The critical Branching Markov Chain is transient

Nina Gantert and Sebastian Müller

23.01.06

Abstract

We investigate recurrence and transience of Branching Markov Chains (BMC) in discrete time. Branching Markov Chains are clouds of particles which move (according to an irreducible underlying Markov Chain) and produce offspring independently. The offspring distribution can depend on the location of the particle. If the offspring distribution is constant for all locations, these are Tree-Indexed Markov chains in the sense of [1]. Starting with one particle at location x, we denote by $\alpha(x)$ the probability that x is visited infinitely often by the cloud. Due to the irreducibility of the underlying Markov Chain, there are three regimes: either $\alpha(x) = 0$ for all x (transient regime), or $0 < \alpha(x) < 1$ for all x (weakly recurrent regime) or $\alpha(x) = 1$ for all x (strongly recurrent regime). We give classification results, including a sufficient condition for transience in the general case. If the mean of the offspring distribution is constant, we give a criterion for transience involving the spectral radius of the underlying Markov Chain and the mean of the offspring distribution. In particular, the critical BMC is transient. Examples for the classification are provided.
KEYWORDS: Branching Markov Chains, recurrence and transience, Lyapunov function, spectral radius
AMS 2000 MATHEMATICS SUBJECT CLASSIFICATION: 60J10, 60J80

1 Introduction

A Branching Markov Chain (BMC) is a system of particles, which move independently according to the transition probabilities of an underlying Markov chain. We take a countable state space X and an irreducible stochastic transition kernel P for the underlying Markov chain (X, P). The BMC starts with one particle in an arbitrary starting position $x_s \in X$ at time 0. Particles move independently according to P. At each position $x \in X$, they independently produce offspring according to some probability distribution $\mu(x)$ on $\{1, 2, 3, \dots\}$ (which can depend on the position x of the particle) and die. We assume that there is always at least one offspring particle, so that the number of particles is always increasing in time. Similar models have been studied in [7].

The transition probabilities of the Markov chain and the offspring distribution can be given as a (typical) realization of a random environment. The behavior of the resulting "Branching Random Walk in Random Environment" has been classified in [3], [5] and [6] for the case where the underlying Markov chain is a Random Walk in Random Environment on $\mathbb{Z}^+$ or on a tree. A similar, but more general model, where movement and offspring production are not independent anymore, is considered in [4].

Let $\alpha(x)$ be the probability that, starting the BMC from $x_s = x$, the location x is visited by infinitely many particles. Using the irreducibility of the underlying Markov Chain, we obtain, similar to Lemma 3.1 in [1], the following classification:

Lemma 1.1. *There are three possible regimes:*

$$\alpha(x) = 0 \quad \forall x \in X \tag{1}$$

(transient regime)

$$0 < \alpha(x) < 1 \quad \forall x \in X \tag{2}$$

(weakly recurrent regime)

$$\alpha(x) = 1 \quad \forall x \in X \tag{3}$$

(strongly recurrent regime).

We write $\alpha \equiv 0$ ($\equiv 1$) if $\alpha(x) = 0$ ($= 1$) $\forall x \in X$. We say that a BMC is *recurrent* if it is not transient, i.e. if (2) or (3) are satisfied. Note that in the weakly recurrent regime, the values of $\alpha(x)$ do in general not coincide.

We first give a sufficient condition, Theorem 3.1, for transience where the Markov chain can be any irreducible Markov chain and the branching distributions can be arbitrary. Under the assumption of constant mean offspring we obtain in Theorem 3.2 a classification in transience and recurrence for all irreducible Markov chains. In particular, we show that in the critical case the BMC is transient. It is left to forthcoming work to study the subdivision of the recurrent phase. Under homogeneity conditions, i.e. quasi-transitivity, on the BMC we show that the strongly recurrent regime coincides with the recurrent regime, i.e. (2) does not occur, see Theorem 3.4.

2 Preliminaries

We give the definition of the spectral radius of an irreducible Markov chain (X, P) and quote a result which characterizes the spectral radius in terms of $t-$superharmonic functions. For further details see e.g. [8].

