Particle systems with locally dependent branching: long-time behaviour, genealogy and critical parameters

Dissertation
zur Erlangen des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Matthias Birkner
aus Frankfurt am Main

Frankfurt 2003
(D F 1)

vom Fachbereich Mathematik der Johann Wolfgang Goethe – Universität
als Dissertation angenommen.

Dekan : Prof. Dr. Johann Baumeister

Gutachter : Prof. Dr. Anton Wakolbinger
Prof. Dr. Achim Klenke
Prof. Dr. Götz Kersting

Datum der Disputation: 18. Juli 2003

Contents

Chapter 1

Introduction

Consider a system of particles performing random walks on $\mathbb{Z}^d$ and branching, with each particle dying and being replaced by a random number of offspring, the average number of offspring being one. Two forces are in conflict here: the branching produces fluctuations in the spatial distribution and the random motion smooths the fluctuations out. Which of the two tendencies will prevail in the long run? If we begin with a spatially homogeneous particle configuration, will the system evolve towards a non-trivial equilibrium, or are the fluctuations so strong that the system is doomed to local extinction? Will the equilibrium, if it exists, have finite second moments, or will the fluctuations destroy them? These are the central questions to which this thesis addresses itself.

The answers are well known in the classical case where the particles branch independently. If the offspring variance is finite, everything hinges on whether the symmetrized random walk is recurrent or transient. With recurrence, the system is driven to local extinction; with transience it has a one-parameter family of non-trivial equilibria, and finiteness of second moments of the initial state persists in the long-term limit.

We investigate the case of dependent branching, in which the branching behaviours of particles at the same site are correlated, whether by genuine interaction between the particles or simply because they share the same local environment. These two alternative mechanisms of dependence are embodied in the following two models.

1) The branching rate at a site depends on the number of particles there: if there are k individuals at x, a branching event occurs at rate $\sigma(k)$. The particle which is chosen to branch then produces a random number of offspring according to a fixed distribution ν with mean one and finite variance. This model is studied in chapter 2. The classical case of independent branching corresponds to $\sigma(k) = \text{const} \times k$.

2) The offspring laws $Q(x, n)$ in generation n at site x are i.i.d. copies of a random distribution Q on $\mathbb{N}_0$ whose expectation $m_1(Q)$ has mean one and finite variance. Given $Q(x, n)$, all particles at site x in generation n independently produce offspring with distribution $Q(x, n)$. Dependencies in the branching are induced by the random environment, not by the particles themselves. These branching systems in random environment are considered in chapter 4.

For systems of type 1 the following two extreme cases turn out to be especially interesting:

a) $\sigma(k) = C_\sigma k^2$.

b) $\sigma(k) = \mathbf{1}_{\{k=1\}}$.

The branching rate in b) is extremely low: particles only branch as long as they are alone. By contrast, the order of growth of the branching rate function in a) is as high as possible consistent with preventing an explosion of the second moment in finite time.

Assume that the symmetrization of the random walk is transient, and consider spatially homogeneous initial conditions with finite second moments. Then, both in case 1a) and in case 2, the long-term behaviour of the system depends on the interplay of three quantities: b, β_2, and β_*. The parameter b is a measure of the variability of the branching: in case 1a) $b := C_\sigma \mathrm{Var}(\nu)$, and in case 2, $b := \mathrm{Var}(m_1(Q))$. β_2 and β_* describe the random walk, as we now briefly explain.

Let V be the total collision time of two independent copies of the random walk. Call $\mathbb{E}[\exp(\beta V)]$ and $\mathbb{E}[(1+\beta)^V]$ the *exponential moment of order* β in cases 1a) and 2, respectively. Let β_2 be the supremum of all β for which the exponential moment of order β of V is finite, and let β_* be the supremum of all β for which the exponential moment of order β of V, *conditional* on one of the two random walk copies, is finite almost surely.

