CATALYTIC BRANCHING PROCESSES VIA SPINE TECHNIQUES AND RENEWAL THEORY

LEIF DÖRING AND MATTHEW ROBERTS

Abstract. In this article we contribute to the moment analysis of branching processes in catalytic media. The many-to-few lemma based on the spine technique is used to derive a system of (discrete space) partial differential equations for the number of particles in a variation of constants formulation. The long-time behavior is then deduced from renewal theorems and induction.

1. Introduction and Results

A classical subject of probability theory is the analysis of branching processes in discrete or continuous time, going back to the study of extinction of family names by Francis Galton. There have been many contributions to the area since, and we present here an application of a recent development in the probabilistic theory. We identify qualitatively different regimes for the long-time behaviour for moments of sizes of populations in a simple model of a branching Markov process in a catalytic environment.

To give some background for the branching mechanism, we recall the discrete-time Galton-Watson process. Given a random variable $X$ with law $\mu$ taking values in $\mathbb{N}$, the branching mechanism is modelled as follows: for a deterministic or random initial number $Z_0 \in \mathbb{N}$ of particles, one defines for $n = 1, 2, \ldots$

$$Z_{n+1} = \sum_{r=0}^{Z_n} X_r(n),$$

where all $X_r(n)$ are independent and distributed according to $\mu$. Each particle in generation $n$ is thought of as giving birth to a random number of particles according to $\mu$, and these particles together form generation $n + 1$. For the continuous-time analogue each particle carries an independent exponential clock of rate 1 and performs its breeding event once its clock rings.

It is well-known that a crucial quantity appearing in the analysis is $m = \mathbb{E}[X]$, the expected number of offspring particles. The process has positive chance of long-term survival if and only if $m > 1$. This is known as the supercritical case. The cases $m = 1$ (critical) and $m < 1$ (subcritical) also show qualitatively different behaviour in the rate at which the probability of survival decays.

As this paper deals with the moment analysis of a spatial relative to this system, we mention the classical trichotomy for the moment asymptotics of Galton-Watson processes:

$$\lim_{t \to \infty} e^{-k(m-1)t} \mathbb{E}[Z_t^k] \in (0, \infty) \quad \forall k \in \mathbb{N} \quad \text{if } m < 1,$$

$$\lim_{t \to \infty} t^{k-1} \mathbb{E}[Z_t^k] \in (0, \infty) \quad \forall k \in \mathbb{N} \quad \text{if } m = 1,$$

$$\lim_{t \to \infty} e^{-k(m-1)t} \mathbb{E}[Z_t^k] \in (0, \infty) \quad \forall k \in \mathbb{N} \quad \text{if } m > 1,$$

so that all moments increase exponentially to infinity if $m > 1$, increase polynomially if $m = 1$, and decay exponentially fast to zero if $m < 1.$

2000 Mathematics Subject Classification. Primary 60J27; Secondary 60J80.

Key words and phrases. Renewal Theorem, Local Times, Branching Process, Many-to-Few Lemma.

LD acknowledges the support of the Fondation Sciences Mathématiques de Paris. MR thanks ANR MADCOF (grant ANR-08-BLAN-0220-01) and WIAS for their support.
In the present article we are interested in a simple spatial version of the Galton-Watson process for which a system of branching particles moves in space and particles branch only in the presence of a catalyst. More precisely, we start a particle $\xi$ which moves on some countable set $S$ according to a continuous-time Markov process with Q-matrix $A$. This particle carries an exponential clock of rate 1 that only ticks if $\xi$ is at the same site as the catalyst, which we assume sits at some fixed site $0 \in S$. If and when the clock rings, then the particle dies and is replaced in its position by a random number of offspring. This number is distributed according to some offspring distribution $\mu$, and all newly born particles behave as independent copies of their parent: they move on $S$ according to $A$ and branch after an exponential rate 1 amount of time spent at 0.

In recent years several authors have studied such branching systems. Often the first quantities that are analyzed are moments of the form

$$M_k(t, x, y) = \mathbb{E}[N_t(y)^k \mid \xi_0 = x] \quad \text{and} \quad M_k(t, x) = \mathbb{E}[N_t^k \mid \xi_0 = x],$$

where $N_t(y)$ is the number of particles alive at site $y$ at time $t$, and $N_t = \sum_{y \in S} N_t(y)$ is the total number of particles alive at time $t$. Under the additional assumption that $A = \Delta$ is the discrete Laplacian on $\mathbb{Z}^d$, the moment analysis was first carried out in [ABY98], [ABY98b], [AB00] via partial differential equations and Tauberian theorems. More recently, the moment analysis, and moreover the study of conditional limit theorems, was pushed forward to more general spatial movement $A$ assuming

(A1) irreducibility,

(A2) spatial homogeneity,

(A3) symmetry,

(A4) finite variance of jump sizes.

Techniques such as Bellman-Harris branching processes (see [TV03], [VT05], [B10], [B11]), operator theory (see [Y10]), and renewal theory (see [HVT10]) have been applied successfully. Some of these tools also apply in a non-symmetric framework. We present a purely stochastic approach avoiding the assumptions (A1)-(A4). In order to avoid many pathological special cases we only assume

(A) the motion governed by $A$ is irreducible.

This assumption is not necessary, and the interested reader may easily reconstruct the additional cases from our proofs.

In order to analyze the moments $M_k$ one can proceed in two steps. First, a set of partial differential equations for $M_k$ is derived. This can be done for instance as in [AB00] via analytic arguments from partial differential equations for the generating functions $\mathbb{E}_x[e^{-zN_t(y)}]$ and $\mathbb{E}_x[e^{-zN_t}]$ combined with Faà di Bruno’s formula of differentiation. The asymptotic properties of solutions to those differential equations are then analyzed in a second step where more information on the transition probabilities corresponding to $A$ implies more precise results on the asymptotics for $M_k$. This is where the finite variance assumption is used via the local central limit theorem.

