rightsquigarrow}_m 0.j\right). \end{aligned}$$

Finally, supposing the lemma holds for all pairs $(m,0), (m,1), \dots$ we check that it holds for the pair $(m+1, 1)$:

$$\begin{aligned} (f^{m+1}(O)^1)_{ij} &= \sum_{n=0}^{\infty} (P_n f^m(O)^n)_{ij} \\ &= \sum_{n=0}^{\infty}\sum_{k\in S} (P_n)_{ik}(f^m(O)^n)_{kj} \\ &= \sum_{n=0}^{\infty}\sum_{k\in S} \mathbb{P}(1.i \to n.k)\,\mathbb{P}\left(n.k \overset{n}{\rightsquigarrow}_m 0.j\right) \\ &= \mathbb{P}\left(1.i \overset{1}{\rightsquigarrow}_{(m+1)} 0.j\right). \end{aligned}$$

This completes the induction. □

Proof of Theorem 6.3. By definition of the extinction matrix,

$$\begin{aligned} E_{ij} &= \mathbb{P}(1.i \text{ halts in } 0.j) \\ &= \mathbb{P}\left(\bigcup_{t\geq 1}\{1.i \xrightarrow{t} 0.j\}\right) \\ &= \mathbb{P}\left(\bigcup_{m\geq 1}\{1.i \rightsquigarrow_m 0.j\}\right) \\ &= \lim_{m\to\infty} \mathbb{P}(\{1.i \rightsquigarrow_m 0.j\}). \end{aligned}$$

In the second to last line we have used (14), and in the last line we have used (13). By Lemma 6.5, the right side equals $\lim_{m\to\infty} f^m(O)_{ij}$. □

7. Open Questions

7.1. **Infinite state space.** We assumed the environment Markov chain has a finite state space. We expect our results to extend to positive recurrent Markov chains (perhaps assuming a tail bound on the offspring distributions and hitting times). We used exponential tails of hitting times to prove Theorem 5.5, but one can check in the proof that $2+\delta$ moments suffice.

The null recurrent case is more subtle. Here the difficulty is that the random variable Δ_1 (the net number of offspring produced in an excursion from the starting environment) is no longer integrable: $\mathbb{E}\Delta_1^+ = \mathbb{E}\Delta_1^- = \infty$.

The transient case can have quite different behavior, as shown by the next two examples.

Example. Let the environment chain be a simple random walk on $\mathbb{Z}^3$, with offspring distribution

$$\xi_t^i = \begin{cases} 0 & \text{with probability } (|i|+2)^{-3} \\ 1 & \text{with probability } 1-(|i|+2)^{-3} \end{cases}$$

where $|i|$ denotes the Euclidean norm of $i \in \mathbb{Z}^3$. Even though $\mathbb{E}\xi_t^i < 1$ for all i, the resulting BPME survives with positive probability. The basic estimate used to prove survival is

$$P(L_r \geq ar^2) \leq c_0 e^{-c_1 a} \tag{15}$$

where $L_r = \sum_{t=0}^\infty \mathbb{1}\{|Q_t| < r\}$ is the total time spent by the random walk in the ball $\{i \in \mathbb{Z}^3 : |i| < r\}$. This estimate can be used to show that the number of times t such that $\xi_t^{Q_t} = 0$ is almost surely finite.

Example. Fix $m > 0$. Let the environment chain be a simple random walk on $\mathbb{Z}^3$, with offspring distribution

$$\xi_t^i = \begin{cases} 0 & \text{with probability } 1-(|i|+2)^{-3} \\ m(|i|+2)^3 & \text{with probability } (|i|+2)^{-3}. \end{cases}$$

Then $\mathbb{E}\xi_t^i = m$ for all i, but the resulting BPME goes extinct almost surely. The proof uses (15) to show that almost surely, only finitely many $\xi_t^{Q_t}$ are nonzero.

7.2. Uniqueness of the fixed point.

By Theorem 1.1, the extinction matrix E is stochastic if and only if $\mu \leq 1$ and $P \neq P_1$. We list some open questions about the matrix generating function $f(M) = \sum P_n M^n$.

1. In the case $\mu \leq 1$ and $P \neq P_1$, is the extinction matrix E the unique fixed point of f?
2. In the case $\mu > 1$, is E the unique fixed point of f that is not stochastic?
3. Does it hold for every substochastic matrix M that is not stochastic, that $\lim_{n\to\infty} f^n(M) = E$?

7.3. Stochastic fixed point?

Iterates of f starting at the identity matrix have a natural interpretation: $f^n(I)_{ij}$ is the probability starting with population 1 in environment i that the environment is j after n generations. One might expect that as $n \to \infty$, the environment after n generations would converge to the stationary distribution π on the event of survival, but this is wrong! Let v be the left Perron-Frobenius eigenvector of the extinction matrix, normalized so that its coordinates sum to 1. Experiments suggest that as $n \to \infty$, the distribution of the environment after n generations converges to v on the event of survival, which motivates the following conjecture.

Conjecture 7.1.

$$\lim_{n\to\infty} f^n(I)_{ij} = E_{ij} + \left(1 - \sum_{k\in S} E_{ik}\right) v_j.$$

7.4. **Multitype BPME.** Consider a network of BPMEs, where the offspring of each BPME are sent to other BPMEs. Formally, let $G = (V, E)$ be a finite directed graph, with the following data associated to each vertex $v \in V$:

(1) A finite set S_v (the **state space** of v).
(2) A stochastic matrix $P_v : \mathbb{R}^{S_v} \to \mathbb{R}^{S_v}$ (the **transition matrix** of v).
(3) A stochastic matrix $R_v : \mathbb{R}^{S_v} \to \mathbb{R}^{\prod_w \mathbb{N}}$ (the **reproduction matrix** of v).