We show that, both in case 1a) and 2, β_2^{-1} is the value at zero of the Green function of the difference random walk $S - S'$, where S and S' are two independent copies of the underlying random walk. We prove for case 2, using tools from large deviation theory, that β_*^{-1} is the sum over $\exp(-H_n)$, where H_n is the entropy of the random walk position S_n at time n. Moreover, we present strong evidence that the same is true in case 1a), with $n \in \mathbb{N}$ replaced by $t \in \mathbb{R}_+$, and the sum replaced by an integral. In particular, these characterizations ensure that, under mild conditions on the random walk dynamics, β_2 is strictly less than β_*.

Both in case 1a) and in case 2) the three parameters b, β_2, and β_* determine the long-term behaviour of the process, as follows.

(i) for $b < \beta_2$, second moments remain bounded in the large-time limit, and the system converges to an equilibrium which only depends on the initial intensity.

(ii) for $\beta_2 \leq b < \beta_*$, second moments grow exponentially in time, but the system still converges to an equilibrium which preserves the initial intensity. Thus contrary to the classical case of independent branching we have a family of equilibria with local particle numbers of infinite variance.

For the proof of (i) we use coupling arguments and expressions for mixed moments involving random walks. The proof of (ii) relies on a representation of the cluster of "siblings" of a randomly sampled particle in terms of its genealogical tree. For particle systems with independent branching, the idea of employing spatially embedded *locally size-biased* genealogical trees in order to analyse the long-term behavior can be traced back (at least) to the seminal paper of Kallenberg [19]. For particle systems with locally dependent branching as in cases 1 and 2, we give a representation of the locally size-biased genealogical trees (or "Kallenberg trees") in section 2.5 and in section 4.4, respectively.

In case 2 the Kallenberg trees are not conceived as conditioned on the environment (as in [16]) but as randomized over the environment. In such an "annealed" situation, a construction of Kallenberg trees for a special example of locally dependent branching in continuous time related to case 2 (the so-called "coupled branching process") appears in [14]. There the environment is given by an i.i.d. family (indexed by the sites) of homogeneous Poisson processes on the time axis, each Poisson time point designating a "local catastrophe" at the respective site. Given the environment, particles branch independently with a mean offspring bigger than one, but whenever a local catastrophe occurs, the population at this site is wiped out.

Part of the present thesis was motivated by an attempt to understand Greven's arguments in [14] better and to discern a more general structure behind them. The latter endeavour has met with some success; progress in the former has been less complete. In particular, a big question remains open: Does $b > \beta_*$ imply local extinction?

A second source of inspiration for the present thesis has been the ongoing investigations of Greven and den Hollander on the diffusion limit of systems of type 1a), also known as the *parabolic Anderson model*, cf. [1]. This is a system $(X_x(t))_{t\geq 0, x\in\mathbb{Z}^d}$ of interacting diffusions

$$dX_x(t) = \sum_y p_{x-y}(X_y(t) - X_x(t))dt + bX_x(t)dB_x(t),$$

where the B_x are independent standard Brownian motions. (We are grateful to the authors for making their manuscript [15] available.) They give a characterization of β_* in terms of a variational problem similar to our (5.6), but they do not give an explicit formula for the maximiser. Our arguments can be carried over to the continuous-time case (see section 5.5), thus yielding a more explicit characterization of β_*. Shiga proved, using techniques from stochastic analysis, that the parabolic Anderson model becomes locally extinct even for transient p if the coefficient b is sufficiently large, cf. [29]. He was not able to give the exact threshold value; in [15] the authors conjecture that it is β_*.

The parallels between systems of type 1a) and 2, which are expressed in the above stated assertions (i) and (ii), are less surprising when one realizes that both types of systems have the property that the variance of the increment of the local density, given the current population, is proportional to the *square* of the current local density. Thus one might expect that these systems should be in the same "universality class" as the parabolic Anderson model.

Let us now turn to branching particle systems of type 1, with a branching rate $\sigma(\cdot)$ more general than in 1a). (In in order to ensure the "attractivity" of the systems, we assume that $\sigma(k)$ is monotone in k or that the branching is critical binary.) For transient symmetrised particle motion and $\limsup \sigma(k)/k^2 < \beta_2/\mathrm{Var}(\nu)$ we prove that there is persistence (of the initial intensity in the large-time limit). Any such system has a family of equilibria parametrised by the local density, all these equilibria having finite (local) variance. This complements a result of Cox and Greven ([5]) on the corresponding interacting diffusion limit.