The approach presented in this article is based on the combinatorial spine representation of [HR11] to derive sets of partial differential equations, in variation of constants form, for the $k$th moments of $N_t(y)$ and $N_t$. A set of combinatorial factors can be given a direct probabilistic explanation, whereas the same factors appear otherwise from Faà di Bruno’s formula. Those equations are then analyzed via renewal theorems. We have to emphasize that under the assumption (A) only, general precise asymptotic results are of course not possible so that we aim at giving a qualitative description. Compared to the fine results in the presence of a local central limit theorem (such as Lemma 3.1 of [HVT10] for finite variance transitions on $\mathbb{Z}^d$) our qualitative description is rather poor. On the other hand, the generality of our results allows for some interesting applications. For example, one can easily deduce asymptotics for moments of the number of particles when the catalyst is not fixed at zero, but rather follows some Markov process of its own, simply by considering the difference walk.
To state our main result, we denote the transition probabilities of \( A \) by \( p_t(x, y) = \mathbb{P}_x(\xi_t = y) \) and the Green function by

\[
G_\infty(x, y) = \int_0^\infty p_t(x, y) \, dt.
\]

Recall that, by irreducibility, the Green function is finite for all \( x, y \in S \) if and only if \( A \) is transient. For the statement of the result let us further denote by

\[
L_t(y) = \int_0^t \mathbb{1}_{\{\xi_s = y\}} \, ds
\]

the time of \( \xi \) spent at site \( y \) up to time \( t \).

**Theorem 1.** Suppose that \( \mu \) has finite moments of all orders; then the following regimes occur for all integers \( k \geq 1 \):

i) If the branching mechanism is **subcritical**, then

\[
\begin{align*}
\lim_{t \to \infty} M^k(t, x) & \in (0, \infty) \quad \text{if } A \text{ is transient,} \\
\lim_{t \to \infty} M^k(t, x) & = 0 \quad \text{if } A \text{ is recurrent,}
\end{align*}
\]

and

\[
\lim_{t \to \infty} M^k(t, x, y) = 0 \quad \text{in all cases.}
\]

ii) If the branching mechanism is **critical**, then

\[
\lim_{t \to \infty} \mathbb{E}_x[L_t(0)^{k-1}] \in (0, \infty) \quad \text{and} \quad M^1(t, x, y) = p_t(x, y).
\]

iii) If the branching mechanism is **supercritical**, then there is a critical constant

\[
\beta = \frac{1}{G_\infty(0, 0)} + 1 \geq 1
\]

such that

a) for \( m < \beta \)

\[
\lim_{t \to \infty} M^1(t, x) \in (0, \infty) \quad \text{and} \quad \lim_{t \to \infty} M^1(t, x, y) = 0;
\]

further, there exist constants \( c \) and \( C \) such that

\[
c\mathbb{E}_x[L_t(0)^{k-1}] \leq M^k(t, x) \leq C t^{k-1}.
\]

b) for \( m = \beta \)

\[
\lim_{t \to \infty} M^k(t, x) = \infty,
\]

and

\[
\begin{align*}
\lim_{t \to \infty} M^k(t, x, y) & = \infty \quad \text{if} \quad \int_0^\infty r p_r(0, 0) \, dr = \infty \\
\lim_{t \to \infty} \frac{M^k(t, x, y)}{t^{k-1}} & \in (0, \infty) \quad \text{if} \quad \int_0^\infty r p_r(0, 0) \, dr < \infty.
\end{align*}
\]

(In both cases the growth is subexponential.)

c) for \( m > \beta \)

\[
\lim_{t \to \infty} e^{-kr(m)t} M^k(t, x, y) \in (0, \infty) \quad \text{and} \quad \lim_{t \to \infty} e^{-kr(m)t} M^k(t, x) \in (0, \infty)
\]

where \( r(m) \) equals the unique solution \( \lambda \) to \( \int_0^\infty e^{-\lambda t} p_t(0, 0) \, dt = \frac{1}{m-1} \).
We did not state all the asymptotics in cases ii) and iii)a). Our methods, see Lemma 3, do allow for investigation of these cases too; in particular they show how $M^k(t, x, y)$ for $i < k$. However, without further knowledge of the underlying motion, it is not possible to give any useful and general information. If more information on the tail of $p_t(x, y)$ is available then the recursive equations can indeed be analyzed: for instance for kernels on $\mathbb{Z}^d$ with second moments the local central limit theorem can be applied leading to $p_t(x, y) \sim Ct^{-d/2}$, and such cases have already been addressed by other authors.

The formulation of the theorem does not include the limiting constants. Indeed, the proofs give some of those (in an explicit form involving the transition probabilities $p_t$) in the supercritical regime but they seem to be of little use. The use of spectral theory for symmetric $Q$-matrices $A$ allows one to derive the exponential growth rate $r(m)$ as the maximal eigenvalue of a Schrödinger operator with one-point potential and the appearing constants via the eigenfunctions. Our renewal theorem based proof gives the representation of $r(m)$ as the inverse of the Laplace transform of $p_t(0,0)$ at $1/(m-1)$ and the eigenfunction expressed via integrals of $p_t(0,0)$. As $p_t(0,0)$ is rarely known explicitly, the integral form of the constants is not very useful (apart from the trivial case of Example 1 below). Only in case iii) b) for $\int_0^\infty rp_t(0,0)\,dr = \infty$ are the proofs unable to give strong asymptotics. This is caused by the use of an infinite-mean renewal theorem which only gives asymptotic bounds up to an unknown factor between 1 and 2. There is basically one example in which $p_t(0,0)$ is trivially known:

**Example 1:** For the trivial motion $A = 0$, i.e. branching particles are fixed at the same site as the catalyst, the supercritical cases iii) a) and b) do not occur as $A$ is trivially recurrent so that $\beta = 1$. Furthermore, in this example $p_t(0,0) = 1$ for all $t \geq 0$ so that $r(m) = m - 1$. In fact by examining the proof of Theorem 1 one recovers (1.1,1.2,1.3) with all constants.

The explicit representation for the exponential growth rate allows for a more careful comparison with the non-spatial case.