Definition 2.1. Let (X, P) be an irreducible Markov chain with countable state space X and transition operator $P = (p(x, y))_{x,y \in X}$. The spectral radius of (X, P) is defined as

$$\rho(P) := \limsup_{n\to\infty} \left(p^{(n)}(x, y)\right)^{1/n} \in (0, 1], \tag{4}$$

where $p^{(n)}(x, y)$ is the probability to get from x to y in n steps. P is interpreted as a (countable) stochastic matrix, so that $p^{(n)}(x, y)$ is the $(x, y)-$entry of the matrix power P^n. We set $P^0 = I$, the identity matrix over X.

The transition operator P acts on functions $f: X \to \mathbb{R}$ by

$$Pf(x) := \sum_y p(x, y)f(y). \tag{5}$$

Definition 2.2. The Green function of (X, P) is the power series

$$G(x, y|z) = \sum_{n=0}^{\infty} p^{(n)}(x, y) z^n, \; x, y \in X, \; z \in \mathbb{C}.$$

Remark 2.1. For all $x, y \in X$ the power series $G(x, y|z)$ has the same radius of convergence $1/\rho(P)$.

Definition 2.3. Fix $t > 0$. A $t-$superharmonic function is a function $f: X \to \mathbb{R}$ satisfying

$$Pf \leq tf.$$

We write $S(P, t)$ for the collection of all $t-$superharmonic functions and $S^+(P, t)$ for the positive cone of $S(P, t)$, i.e. $S^+(P, t) = \{f \in S(P, t) : f \geq 0\}$.

A base of the cone $S^+(P, t)$ can be defined with the help of a reference point $x_0 \in X$ by

$$B(P, t) := \{f \in S^+(P, t) : \; f(x_0) = 1\}.$$

Lemma 2.1. *$B(P, t)$ is compact in the topology of pointwise convergence.*

Proof. The closedness of $B(P, t)$ follows from Fatou's lemma. Let $x \in X$, then irreducibility implies the existence of n_x such that $p^{(n_x)}(x_0, x) > 0$. If $f \in B(P, t)$ then

$$p^{(n)}(x_0, x) f(x) \leq P^n f(x_0) \leq t^n f(x_0) = t^n.$$

Hence

$$f(x) \leq \frac{t^{n_x}}{p^{(n_x)}(x_0, x)} \quad \forall f \in B(P, t),$$

and the desired compactness follows. □

Lemma 2.2.

$$\rho(P) = \min\{t > 0 : \; \exists f(\cdot) > 0 \textit{ such that } Pf \leq tf\}$$

Proof. If there exists a function $f \neq 0$ in $S^+(P, t)$, then $p^{(n)}(x, x) f(x) \leq P^n f(x) \leq t^n f(x)$. Hence $\rho(P) = \limsup_n \left(p^{(n)}(x, x)\right)^{1/n} \leq t$. Conversely, for $t > \rho(P)$ the function $f(x) = G(x, x_0|1/t)$ is by Remark 2.1 well-defined. It is clear that $f(\cdot)$ is non-zero and in $S^+(P, t)$. Hence, $B(P, t) \neq \emptyset$. We have $B(P, t_1) \subseteq B(P, t_2)$ for $t_1 < t_2$. By compactness of the sets $B(P, t)$, it follows that $B(P, \rho(P)) = \bigcap_{t > \rho(P)} B(P, t) \neq \emptyset$. □

2.1 Branching Markov Chains

We consider an irreducible Markov chain (X, P) in discrete time. For all $x \in X$ let

$$\mu_1(x), \mu_2(x), \ldots$$

be a sequence of non-negative numbers satisfying

$$\sum_{k=1}^{\infty} \mu_k(x) = 1 \text{ and } m(x) := \sum_{k=1}^{\infty} k \mu_k(x) < \infty.$$