For particle systems of type 1 with general particle motion we prove a comparison result showing that if the system ξ uses branching rate function σ, the system ξ' uses branching rate function σ', and $\sigma(\cdot) \geq \sigma'(\cdot)$, then certain convex functionals of $\xi(t)$ have larger expectations than their counterparts involving $\xi'(t)$. This is a particle analogue of the main result of [4] who treated interacting diffusions. Applying the comparison theorem to Laplace functionals we get as an immediate consequence that local extinction of ξ implies the same for ξ', which is intuitively plausible because more branching should lead to more fluctuations, driving the system more easily to its (absorbing) vacuum state.

Finally, the state-dependent branchers with *recurrent* symmetrised motion pose intriguing questions. In view of the comparison theorem and well-known properties of systems with independent branching, there must obviously be local extinction whenever $\sigma(\cdot)$ grows at least linearly. On the other hand, can a strong "down-regulation" of the branching rate make possible a non-trivial equilibrium? Taking this to the extreme leads us back to the above-mentioned systems of type 1b), i.e. to $\sigma(k) = \mathbf{1}_{\{k=1\}}$. If such a system of "lonely branchers" suffers local extinction, the answer must generically be "no". The question whether the lonely branchers in $\mathbb{Z}^1$ and $\mathbb{Z}^2$ die out locally was posed by Ted Cox a few years ago (private communication), and is still open.

One tool for investigating the long-time behavior of these systems is the Kallenberg tree: its local clumping is equivalent to local extinction of the particle system, cf. Lemma 9. As we have not yet succeeded in analysing the long-term behavior of the "full" Kallenberg tree of the lonely branchers, instead we have considered a caricature, where the trunk does not move and the side-lines do not branch. This leads to systems of random walks with "self-blocking immigration", which are studied in chapter 3. We have been able to show that these systems experience local overflow whenever the motion is recurrent. This leads us to the conjecture that the same holds for the *true* Kallenberg tree and thus that the lonely branchers die out locally.

We also study the quantitative long-time behaviour of self-blocking immigration systems starting from the empty configuration. We obtain a fairly complete picture for positive recurrent motion: the total number of particles (or, equivalently in this case, the local density) grows logarithmically in time. By applying ideas from the theory of hydrodynamic limits we derive an "effective equation" for the local density in the case of simple random walk on $\mathbb{Z}$, and analyse the precise asymptotic long-time behaviour of this equation. This predicts that the number of particles should grow like $C\sqrt{t}\log t$, and again that the local density grows like $\log t$. Finally, we use the relative entropy method to show that the prediction captures at least the correct power of t, namely that the number grows more slowly than $t^{1/2+\varepsilon}$ for any $\varepsilon > 0$. It remains a challenging problem to make the $\sqrt{t}\log t$-prediction rigorous. Together with the elimination of the "caricature step" this would say something interesting about one-dimensional lonely branchers: The family of an individual alive at time t should be of size $\sqrt{t}\log t$, not of size t as in non-interacting systems.

Acknowledgement

First of all I would like to thank my supervisor, Professor Anton Wakolbinger, for the many interesting discussions and the constant encouragement while this thesis was prepared. He always had a positive outlook on things and believed in me and my work at times when I was sceptical.

Furthermore I have the pleasure to thank Ted Cox, Donald Dawson, Alison Etheridge, Brooks Ferebee, Jochen Geiger, Andreas Greven, Frank den Hollander, Lars Kauffmann, Götz Kersting, Achim Klenke, Bas Lemmens, José Alfredo López Mimbela, Kaya Memişoğlu, Dirk Metzler, Jeremy Quastel, Anja Sturm, Jan Swart and Iljana Zähle for numerous mathematical discussions that shaped this work. I gratefully acknowledge the heroic efforts of Daniela Fischer and Angelika Esser, who proofread the manuscript. The remaining errors are of course all mine.

Last but not least my heartfelt thanks go to Daniela and my parents for their love and support.