**Corollary 1.** Let $r(m)$ be the exponential growth rate obtained in the supercritical case of Theorem 1. Then

$$m \mapsto r(m)$$

is convex, with

$$r(1) = 0, \quad r(m) \leq m - 1, \quad \lim_{m \to \infty} \frac{r(m)}{m - 1} = 1.$$  

**Proof.** This follows from elementary manipulations of the defining equation for $r(m)$.  

**Remark 1.** We reiterate here that our results can be generalized when the fixed branching source is replaced by a random branching source moving according to a random walk independent of the branching particles. For the proofs the branching particles only have to be replaced by branching particles relative to the branching source.

2. Proofs

The key tool in our proofs will be the many-to-few lemma proved in [HR11] which relies on modern spine techniques. These emerged from work of Kurtz, Lyons, Pemantle and Peres in the mid-1990s KLPP97, L97, LPP95. The idea is that to understand certain functionals of branching processes, it is enough to carefully study the behaviour of one special particle, the spine. In particular very general many-to-one lemmas emerged, allowing one to easily calculate expectations of sums over particles like

$$\mathbb{E}\left[\sum_{v \in N_t} f(v)\right],$$

where $f(v)$ is some well-behaved functional of the behaviour of the particle $v$ up to time $t$, and $N_t$ here is viewed as the set of particles alive at time $t$, rather than the number. It will always be clear from the context which meaning for $N_t$ is intended.
It is natural to ask whether similar results exist for higher moments of sums over $N_t$. This is the idea behind [HR11], wherein it turns out that to understand the $k$th moment one must consider a system of $k$ particles. The $k$ particles introduce complications compared to the single particle required for first moments, but this is still significantly simpler than controlling the behaviour of the potentially huge random number of particles in $N_t$.

While we do not need to understand the full spine setup here, we shall require some explanation. For each $k \geq 0$ let $p_k = P(X = k)$ and $m_k = E[X^k]$, the $k$th moment of the offspring distribution (in particular $m_1 = m$). We define a new measure $Q = Q^k$, under which there are $k$ distinguished lines of descent known as spines. The construction of $Q$ relies on a carefully chosen change of measure, but we do not need to understand the full construction and instead refer to [HR11]. In order to use the technique, we simply have to understand the dynamics of the system under $Q$.

Under $Q^k$, particles behave as follows:

- We begin with one particle at position $x$ which (as well as its position) has a mark $k$. We think of a particle with mark $j$ as carrying $j$ spines.
- Whenever a particle with mark $j$, $j \geq 1$, spends an (independent) exponential time with parameter $m_j$ in the same position as the catalyst, it dies and is replaced by a random number of new particles with law $A_j$.
- The probability of the event $\{A_j = a\}$ is $a^j p_a m_j^{-1}$. (This is the $j$th size-biased distribution relative to $\mu$.)
- Given that $a$ particles $v_1, \ldots, v_a$ are born, the $j$ spines each choose a particle to follow independently and uniformly at random. Thus particle $v_i$ has mark $l$ with probability $a^{-l}(1-a^{-1})^{j-l}$, $l = 0, \ldots, j$, $i = 1, \ldots, a$. We also note that this means that there are always $k$ spines amongst the particles alive; equivalently the sum of the marks over all particles alive always equals $k$.
- Particles with mark 0 are no longer of interest (in fact they behave just as under $P$, branching at rate 1 when in the same position as the catalyst and giving birth to numbers of particles with law $\mu$, but we will not need to use this).

For a particle $v$, we let $X_v(t)$ be its position at time $t$ and $B_v$ be its mark (the number of spines it is carrying). Let $\sigma_v$ be the time of its birth and $\tau_v$ the time of its death, and define $\sigma_v(t) = \sigma_v \land t$ and $\tau_v(t) = \tau_v \land t$. Let $\chi^i_t$ be the current position of the $i$th spine. We call the collection of particles that have carried at least one spine up to time $t$ the skeleton at time $t$, and write $\text{ske}(t)$. Figure 1 gives an impression of the skeleton at the start of the process.

![Figure 1](image)

**Figure 1.** An impression of the start of the process: each particle in the skeleton is a different colour, and particles not in the skeleton are drawn in pale grey. The circles show the number of spines being carried by each particle in the skeleton.

A much more general form of the following lemma was proved in [HR11].
Lemma 1 (Many-to-few). Suppose that $f : \mathbb{R} \to \mathbb{R}$ is measurable. Then, for any $k \geq 1$,
\[
\mathbb{E} \left[ \sum_{v_1, \ldots, v_k \in N_t} f(X_{v_1}(t)) \cdots f(X_{v_k}(t)) \right] = Q^k \left[ f(\chi_t^1) \cdots f(\chi_t^k) \prod_{v \in \text{select}(t)} \exp \left( (m_{B_v} - 1) \int_{\sigma_v(t)}^{\tau_v(t)} 1_{0}(X_v(s)) ds \right) \right].
\]

Clearly if we take $f \equiv 1$, then the left hand side is simply the $k$th moment of the number of particles alive at time $t$. The lemma is useful since the right-hand side depends on at most $k$ particles at a time, rather than the arbitrarily large random number of particles on the left-hand side.

Having introduced the spine technique, we can now proceed with the proof of Theorem 1. We first use Lemma 1 for the case $k = 1$, which is simply the many-to-one lemma, to deduce two convenient representations for the first moments: a Feynman-Kac expression and a variation of constants formula. Indeed, the exponential expression equally works for other random potentials and, hence, is well known for instance in the parabolic Anderson model literature. More interestingly, the variation of constants representation is most useful in the case of a one-point potential: it simplifies to a renewal type equation. Understanding when those are proper renewal equations replaces the spectral theoretic arguments of [ABY98] and explains the different cases appearing in Theorem 1.