We define the Branching Markov Chain (BMC) on (X, P) following [7]. At time 0 we start with one particle in an arbitrary starting position $x_s \in X$. When a particle is in x, it generates k offspring particles at x with probability $\mu_k(x)$ $(k = 1, 2, \ldots)$ and dies. The k offspring particles then move independently according to the Markov chain (X, P) and generate their offspring as well. At any time, all particles move and branch independently of the other particles and the previous history of the process. The resulting BMC is a Markov chain with countable state space X', namely the space of all particle configurations

$$\omega(n) = \{x_1(n), x_2(n), \ldots, x_{\eta(n)}(n)\},$$

where $x_i(n) \in X$ is the position of the ith particle at time n and $\eta(n)$ is the total number of particles at time n. Since there is always at least one offspring particle, the number of particles is always increasing in time. In most cases under consideration the number of particles $\eta(n)$ tends to infinity as $n \to \infty$ almost surely. Therefore, it is not interesting to ask if a BMC is recurrent as a Markov chain on X': $\eta(n) \to \infty$ implies its transience. It is more reasonable to define transience and recurrence as in Lemma 1.1. With the notations above we can write $\alpha(x)$ as

$$\alpha(x) = \mathbb{P}_x \left(\sum_{n=1}^{\infty} \sum_{i=1}^{\eta(n)} \mathbf{1}_{\{x_i(n)=x\}} = \infty \right),$$

where $\mathbb{P}_x(\cdot) = \mathbb{P}(\cdot | x_s = x)$ and $x \in X$. Note that a BMC in our setting is strongly recurrent $(\alpha \equiv 1)$ if every state $x \in X$ is visited with probability 1. In analogy to [7], we introduce the following modified version of the BMC. We fix an arbitrary position $x_0 \in X$, which we denote the origin of X. After the first time step we conceive the origin as an *absorbing* point: if a particle reaches the origin it stays there forever and stops producing offspring. We denote this new process with BMC*. The process BMC* is analogous to the original process BMC except that $p(x_0, x_0) = 1,\ p(x_0, x) = 0\ \forall x \neq x_0$ and $\mu_1(x_0) = 1$ from the second time step on. Let $\eta_0(n, x_s)$ be the number of particles at position x_0 at time n, given that the BMC* started in $x_s \in X$. We define the random variable $\nu(x_s)$ as

$$\nu(x_s) = \lim_{n \to \infty} \eta_0(n, x_s).$$

The random variable ν takes values in $\{0, 1, 2, \ldots\} \cup \{\infty\}$.

3 Results

We present a sufficient condition for transience of a Branching Markov Chain (BMC), which is inspired by the Lyapunov methods developed in [3] and [7].

Theorem 3.1. *A BMC with irreducible underlying Markov chain (X, P) and $m(y) > 1$ for some $y \in X$ is transient if there exists a strictly positive function $f(\cdot)$ such that*

$$Pf(x) \leq \frac{f(x)}{m(x)} \quad \forall x \in X. \tag{6}$$

Proof. We show that the total number of particles returning to a starting point $x_s = x_0 \neq y$ is finite. The total number of particles in x_0 can be interpreted as the total number of progeny in a branching process $(Z_n)_{n\geq 0}$. We show that this process dies out with probability one. The branching process $(Z_n)_{n\geq 0}$ is defined as follows: Note that each particle has a unique ancestry line which leads back to the starting particle at time 0 at x_0. Let $Z_0 = 1$ and let Z_1 be the number of particles being the first particle in their ancestry line (after the starting particle) to visit x_0. Inductively we define Z_n as the number of particles being the nth particle in their ancestry line to visit x_0. This defines a Galton-Watson process with offspring distribution $Z \stackrel{d}{=} Z_1$. We have that

$$\sum_{n=1}^{\infty} Z_n = \sum_{n=1}^{\infty} \sum_{i=1}^{\eta(n)} \mathbf{1}_{\{x_i(n)=x_0\}}$$

and $Z \stackrel{d}{=} \nu(x_0)$. In order to show that (Z_n) dies out almost sure it suffices to show that $E\nu(x_0) \leq 1$ and $\mathbb{P}_{x_0}(\nu(x_0) < 1) > 0$. Given the first statement the latter is true since $m(y) > 1$ and hence $\mathbb{P}_{x_0}(\nu(x_0) > 1) > 0$. It remains to show the first statement: Consider the corresponding BMC* and define

$$Q(n) := \sum_{i=1}^{\eta(n)} f(x_i(n)),$$

where $x_i(n)$ is the position of the ith particle at time n. One can show that $Q(n)$ is a supermartingale and that

$$\nu(x_s) \leq \frac{Q_\infty}{f(x_s)}.$$

We refer the reader for the technical details to the proof of Theorem 3.2 in [7]. We obtain by taking expectations and starting the BMC* in $x_s = x_0$

$$E\nu(x_0) \leq \frac{EQ_\infty}{f(x_0)} \leq \frac{EQ(0)}{f(x_0)} = \frac{f(x_0)}{f(x_0)} = 1. \tag{7}$$