Chapter 2

Spatial branching processes with state dependent branching rate

2.1 Scenario

We deal with spatially extended systems of branching particles in continuous time, that is, we consider populations of particles that move around randomly on some discrete space S – we will mostly consider $\mathbb{Z}^d$ – and branch into a random number of offspring. The spatial motion is given by some Markov dynamics which is independent for different particles, but the *rate* at which particles die is a function of the overall number of particles which are currently at the same site. Upon her death, a particle is replaced by a random number of offspring, possibly none. The number of children is independent of everything else with a fixed distribution ν with mean one and finite variance. Here is our list of formal ingredients:

S .... "basic space", some countable set
$\xi_x(t)$ .... no. of particles at $x \in S$ at time t
$\sigma : \mathbb{N}_0 \to \mathbb{R}_+$ .... branching rate function ($\sigma(0) = 0$), each individual at x dies at rate $\sigma(\xi_x)/\xi_x$
$\nu \in \mathcal{M}_1(\mathbb{N}_0)$ .... individual offspring distribution: at her death, a particle is replaced by k offspring with probability ν_k. We assume $\sum_k k\nu_k = 1$,
$\mathrm{Var}(\nu) := \sum_k k^2\nu_k - 1 < \infty$.
$p_{x,y}$ ....irreducible stochastic matrix on S, particles follow independent p-motion with rate κ.

The time evolution of the configuration $\xi(t)$ is a Markov process on a suitably chosen subspace of $\mathbb{Z}_+^S$ with formal generator given by

$$
\begin{aligned}
L\,F(\xi) \quad &= \quad \kappa \sum_{x,y\in S} \xi_x p_{xy}\Big(F(\xi^{(x,y)}) - F(\xi)\Big) \\
&\qquad + \sum_{x\in S} \sigma(\xi_x) \sum_{k\geq 0} \nu_k \Big(F(\xi + (k-1)\delta_x) - F(\xi)\Big) \qquad (2.1) \\
&=: \quad \kappa L_{rw} F(\xi) + L_{br} F(\xi),
\end{aligned}
$$

where δ_x is the configuration with exactly one particle at position x, we use the obvious component-wise addition of particle configurations, and $\xi^{(x,y)} = \xi + \delta_y - \delta_x$ is obtained from ξ by moving one particle from x to y.

We assume further on that

$$\sigma(k) \leq C_\sigma k^2 \quad \text{for some } C_\sigma > 0, \tag{2.2}$$

and that

$$A := \sup_{z\in S} \sum_{y\in S} p_{yz} < \infty. \tag{2.3}$$

Remark 1 (2.2) ensures for $n \in \mathbb{N}$ that if $\xi(0)$ and ν have n moments, this will also be true of $\xi(t)$ for any $t > 0$, see Lemma 3. One checks easily, e.g. by considering a one-point space, that if $\sigma(k) = k^{2+\epsilon}$ for some $\epsilon > 0$ this would not be true. In fact, the system would then loose even second moments in finite time.

(2.3) excludes the possibility of instantaneous implosion of mass moving in from infinity, if we start from an initial condition $\xi(0)$ with uniformly bounded intensity. It guarantees that then the number of particles at any given site and any time is almost surely finite. Observe that (2.3) is satisfied whenever p has an invariant measure α satisfying $0 < \inf \alpha_x \leq \sup \alpha_x < \infty$. It is automatically satisfied for doubly stochastic p, thus in particular for p of random walk type, i.e. p satisfying $p_{xy} = p_{0,y-x}$, and also for p corresponding to a random walk on any directed graph where the vertices have uniformly bounded degree. From (2.3) we easily get by induction that $\sup_z \sum_y p^n_{yz} \leq A^n$ (note e.g. that $\sum_y p^2_{yz} = \sum_y p_{yz} \sum_x p_{xy} \leq A \sum_y p_{yz} \leq A^2$).

Additionally, we impose
Assumption A We assume that one of the following cases holds

$$\nu = \frac{1}{2}(\delta_0 + \delta_2), \text{ i.e. the offspring law is critical binary, or} \tag{2.4}$$

$$\sigma : \mathbb{N}_0 \to \mathbb{R}_+ \text{ is a non-decreasing function.} \tag{2.5}$$

Observe that either of (2.4) or (2.5) ensures that the ξ-process is attractive, as then e.g. the "obvious" coupling provided by (2.18) is order-preserving. By considering a one-point space one easily sees that if both fail, the system will in general *not* be attractive.