Lemma 2. The first moments can be expressed as
\begin{align*}
(2.1) \quad M^1(t, x) &= \mathbb{E}_x \left[ e^{(m-1) \int_0^t 1_{0}(\xi_r) \, dr} \right], \\
(2.2) \quad M^1(t, x, y) &= \mathbb{E}_x \left[ e^{(m-1) \int_0^t 1_{0}(\xi_r) \, dr} 1_y(\xi_t) \right],
\end{align*}
where $\xi_t$ is a single particle moving with $Q$-matrix $A$. Furthermore, these quantities fulfill
\begin{align*}
(2.3) \quad M^1(t, x) &= 1 + (m-1)p_t(x, 0) \ast M^1(t, 0), \\
(2.4) \quad M^1(t, x, y) &= p_t(x, y) + (m-1)p_t(x, 0) \ast M^1(t, 0, y),
\end{align*}
where $\ast$ denotes ordinary convolution in $t$.

For completeness we include a proof of these well-known relations. First let us briefly mention why the renewal type equations occur naturally. The Feynman-Kac representation can be proved in various ways; we derive it simply from the many-to-few lemma. The Feynman-Kac formula then leads naturally to solutions of discrete-space heat equations with one-point potential:
\[
\begin{aligned}
\frac{\partial}{\partial t} u(t, x) &= Au(t, x) + (m-1)1_0(x)u(t, x) \\
u(0, x) &= 1_y(x).
\end{aligned}
\]

Applying the variation of constants formula for solutions gives
\[
u(t, x) = P_t u(0, x) + \int_0^t P_{t-s}(m-1)1_0(x)u(s, x) \, ds
\]
\[
= p_t(x, y) + (m-1) \int_0^t p_{t-s}(x, 0)u(s, x) \, ds,
\]
where $P_t$ is the semigroup corresponding to $A$, i.e. $P_t f(x) = \mathbb{E}_x [f(\xi_t)]$.

Proof of Lemma 2. To prove (2.1) and (2.2) we apply the easiest case of Lemma 1. Choose $f(x) = 1_y(\xi_t)$.

which is what we claimed but with expectations taken under $\mathbb{Q}$ rather than the original measure $\mathbb{P}$. However we note that the motion of the single spine is the same (it has $Q$-matrix $A$) under both $\mathbb{P}$ and $\mathbb{Q}$, so we may simply replace $\mathbb{Q}$ with $\mathbb{P}$, giving (2.1) and (2.2).

The variation of constants formulas can now be derived from the Feynman-Kac formulas. We only prove the second identity, as the first can be proved similarly. We use the exponential series to get

$$
\mathbb{E}_x \left[ e^{(m-1) \int_0^t 1_0(\xi_r) \, dr} \right] y(\xi_t) \\
= \mathbb{E}_x \left[ \sum_{n=0}^{\infty} \frac{(m-1)^n}{n!} \left( \int_0^t 1_0(\xi_r) \, dr \right)^n y(\xi_t) \right] \\
= \mathbb{P}_x(\xi_t = y) + \mathbb{E}_x \left[ \sum_{n=1}^{\infty} \frac{(m-1)^n}{n!} \int_0^t \int_0^t \cdots \int_0^t 1_0(\xi_{r_1}) \cdots 1_0(\xi_{r_n}) \, dr_n \cdots dr_1 y(\xi_t) \right] \\
= p_t(x, y) + \mathbb{E}_x \left[ \sum_{n=1}^{\infty} \frac{(m-1)^n}{n!} \int_0^t \int_0^t \cdots \int_0^t 1_0(\xi_{r_1}) \cdots 1_0(\xi_{r_n}) \, dr_n \cdots dr_1 y(\xi_t) \right].
$$

The last step is justified by the fact that the function that is integrated is symmetric in all arguments and, thus, it suffices to integrate over a simplex. We can exchange sum and expectation and obtain that the last expression equals

$$
p_t(x, y) + (m-1) \int_0^t \int_0^t \cdots \int_0^t \mathbb{E}_x[\xi_{r_1} = 0, \ldots, \xi_{r_n} = 0] \, dr_n \cdots dr_1.
$$

Due to the Markov property, the last expression equals

$$
p_t(x, y) + (m-1) \int_0^t p_1(x, 0) \sum_{n=1}^{\infty} \frac{(m-1)^{n-1}}{n!} \int_0^t \int_0^t \cdots \int_0^t \mathbb{P}_0[\xi_{r_2-r_1} = 0, \ldots, \xi_{r_{n-1}} = 0] \, dr_n \cdots dr_1
$$

and can be rewritten as

$$
p_t(x, y) + (m-1) \int_0^t p_1(x, 0) \left( \sum_{n=1}^{\infty} \frac{(m-1)^{n-1}}{n!} \int_0^{t-r_1} \cdots \int_0^{t-r_{n-1}} \mathbb{P}_0[\xi_{r_2} = 0, \ldots, \xi_{r_n} = 0] \, dr_n \cdots dr_1 \right) \, dr_1.
$$

Using the same line of arguments backwards for the term in parentheses, the assertion follows. □

Having derived variation of constants formulas, there are different ways to analyze the asymptotics of the first moments. Assuming more regularity for the transition probabilities, this can be done as sketched in the next remark.

**Remark 2.** Taking Laplace transforms $\mathcal{L}$ in $t$, one can transform (2.3), and similarly (2.4), into the algebraic equation

$$
\mathcal{L}M^1(\lambda, x) = \frac{1}{\lambda} + (m-1)\mathcal{L}M^1(\lambda, 0)\mathcal{L}p_\lambda(x, 0), \quad \lambda > 0,
$$

which can be solved explicitly to obtain

$$
\mathcal{L}M^1(\lambda, x) = \frac{1}{\lambda(1 - (m-1)\mathcal{L}p_\lambda(x, 0))}, \quad \lambda > 0.
$$

Assuming the asymptotics of $p_\lambda(x, 0)$ are known for $t$ tending to infinity (and are sufficiently regular), the asymptotics of $\mathcal{L}p_\lambda(x, 0)$ for $\lambda$ tending to zero can be deduced from Tauberian theorems. Hence, from Equation (2.5) one can then deduce the asymptotics of $\mathcal{L}M^1(\lambda, x)$ as $\lambda$ tends to zero.