□

Remark 3.1. In contrast to Theorem 2.2. in [3] and Corollary 3.1 in [7] we demand that the condition (6) holds for all $x \in X$ but don't require that $f(x) \to 0$. Note that in [3] the BMC* is defined in a slightly different way: the origin x_0 is always absorbing.

Remark 3.2. The converse of Theorem 3.1 does not hold in general, for a counterexample see Section 5 in [3].

3.1 BMC with constant mean offspring

We assume that the mean number of offspring is constant, i.e. $m(x) = m > 1$ for all $x \in X$. Note that we do not assume $(\mu_k(x))_k = (\mu_k(y))_k$ for $x, y \in X$, and the BMC therefore needs not to be a Tree-Indexed Markov Chain as in [1].

Under these assumptions, we have the following.

Theorem 3.2. *For a BMC with irreducible underlying Markov chain (X, P) and constant mean offspring $m > 1$, it holds that the BMC is transient if $m \leq 1/\rho(P)$ and recurrent if $m > 1/\rho(P)$.*

Remark 3.3. If $m = \infty$ then the BMC is recurrent, since one can compare the process with a suitable BMC with $\widetilde{m} > 1/\rho(P)$.

Proof. The first part follows from Lemma 2.2 and Theorem 3.1. To show the recurrence we use ideas developed in [1], [3] and [7]: We compare the original BMC with some new process with fewer particles and show that the new process is recurrent. We start the BMC in $x_0 \in X$. We know from the hypothesis and the definition of $\rho(P)$, that there exists a $k = k(x_0)$ such that

$$p^{(k)}(x_0, x_0) > m^{-k}.$$

We observe the BMC only at times $k, 2k, 3k, \ldots$. Particles that are not in x_0 at these times are neglected. Let $\xi(n)$ be the number of the remaining particles in x_0 at time nk. The process $\xi(\cdot)$ is a Galton-Watson process with mean $p^{(k)}(x_0, x_0) \cdot m^k > 1$, thus survives with positive probability and hence the origin is visited infinitely often by the original BMC with positive probability. □

Remark 3.4. Theorem 3.2 implies in particular that Markov chains indexed by Galton-Watson trees are transient in the critical case $m = 1/\rho(P)$, since if $(\mu_k(x))_k = (\mu_k(y))_k$ for all $x, y \in X$ the BMC is a Markov chain indexed by a Galton-Watson tree, compare to [1].

3.2 Quasi-transitive BMC

Let X be a locally finite, connected graph and $Aut(X)$ be the group of automorphisms of X. Let P be the transition matrix of an irreducible Markov chain on X and $Aut(X, P)$ be the group of all $\gamma \in Aut(X)$ which satisfy $p(\gamma x, \gamma y) = p(x, y)$ for all $x, y \in X$. We say the Markov chain (X, P) is transitive, if the group $Aut(X, P)$ acts transitively on X and quasi-transitive if $Aut(X, P)$ acts with finitely many orbits on X, that is if each vertex of X belongs to one of finitely many orbits.

We say a BMC is quasi-transitive if the group $Aut(X, P, \mu)$ of all $\gamma \in Aut(X, P)$ which satisfy $\mu_k(x) = \mu_k(\gamma x)$ $\forall k \geq 1$ for all $x \in X$ acts with finitely many orbits on X. Using induction on n, one can show the following.

Lemma 3.3. *For a quasi-transitive BMC it holds that for all $x, y \in X$ and all $\gamma \in Aut(X, P, \mu)$*

$$\mathbb{P}_x\left(\sum_{i=1}^{\eta(n)} \mathbf{1}\{x_i(n) = y\} = k\right) = \mathbb{P}_{\gamma x}\left(\sum_{i=1}^{\eta(n)} \mathbf{1}\{x_i(n) = \gamma y\} = k\right) \quad \forall n \in \mathbb{N}. \tag{8}$$

For quasi-transitive BMC we have a $0-1-$ law for the return probability. In other words, $\alpha \in \{0, 1\}$ in this case.