Here the product is over out-neighbors w of v. When an individual at vertex v reproduces, the state of v updates according to the transition matrix P_v, and the individual at v is replaced by a random number of offspring at each out-neighbor w of v. The reproduction matrix R_v specifies the distribution of this offspring vector, which depends on the state of vertex v. This process continues unless there are no individuals left, in which case the network is said to **halt**. The **abelian property** [BL16a] ensures that the probability of halting does not depend on the order in which individuals reproduce. Moreover, on the event that the network halts, the distribution of the final states of the vertices does not depend on the order in which individuals reproduce.

If the transition matrix P_v is irreducible, then it has a unique stationary distribution π_v. Let $\mu_{vw}(i)$ denote the mean number of offspring sent to vertex w when an individual at vertex v reproduces in state i. Then the long-term average number of offspring sent from v to w when an individual at vertex v reproduces is

$$M_{vw} := \sum_{i\in S_v} \pi_v(i)\mu_{vw}(i).$$

Denote by μ the Perron-Frobenius eigenvalue of the $V \times V$ matrix M.

Conjecture 7.2. *If $\mu < 1$, then the network halts almost surely for any initial state and population.*

If $\mu = 1$ and there are no conserved quantities, then the network halts almost surely for any initial state and population.

If $\mu > 1$, then for sufficiently large initial population the network has a positive probability not to halt.

Here a **conserved quanity** is a collection of real numbers a_v and functions $\varphi_v : S_v \to \mathbb{R}$ for each $v \in V$, such that

$$\sum_{v\in V} a_v X_v + \varphi_v(Q_v)$$

is an almost sure constant, where X_v is the number of individuals at vertex v, and Q_v is the state of vertex v.

Conjecture 7.2 is a common generalization of Theorem 1.1 and the main result of [BL16b]: The former is the case $\#V = 1$, and the latter is the case that all offspring distributions are deterministic.

Acknowledgment

We thank Rodrigo Delgado for his careful reading of an early draft.

References

[AF02] David Aldous and James Allen Fill, *Reversible Markov chains and random walks on graphs*, 2002, Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[AK71] Krishna B. Athreya and Samuel Karlin, *On Branching Processes with Random Environments: I: Extinction Probabilities*, The Annals of Mathematical Statistics **42** (1971), no. 5, 1499 – 1520.

[AN72] K. B. Athreya and P. E. Ney, *Branching processes*, Dover Publications, Inc., Mineola, NY, 1972. MR 2047480

[BL16a] Benjamin Bond and Lionel Levine, *Abelian networks I. Foundations and examples*, SIAM Journal on Discrete Mathematics **30** (2016), no. 2, 856–874.

[BL16b] ———, *Abelian networks II. Halting on all inputs*, Selecta Mathematica **22** (2016), no. 1, 319–340.

[BMP97] C. Boldrighini, R.A. Minlos, and A. Pellegrinotti, *Almost-sure central limit theorem for a markov model of random walk in dynamical random environment*, Probability theory and related fields **109** (1997), no. 2, 245–273.

[BZ06] Antar Bandyopadhyay and Ofer Zeitouni, *Random walk in dynamic Markovian random environment*, ALEA Lat. Am. J. Probab. Math. Stat. **1** (2006), 205–224.

[CL21] Swee Hong Chan and Lionel Levine, *Abelian networks IV. Dynamics of nonhalting networks*, Memoirs of the American Mathematical Society (2021), to appear.

[DKL08] Dmitry Dolgopyat, Gerhard Keller, and Carlangelo Liverani, *Random walk in Markovian environment*, The Annals of Probability **36** (2008), no. 5, 1676–1710.

[Dur19] Rick Durrett, *Probability: theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, 2019, Fifth edition.

[Har02] Theodore E. Harris, *The theory of branching processes*, Dover Publications, 2002, Corrected reprint of the 1963 original.

[Jon04] Galin L. Jones, *On the Markov chain central limit theorem*, Probab. Surv. **1** (2004), 299–320. MR 2068475

[KLPP97] Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres, *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes*, Classical and modern branching processes, Springer, 1997, pp. 181–185.

[KS66] Harry Kesten and Bernt P. Stigum, *A limit theorem for multidimensional Galton-Watson processes*, The Annals of Mathematical Statistics **37** (1966), no. 5, 1211–1223.

[LP17] David A. Levin and Yuval Peres, *Markov chains and mixing times*, vol. 107, American Mathematical Society, 2017.

[Luc91] David M Lucantoni, *New results on the single server queue with a batch Markovian arrival process*, Communications in Statistics. Stochastic Models **7** (1991), no. 1, 1–46.

[RV13] Frank Redig and Florian Völlering, *Random walks in dynamic random environments: a transference principle*, The Annals of Probability **41** (2013), no. 5, 3157–3180.

Lila Greco, Department of Mathematics, Cornell University, Ithaca, NY 14853.

Lionel Levine, Department of Mathematics, Cornell University, Ithaca, NY 14853.
`https://pi.math.cornell.edu/~levine`