This, using Tauberian theorems in the reverse direction, allows one to deduce the asymptotics of $M^1(t, x)$ for $t$ tending to infinity.

Unfortunately, to make this approach work, ultimate monotonicity and asymptotics of the type $p_\lambda(x, 0) \sim Ct^{-\alpha}$ are needed. This motivated the authors of [ABY98] to assume (A4) so that by the local central limit theorem

$$
p_\lambda(x, 0) \sim \left( \frac{d}{2\pi} \right)^{d/2} t^{-d/2}.
$$
As we did not assume any regularity for \( p_t \), the aforementioned approach fails in general. We instead use an approach based on renewal theorems recently seen in [DS10].

**Proof of Theorem 2** for \( M^1 \). Taking into account irreducibility and the Markov property of \( \mathcal{A} \), we see that the property “\( \int_0^\infty \mathbb{1}_0(\xi_t)\,dr = \infty \) almost surely” does not depend on the starting value \( \xi_0 \). To prove case i), we simply apply dominated convergence to (2.1) and (2.2). If \( \mathcal{A} \) is transient, then \( \int_0^\infty \mathbb{1}_0(\xi_t)\,dr < \infty \) almost surely and \( M^1(t,x) \) converges to a constant. On the other hand if \( \mathcal{A} \) is recurrent, then \( \int_0^\infty \mathbb{1}_0(\xi_t)\,dr = \infty \) almost surely and \( M^1(t,x) \to 0 \). In both cases \( M^1(t,x) \to 0 \), because if \( \mathcal{A} \) is transient then \( \mathbb{1}_{\{\xi_t=y\}} \to 0 \) almost surely, and if \( \mathcal{A} \) is recurrent then \( M^1(t,x,y) \leq M^1(t,x) \to 0 \).

Regime ii) is trivial as here \( M^1(t,x) = 1 \) and \( M^1(t,x,y) = p_t(x,y) \). Next, for regime iii) a) we exploit both the standard and the reverse Hölder inequality for \( p > 1 \):

\[
\begin{align*}
\tag{2.6} M^1(t,x,y) &\geq \mathbb{E}_x\left[ e^{-(1/(p-1))(m-1)\int_0^\infty \mathbb{1}_0(\xi_t)\,dr} \right]^{-1}\left(\frac{1}{p}ight)p_t(x,y)^p, \\
\tag{2.7} M^1(t,x,y) &\leq \mathbb{E}_x\left[ e^{p(m-1)\int_0^\infty \mathbb{1}_0(\xi_t)\,dr} \right]^{1/p}\left(\frac{1}{p}ight)p_t(x,y)^{(p-1)/p}.
\end{align*}
\]

In the recurrent case \( G_\infty(0,0) = \infty \) and thus \( \beta = 1 \), so this case has already been dealt with in regime ii). Hence we may assume that \( \mathcal{A} \) is transient so that \( \int_0^\infty \mathbb{1}_0(\xi_t)\,dr < \infty \) with positive probability. This shows that the expectation in the lower bound (2.6) converges to a finite constant. By assumption \( m-1 < \beta \) so that there is \( p > 1 \) satisfying \( p(m-1) < \beta \). With this choice of \( p \), part 3) of Theorem 1 of [DS10] implies that also the expectation in the upper bound (2.7) converges to a finite constant. In total this shows that

\[
Cp_t(x,y)^p \leq M^1(t,x,y) \leq C'p_t(x,y)^{(p-1)/p}
\]

and the claim for \( M^1(t,x,y) \) follows. For \( M^1(t,x) \) we can directly refer to Theorem 1 of [DS10].

For regimes iii) b) and c) we give arguments based on renewal theorems. A closer look at the variation of constants formula (2.4) shows that only for \( x = 0 \), \( M^1(t,x,y) \) occurs on both sides of the equation. Hence, we start with the case \( x = 0 \) and afterwards deduce the asymptotics for \( x \neq 0 \).

Let us begin with the simpler case iii) c). As mentioned above, in this case we may assume that \( \mathcal{A} \) is transient so that \( \int_0^\infty p_t(0,0)\,dr < \infty \). Hence, dominated convergence ensures that the equation \( \int_0^\infty e^{-\lambda t}p_t(0,0)\,dt = 1/(m-1) \) has a unique positive root \( \lambda \), which we call \( r(m) \). The definition of \( r(m) \) shows that \( U(dt) := (m-1)e^{-r(m)t}p_t(0,0)\,dt \) is a probability measure on \([0,\infty)\) and furthermore \( e^{-r(m)t}p_t(0,0) \) is directly Riemann integrable. Hence the classical renewal theorem (see page 349 of [F71]) can be applied to the (complete) renewal equation

\[
f(t) = g(t) + f \ast U(t),
\]

with \( f(t) = e^{-r(m)t}M^1(t,0,y) \) and \( g(t) = e^{-r(m)t}p_t(0,y) \). The renewal theorem implies that

\[
\lim_{t \to \infty} f(t) = \frac{\int_0^\infty g(s)\,ds}{\int_0^\infty U((s,\infty))\,ds} \in (0,\infty)
\]

so that the claim for \( M^1(t,0,y) \) follows including the limiting constants.