While (2.2) and (2.3) are mild and rather natural in our scenario, Assumption A is more stringent. Some parts of our program, namely the construction of the process and the moment computations, could as well be carried through without it. In general, we might not be able to prove that the Markov semigroup specified by L is unique, unless we imposed additional Lipschitz conditions on σ, see the beginning of section 2.2. Other parts, in particular the coupling techniques used in section 2.4 to prove existence of equilibria and find their domains of attraction, depend heavily on attractivity. Nonetheless, it may well be that attractivity is just convenient to work with, but the results hold more generally.

We have to clarify which set of configurations our processes will live on. In order to do so choose a reference function $\gamma : S \to (0, \infty)$ satisfying for some $M \geq 0$

$$\sum_{y\in S} p_{xy}\gamma_y \leq M\gamma_x \quad \text{for all } x \in S \tag{2.6}$$

and $\sum_x \gamma_x < \infty$. We assume without loss of generality that $\sum_x \gamma_x = 1$. We shall see below that then

$$E_1 := \big\{\xi : S \to \mathbb{N}_0 \,\big|\, ||\xi||_{\gamma,1} := \sum_x \gamma_x |\xi_x| < \infty\big\} \tag{2.7}$$

or, when dealing with initial conditions with finite second moments,

$$E_2 := \big\{\xi : S \to \mathbb{N}_0 \,\big|\, ||\xi||^2_{\gamma,2} := \sum_x \gamma_x |\xi_x|^2 < \infty\big\} \quad (\subset E_1) \tag{2.8}$$

are reasonable state spaces for our purposes. We will also have occasion to consider

$$E_{\mathit{fin}} := \big\{\xi : S \to \mathbb{N}_0 \,\big|\, |\xi| := \sum_x |\xi_x| < \infty\big\},$$

the set of all configurations consisting of finitely many particles. Of course, E_1 and E_2 can be given the topology induced by their respective metrics. On the other hand, we will be mostly considering convergence in finite dimensional distributions for E_i-valued processes, $i \in \{1,2\}$. Thus, we usually equip E_i with the (coarser) product topology. Observe that E_1 and E_2 are measurable subsets of $(\mathbb{N}_0)^S$ (if the latter is given the product σ-algebra), and that any probability measure ν on $(\mathbb{N}_0)^S$ with $\sup_{x\in S} \int \xi_x \, d\nu < \infty$ automatically satisfies $\nu(E_1) = 1$ because $\int ||\xi|| \, d\nu \le \sup_{x\in S} \int \xi_x \, d\nu \times \sum_y \gamma_y$. Similarly, $\sup_{x\in S} \int (\xi_x)^2 \, d\nu < \infty$ implies $\nu(E_2) = 1$. As we will be mostly dealing with shift-invariant situations with locally finite intensities our statements will not depend on the particular choice of γ.

The standard way of obtaining a suitable γ, following Liggett and Spitzer, is to choose $\beta : S \to (0,\infty)$ summable and put $\gamma_x := \sum_{n=0}^\infty M^{-n} \sum_{y\in S} p^n_{xy} \beta_y$, where $M > A \vee 1$ and p^n is the n-th matrix power of p, see e.g. [24], p. 444 (observe that $\sum_x \gamma_x \le \sum_n (A/M)^n \sum_y \beta_y$). (2.6) implies $\sum_{y\in S} p^n_{xy}\gamma_y \le M^n \gamma_x$ by induction, hence

$$\sum_{y\in S} p_{xy}(t)\gamma_y \le \exp(t\kappa M)\gamma_x \quad \text{for all } x \in S.$$

For a metric space E, $\mathcal{B}(E)$ denotes the bounded, measurable functions from E to $\mathbb{R}$, $\mathcal{M}_1(E)$ the probability measures on E, and $\mathcal{N}(E)$ the set of purely atomic, integer-valued measures on E.