Theorem 3.4. *For a quasi-transitive BMC with underlying Markov chain (X, P) and branching distribution $(\mu_k(x))_{k\geq 1}$ with constant mean offspring $m(x) = m > 1$ $\forall x$, it holds that*

- *the BMC is transient ($\alpha \equiv 0$) if $m \leq 1/\rho(P)$.*
- *the BMC is strongly recurrent ($\alpha \equiv 1$) if $m > 1/\rho(P)$.*

Proof. The statement for the case $m \leq 1/\rho(P)$ follows from Theorem 3.2. Recurrence in the case $m > 1/\rho(P)$ also follows from Theorem 3.2. In order to show the strong recurrence ($\alpha \equiv 1$) in the case $m > 1/\rho(P)$, we use ideas from the proof of Theorem 4.3 in [3].

Constructing infinitely many supercritical Galton-Watson processes whose extinction probabilities are bounded away from 1, we show that at least one location is hit infinitely often. We start the BMC in $x_{s_1} \in X$. We know from the hypothesis and the definition of $\rho(P)$, that there exists a $k_1 = k_1(x_{s_1})$ such that

$$p^{(k_1)}(x_{s_1}, x_{s_1}) > m^{-k_1}.$$

We construct a new process $\xi_1(\cdot)$ by observing the BMC only at times $k_1, 2k_1, 3k_1, \ldots$ and by neglecting all the particles not being in position x_{s_1}. Then, $\xi_1(n)$ is the number of particles of the new process in x_{s_1} at time nk_1. In this way, we obtain a Galton-Watson process $\xi_1(\cdot)$. The number of particles in x_{s_1} at time nk_1 of the original BMC is at least $\xi_1(n)$. The process $\xi_1(\cdot)$ is a Galton-Watson process with mean $p^{(k_1)}(x_{s_1}, x_{s_1}) \cdot m^{k_1} > 1$. Hence $\xi_1(\cdot)$ dies out with a probability $q_1 = q_1(x_{s_1}) < 1$. If this first process dies out, we start a second process $\xi_2(\cdot)$, defined in the same way with a starting position x_{s_2} (x_{s_2} can be any location which is occupied by a particle at the time where the first process dies out) and $k_2 = k_2(x_{s_2})$ such that

$$p^{(k_2)}(x_{s_2}, x_{s_2}) > m^{-k_2}.$$

This process dies out with probability $q_2 = q_2(x_{s_2})$. If the second process dies out we construct a third one, and so on. We obtain a sequence of processes $\xi_i(\cdot)$ with extinction probabilities q_i. It suffices now to show that the q_i are bounded away from 1: the probability that all the processes die out is then $\prod_i q_i = 0$. Due to Lemma 3.3 we have that for two starting positions x and y of the same orbit

$$\mathbb{P}_x\left(\sum_{i=1}^{\eta(n)} \mathbf{1}\{x_i(n) = x\} = k\right) = \mathbb{P}_y\left(\sum_{i=1}^{\eta(n)} \mathbf{1}\{x_i(n) = y\} = k\right) \quad \forall n \in \mathbb{N}.$$

Hence two processes started in x and y have the same distributions and hence the same extinctions probabilities. Since there are only finitely many orbits, there are only finitely many different extinction probabilities q_i. □

Remark 3.5. Instead of considering quasi-transitive Markov Chains, we could also assume that $(p^{(l)}(x,x))^{1/l}$ converges uniformly in x, i.e. $\forall\ \varepsilon > 0\ \exists\ l$ such that $(p^{(l)}(x,x))^{1/l} > \rho(P) - \varepsilon, \forall x \in X$, and that there is a $k \in \mathbb{N}$ such that $\inf_x \sum_{i=1}^k i\mu_i(x) \geq 1/\rho(P)$. Observing in the same way as in the proof of Theorem 3.4 the BMC with branching distributions $\tilde{\mu}_0(x) = \sum_{i=k+1}^{\infty} \mu_i(x)$ and $\tilde{\mu}_i(x) = \mu_i(x)$ for $i = 1, \ldots, k$ and $x \in X$, we obtain supercritical Galton-Watson processes ξ_i with bounded variances and means bounded away from 1, since l and k do not depend on x_{s_i}. Hence the extinction probabilities q_i are bounded away from 1.