For iii) b), we need to be more careful as the criticality implies that \( (m-1)\int_0^\infty p_t(0,0)\,dr = 1 \). Hence, the measure \( U \) as defined above is already a probability measure so that the variation of constants formula is indeed a proper renewal equation. The renewal measure \( U \) only has finite mean if additionally

\[
\int_0^\infty rp_t(0,0)\,dr < \infty.
\]

In the case of finite mean the claim follows as above from (2.8) without the exponential correction (i.e. \( r(m) = 0 \)). Note that \( p_t(0,y) \) is directly Riemann integrable as the case \( \beta > 0 \) implies that \( \mathcal{A} \) is transient and \( p_t(0,y) \) is decreasing.
If (2.9) fails, we need a renewal theorem for infinite mean variables. Iterating Equation (2.4) reveals the representation
\[(2.10)\]
\[M^1(t,0,y) = p_t(0,y) \ast \sum_{n \geq 0} (m-1)^n p_t(0,0)^n,\]
where \(\ast^n\) denotes \(n\)-fold convolution in \(t\) and \(p_t(0,y) \ast p_t(0,0)^0 = p_t(0,y)\). Note that convergence of the series is justified by
\[(m-1)^n p_t(0,0)^n \leq \left((m-1) \int_0^t p_r(0,0) \, dr\right)^n\]
and the assumption on \(m\). Lemma 1 of [273] now implies that
\[(2.11)\]
\[\sum_{n \geq 0} (m-1)^n p_t(0,0)^n \approx t \int_0^t \int_0^s p_r(0,0) \, dr \, ds\]
which tends to infinity as \((m-1) \int_0^s p_r(0,0) \, dr \to 0\) for \(s \to \infty\) since we assumed that \((m-1)p_r(0,0)\) is a probability density in \(r\). To derive from this observation the result for \(M^1(t,0,y)\), note that the simple bound \(p_t(0,y) \leq 1\) gives the upper bound
\[(2.12)\]
\[M^1(t,0,y) \leq \int_0^t \sum_{n \geq 0} (m-1)^n p_r(0,0)^n \, dr.\]
For a lower bound, we use that due to irreducibility and continuity of \(p_t(0,y)\) in \(t\), there are \(0 < t_0 < t_1\) and \(\epsilon > 0\) such that \(p_t(0,y) > \epsilon\) for \(t_0 \leq t \leq t_1\). This shows that
\[(2.13)\]
\[M^1(t,0,y) \geq \epsilon \int_{t-t_0}^{t-t_1} \sum_{n \geq 0} (m-1)^n p_r(0,0)^n \, dr.\]
Combined with (2.11), the lower and upper bounds directly prove the claim for \(M^1(t,0,y)\).

It remains to deal with regime iii) b) and c) for \(x \neq 0\). The results follow from the asymptotics of the convolutions as those do not vanish at infinity. But this can be deduced from simple upper and lower bounds similar to (2.12) and (2.13).

The asymptotic results for the expected total number of particles \(M^1(t,x)\) follow from similar ideas: estimating as before
\[1 + \epsilon \int_{t-t_0}^{t-t_1} M^1(r,0) \, dr \leq M^1(t,x) \leq 1 + \int_0^t M^1(r,0) \, dr,\]
and applying case 2) of Theorem 1 of [DS10] to (2.1) with \(x = 0\), the result follows.

We now come to the crucial lemma of our paper. We use the the many-to-few lemma to reduce higher moments of \(N_t\) and \(N_t(y)\) to the first moment. More precisely, a system of equations is derived that can be solved inductively once the first moment is known. This particular useful form is caused by the one-point catalyst. A similar system can be derived in the same manner in the deterministic case if the one-point potential is replaced by a \(n\)-point potential. However the case of a random \(n\)-point potential is much more delicate as the sources are “attracted” to the particles, destroying any chance of a renewal theory approach.

**Lemma 3.** For \(k \geq 2\) the \(k\)th moments fulfill
\[(2.14)\]
\[M^k(t,x) = M^1(t,x) + M^1(t,x,0) \ast g_k\left(M^1(t,0), \ldots, M^{k-1}(t,0)\right),\]
\[(2.15)\]
\[M^k(t,x,y) = M^1(t,x,y) + M^1(t,x,0) \ast g_k\left(M^1(t,0,y), \ldots, M^{k-1}(t,0,y)\right),\]
where
\[g_k(M^1, \ldots, M^{k-1}) = \sum_{j=2}^{k} \mathbb{E} \left[\left(\begin{array}{c} X \\ j \end{array}\right)\right] \sum_{i_1, \ldots, i_j > 0} \frac{k!}{i_1! \cdots i_j!} M^{i_1} \cdots M^{i_j}.\]
Proof. We shall only prove equation \((2.14)\); the proof of equation \((2.15)\) is almost identical. We recall the spine setup and introduce some more notation. To begin with, all \(k\) spines are carried by the same particle \(\xi\) which branches at rate \(m_k = \text{E}[X_k]\) when at 0. Thus the \(k\) spines separate into two or more particles at rate \(m_k - m\) when at 0 (since it is possible that at a birth event all \(k\) spines continue to follow the same particle, which happens at rate \(m\)). We consider what happens at this first “separation” time, and call it \(T\).

Let \(i_1, \ldots, i_j > 0\), \(i_1 + \ldots + i_j = k\), and define \(A_k(j; i_1, \ldots, i_j)\) to be the event that at a separation event, \(i_1\) spines follow one particle, \(i_2\) follow another, \(\ldots\), and \(i_j\) follow another. The first particle splits into a new particles with probability \(a^k p_{m_k} m_k^{-1}\) (see the definition of \(Q^k\)). Then given that the first particle splits into a new particles, the probability that \(i_1\) spines follow one particle, \(i_2\) follow another, \(\ldots\), and \(i_j\) follow another is

\[
\frac{1}{a^k} \binom{a}{j} \frac{k!}{i_1! \cdots i_j!}
\]

(the first factor is the probability of each spine making a particular choice from the \(a\) available; the second is the number of ways of choosing the \(j\) particles to assign the spines to; and the third is the number of ways of rearranging the spines amongst those \(j\) particles). Thus the probability of the event \(A_k(j; i_1, \ldots, i_j)\) under \(Q^k\) is

\[
\frac{1}{m_k} \text{E}\left[\left(\frac{X}{j}\right)\right] \frac{k!}{i_1! \cdots i_j!}.
\]

(Note that, as expected, this means that the total rate at which a separation event occurs is

\[
m_k \cdot \frac{1}{m_k} \sum_{j=2}^{k} \text{E}\left[\left(\frac{X}{j}\right)\right] \frac{k!}{i_1! \cdots i_j!} = m_k - m
\]

since the double sum is just the expected number of ways of assigning \(k\) things to \(X\) boxes without assigning them all to the same box.)