4 Examples

1. A BMC with transient underlying Markov chain (X, P) is transient if
$$\sup_{x\in X} m(x) \leq 1/\rho(P).$$
2. A branching symmetric random walk on $\mathbb{Z}^d$, $d \in \mathbb{N}$, is strongly recurrent for all branching distributions with constant mean offspring $m > 1$.

3. Consider a random walk on $\mathbb{Z}$ with drift: Let $X = \mathbb{Z}$, $p \in (0,1)$ and P given by

$$p(x, x+1) = p = 1 - p(x, x-1).$$

Take branching distributions with constant mean offspring m. The spectral radius is $\rho(P) = 2 \cdot \sqrt{p(1-p)}$. Hence, the corresponding BMC is transient if

$$m \leq \frac{1}{2 \cdot \sqrt{p(1-p)}}$$

and strongly recurrent if

$$m > \frac{1}{2 \cdot \sqrt{p(1-p)}}.$$

(This reproduces a result of [3] in section 4, noted that there is a calculation error in the formula after Theorem 4.3 of [3] so that the "<" should become a " $\leq$".)

4. More generally, take $X = \mathbb{Z}^d$ and $e_i \in \mathbb{Z}^d$ with $(e_i)_j = \delta_{ij}$ for $i, j \in \{1, \dots, d\},\ d \geq 1$. Let P be defined by

$$p(x, x+e_i) = p_i^+,\ p(x, x-e_i) = p_i^- \quad \text{such that}$$

$$\sum_{i=1}^d p_i^+ + \sum_{i=1}^d p_i^- = 1, \quad \forall x \in \mathbb{Z}^d$$

and such that P is irreducible. Take branching distributions with constant mean offspring m. The spectral radius can be calculated with the help of the Perron-Frobenius Theorem (see for example [8]):

$$\rho(P) = 2 \sum_{i=1}^d \sqrt{p_i^+ p_i^-}.$$

The corresponding BMC is strongly recurrent if

$$m > \frac{1}{2 \sum_{i=1}^d \sqrt{p_i^+ p_i^-}}.$$

Otherwise it is transient.

References

[1] I. Benjamini and Y. Peres, *Markov chains indexed by trees*, The Annals of Probability **22**, no. 1 (1994), 219–243

[2] I. Benjamini and Y. Peres, *Tree-Indexed Random Walks on Groups and First Passage Percolation,* Probab. Theory and Related Fields **98** (1994), 91–112

[3] F. Comets, M. V. Menshikov and S. Yu. Popov, *One-dimensional branching random walk in random environment: a classification*, Markov Processes Relat. Fields **4**(1998), 465–477

[4] F. Comets and S. Popov, *On multidimensional branching random walks in random environment,* arXiv:math.PR/0507126, (2005)

[5] F. P. Machado and S. Yu. Popov, *One dimensional branching random walk in a Markovian random environment*, Journal Appl. Probab. **37**, no. 4 (2000), 1157–1163

[6] F. P. Machado and S. Yu. Popov, *Branching random walk in random environment on trees*, Stochastic Process. Appl. **106**, no. 1 (2003), 95–106

[7] M. V. Menshikov and S. E. Volkov, *Branching Markov Chains: Qualitative Characteristics*, Markov Processes Relat. Fields **3** (1997), 1–18

[8] W. Woess, *Random walks on infinite graphs and groups,* Cambridge University Press, (2000)

Nina Gantert
Institut für Mathematische Statistik
Universität Münster
Einsteinstr. 62
D-48149 Münster
Germany
gantert@math.uni-muenster.de

Sebastian Müller
Institut für Mathematische Statistik
Universität Münster
Einsteinstr. 62
D-48149 Münster
Germany
Sebastian.Mueller@math.uni-muenster.de