However, for \(j \geq 2\), given that we have a separation event, \(A_k(j; i_1, \ldots, i_j)\) occurs with probability

\[
\frac{1}{m_k} \text{E}\left[\left(\frac{X}{j}\right)\right] \frac{k!}{i_1! \cdots i_j!} \left(\frac{m_k}{m_k - m}\right).
\]

Write \(\chi_t\) for the position of the particle carrying the \(k\) spines for \(t \in [0, T]\), and define \(\mathcal{F}_t\) to be the filtration containing all information (including about the spines) up to time \(t\). Recall that the skeleton \(\text{skel}(t)\) is the tree generated by particles containing at least one spine up to time \(t\); let \(\text{skel}(s; t)\) similarly be the part of the skeleton falling between times \(s\) and \(t\). Using the many-to-few lemma with \(f = 1\), the fact that by definition before \(T\) all spines sit on the same particle and integrating out \(T\), we obtain

\[
\text{E}[X^k] = Q^k \left[ \lim_{\tau \to 0} e^{(m_k - 1) \int_{u}^{\tau} f_{\text{skel}(u)}(X_u(s))ds} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]

\[
= \int_{0}^{t} Q^k \left[ \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{u} f_{\text{skel}(v; t)}(X_v(s))ds} \mathcal{F}_u; T = u \right] du
\]

\[
+ Q^k \left[ e^{(m_k - 1) \int_{0}^{t} 1_{\{T > t\}} e^{-(m_k - m) \int_{0}^{T} 1_{\{T > t\}}} ds} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]

\[
= Q^k \left[ e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \prod_{v \in \text{skel}(u; t)} e^{(m_k - 1) \int_{0}^{T} 1_{\{T \leq t\}} \mathcal{F}_t} \right]
\]
To prove equation (2.15), the same arguments are used with \( f = 1 \) in place of \( f = 1 \). Now we split the sample space according to the distribution of the numbers of spines in the skeleton at time \( T \). Since, given their positions and marks at time \( T \), the particles in the skeleton behave independently, we may split the product up into \( j \) independent factors. Thus

\[
\mathbb{E}[N^t] = \int_0^t \sum_{j=2}^k \sum_{i_1, \ldots, i_j > 0} \mathbb{E} \left[ \left( \frac{X}{j} \right) \right] \frac{k!}{i_1! \cdots i_j!} Q^k \left[ e^{(m-1) f_1} 1_{\mathbb{R}^+} ds \right] \left[ e^{(m-1) f_1} 1_{\mathbb{R}^+} ds \right] \]

where we have used the many-to-few lemma backwards with \( f = 1 \). For Equation (2.15) we again use \( f = 1 \) in place of \( f = 1 \) and copy the same lines of arguments. \( \square \)

**Remark 3.** The factors appearing in \( g_k \) are derived combinatorially from splitting the spines. In Lemma 3.1 of [AB00] they appeared from Faà di Bruno’s differentiation formula.

We need the following elementary lemma before we can complete our proof.

**Lemma 4.** For any non-negative integer-valued random variable \( Y \), and any integers \( a \geq b \geq 1 \),

\[
\mathbb{E}[Y^a] \mathbb{E}[Y] \geq \mathbb{E}[Y^b] \mathbb{E}[Y^{a-b+1}].
\]

**Proof.** Assume without loss of generality that \( b \geq a/2 \). Note that for any two positive integers \( j \) and \( k \),

\[
j^a k + j^b k^{a-b+1} - j^{a-b+1} k^b = j k(j - k)^2 \left( j^{a-3} + 2 j^{a-4} k + 3 j^{a-5} k^2 + \ldots \right.
\]

\[
+ (a - b) j^{b-2} k^{a-b-1} + (a - b - 1) j^{b-3} k^{a-b} + \ldots
\]

\[
+ 2 j k^{a-4} + k^{a-3} \right) \geq 0.
\]

Thus

\[
\mathbb{E}[Y^a] \mathbb{E}[Y] - \mathbb{E}[Y^b] \mathbb{E}[Y^{a-b+1}]
\]

\[
= \sum_{j \geq 1} j^a P(Y = j) \sum_{k \geq 1} k P(Y = k) - \sum_{j \geq 1} j^b P(Y = j) \sum_{k \geq 1} k^{a-b+1} P(Y = k)
\]

\[
= \sum_{j \geq 1} \sum_{k > j} (j^a k + j^b k^{a-b} - j^{a-b+1} k^b) P(Y = j) P(Y = k)
\]

\[
\geq 0
\]

as required. \( \square \)

We can now finish the proof of the main result.

**Proof of Theorem 1 for \( M^k \).** Case i) follows just as for \( M^1 \), applying dominated convergence to the \( Q^k \)-expectation in Lemma 1. Note that if \( T \) is the first split time of the \( k \) spines (as in Lemma 3) then \( e^{(m_k-1) f_1(\zeta_j) ds} \) is stochastically dominated by \( e^{(m_{k-1}) \tau} \) where \( \tau \) is an exponential random variable of parameter \( m_k - m_1 \); this allows us to construct the required dominating random variable.
For case ii), using Lemmas 2 and 3 we find the lower bound

\begin{equation}
M^k(t, x) \geq 1 + C \int_0^t p_s(x, 0) M^{k-1}(t - s, 0) ds \geq C \int_0^t p_s(x, 0) M^{k-1}(t - s, 0) ds.
\end{equation}

An upper bound can be obtained by additionally using Lemma 4 (to reduce \( g_k \) to the leading term \( M^1 M^{k-1} \)) to obtain

\begin{equation}
M^k(t, x) \leq 1 + C \int_0^t p_s(x, 0) M^{k-1}(t - s, 0) ds.
\end{equation}

Using inductively the lower bound (2.16) and furthermore the iteration

\begin{align*}
&\int_0^t \mathbb{P}_x(X_{s_1} = 0) \mathbb{P}_0(X_{s_2} = 0) \cdots \int_0^{t-s_1-\cdots-s_{k-2}} \mathbb{P}_0(X_{s_{k-1}} = 0) ds_{k-1} \cdots ds_2 ds_1 \\
&= \int_0^t \int_0^{t-s_1} \cdots \int_0^{t-s_{k-2}} \mathbb{P}_x(X_{s_1} = 0, X_{s_2} = 0, \ldots, X_{s_{k-1}} = 0) ds_{k-1} \cdots ds_2 ds_1 \\
&= \frac{1}{(k-1)!} \int_0^t \int_0^{t-s_1} \cdots \int_0^{t-s_{k-2}} \mathbb{P}_x(X_{s_1} = 0, X_{s_2} = 0, \ldots, X_{s_{k-1}} = 0) ds_{k-1} \cdots ds_2 ds_1 \\
&= \frac{1}{(k-1)!} \mathbb{E}_x \left[ \left( \int_0^t 1_{\{X_s = 0\}} ds \right)^{k-1} \right] \\
&= \frac{1}{(k-1)!} \mathbb{E}_x \left[ L_t(0)^{k-1} \right]
\end{align*}

we see that \( M^k(t, x) \) goes to infinity if \( A \) is recurrent and to a constant if \( A \) is transient. This implies that the additional summand 1 in (2.17) can be omitted asymptotically in both cases. The claim follows.

The lower bound of case iii)a) follows by the same argument as for case ii), and the upper bound is a straightforward induction using Lemmas 3 and 4. The cases iii)b) and c) also follow from Lemma 3 and induction based on the asymptotics for \( M^1 \).

\[ \square \]

3. Acknowledgements

LD would like to thank Martin Kolb for drawing his attention to [ABY98] and Andreas Kyprianou for his invitation to the Bath-Paris workshop on branching processes, where he learnt of the many-to-few lemma from MR. The authors would also like to thank Piotr Milos for checking an earlier draft, and a referee for pointing out several relevant articles.

References

[ABY98] Albeverio, S.; Bogachev, L.; Yarovaya, E. “Asymptotics of branching symmetric random walk on the lattice with a single source” C. R. Acad. Sci. Paris, 326, Serie I, pp. 975-980, 1998
[ABY98b] Albeverio, S.; Bogachev, L.; Yarovaya, E. “Exit from a catalytic domain and extinction of branching random walk on the lattice with a single source” C. R. Acad. Sci. Paris, 326, Serie I, pp. 975-980, 1998
[AB00] Albeverio, S.; Bogachev, L. “Branching Random Walk in a Catalytic Medium. I. Basic Equations” Positivity 4: 41100, 2000
[AA87] Anderson, K. K.; Athreya, K. B. “A Renewal Theorem in the Infinite Mean Case” Annals of Probability, 15, (1987), 388-393
[B11] Bulinskaya E.V. “Limit distributions arising in branching random walks on integer lattices” Lithuan. Math. J., 2011, 51(3), p. 310-321
[B10] Bulinskaya E.V. “Catalytic branching random walk on three-dimensional lattice” Theory Stoch. Process., 16(2):23-32, 2010
[DS10] Döring, L.; Savov, M. “An Application of Renewal Theorems to Exponential Moments of Local Times” Elect. Comm. in Probab. 15 (2010), 263-269
[E73] Erickson, B. “The strong law of large numbers when the mean is undefined” Trans. Amer. Math. Soc., 54, (1973), 371-381
[F71] Feller, W. “An introduction to probability theory and its applications. Vol. II” John Wiley & Sons, Inc., New York-London-Sydney, (1966)
[GH06] Gärtner, J.; Heydenreich, M. “Annealed Asymptotics for the Parabolic Anderson model with a Moving Catalyst” Stoch. Processes and Appl., 116, (2006), pp. 1511-1529
[HR11] Harris, S.; Roberts, M. “The many-to-few lemma and multiple spines” arXiv:1106.4761v1
[HVT10] Hu, Y.; Vatutin, V.A.; Topchii, V.A. “Branching random walk in $\mathbb{Z}^d$ with branching at the origin only” arXiv:1006.4769v1
[KLPP97] Kurtz, T.; Lyons, R.; Pemantle, R.; Peres, Y. “A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes” In K. B. Athreya and P. Jagers, editors, Classical and modern branching processes (Minneapolis, MN, 1994), volume 84 of IMA Vol. Math. Appl., pages 181–185. Springer, New York, 1997.
[L97] Lyons, R. “A simple path to Biggins’ martingale convergence for branching random walk” In K. B. Athreya and P. Jagers, editors, Classical and modern branching processes (Minneapolis, MN, 1994), volume 84 of IMA Vol. Math. Appl., pages 217–221. Springer, New York, 1997.
[LPP95] Lyons, R.; Pemantle, R.; Peres, Y. “Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes” Ann. Probab., 23(3):1125–1138, 1995.
[TV03] Topchii, V. A.; Vatutin, V. A. “Individuals at the origin in the critical catalytic branching random walk” Discrete Math. Theor. Comput. Sci., 6:325-332, 2003.
[VT05] Vatutin, V.A.; Topchii, V.A. “Limit theorem for critical catalytic branching random walks” Theory Prob. Appl., 2005, Vol. 49, No. 3, pp. 498-518.
[VTY04] Vatutin V.A.; Topchii V.A.; Yarovaya E.B. “Catalytic branching random walk and queueing systems with random number of independent servers” Theory Probab. Math. Stat., 69:1-15, 2004.
[Y91] Yarovaya, E. B. “Use of spectral methods to study branching processes with diffusion in a noncompact phase space” Teor. Mat. Fiz. 88 (1991), 25-30 (in Russian); English translation: Theor. Math. Phys. 88 (1991)
[Y10] Yarovaya, E. B. “The monotonicity of the probability of return into the source in models of branching random walks” Mosc. Univ. Math. Bull., 65(2):78-80, 2010.

Fondation Mathématique de Paris

Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany

E-mail address: leif.doering@upmc.fr

E-mail address: mattiroberts@gmail.com