We consider $E_{\mathit{fin}} \subset E_2 \subset E_1 \subset (\mathbb{N}_0)^S$ to be equipped with the product topology. For $\mu, \mu_1, \mu_2, \ldots \in \mathcal{M}_1(E_1)$ we write $\mu_n \Rightarrow \mu$ for convergence in finite dimensional distributions.

2.2 Formal construction

Here we formally construct the Markov processes $(\xi(t))_{t\ge 0}$ that this chapter is all about. Readers who are satisfied with the informal description given at the beginning of section 2.1 are cordially invited to skip on to the next section where they will find more meaty statements about properties of $(\xi(t))$. But observe that the construction is not completely standard because in general, $\xi_x(t)$ need not have a second moment and $\sigma(\cdot)$ is allowed to grow quadratically. So the expected rate of local changes can be infinite, which causes some trouble for in principle desirable formulas like "$(d/dt)\mathbb{E}\, f(\xi(t)) = \mathbb{E}\, Lf(\xi(t))$", even for bounded f depending only on finitely many coordinates: the righthand side might not be well defined. This kind of problem does not occur if we assume that σ is globally Lipschitz. Then the existence of $(\xi(t))$ as a process on E_1 follows from more or less standard arguments, see e.g. [2], Thm. 13.17. As the case of a quadratic function $\sigma(k) = Ck^2$ exhibits interesting features (see section 2.6), we have refrained from imposing Lipschitz conditions on σ and instead chosen to work with Assumption A.

Let us denote by $\mathrm{Lip}(E_1)$ the set of Lipschitz-continuous functions on E_1, that is all functions $f : E_1 \to \mathbb{R}$ for which there exists a $C_f < \infty$ such that

$$|f(\eta) - f(\eta')| \le C_f ||\eta - \eta'||_{\gamma,1} \quad \text{for all } \eta, \eta' \in E_1.$$

Our strategy for the construction, inspired by [24], is to first consider finite initial conditions, which are then used to approximate more general initial states. On the way, coupling arguments will be our main tool. In doing so, the following system of stochastic differential equations driven by Poisson processes is useful, as it provides a natural simultaneous coupling for all (finite) initial conditions. Assumption A ensures that this coupling is also order-preserving.

Let $\mathcal{F} = (\mathcal{F}_t)_{t\ge0}$ be a (right-continuous, complete) filtration, $N^{\mathrm{rw}}_{x,y}$ for $x, y \in S, x \neq y$, Poisson point processes on $\mathbb{R}_+ \times \mathbb{N}$ with intensity measure $\kappa p_{xy} dt \otimes d\ell$, and $N^{\mathrm{br}}_{z,k}$ for $z \in S, k \in \{0, 2, 3, \dots\}$ Poisson point processes on $\mathbb{R}_+ \times \mathbb{N} \times [0, 1]$ with intensity measure $C_\sigma \nu_k dt \otimes d\ell \otimes du$, all independent and $\mathcal{F}$-adapted in the "time" component. The superscript "rw" refers to random walk, "br" to branching, and ℓ denotes counting measure on $\mathbb{N}$. For brevity, let us denote by $\tilde{N} := \big(N^{\mathrm{rw}}_{x,y}, N^{\mathrm{br}}_{z,k}\big)_{x,y,z\in S,k\in\mathbb{N}_0\setminus\{1\}}$ the collection of all the driving Poisson processes.

We require $(\xi(t))_{t\ge0}$ to be a solution of

$$\begin{aligned}\xi_x(t) \;=\; & \xi_x(0) + \sum_{y\neq x} \Big\{ \int_0^t \mathbf{1}(\xi_y(s-) \ge n) N^{\mathrm{rw}}_{y,x}(ds\, dn) \\ & \qquad\qquad - \int_0^t \mathbf{1}(\xi_x(s-) \ge n) N^{\mathrm{rw}}_{x,y}(ds\, dn) \Big\} \\ & + \sum_{k\ge0} \int_0^t (k-1)\mathbf{1}\Big(\xi_x(s-) \ge \sqrt{n}, \frac{\sigma(\xi_x(s-))}{C_\sigma \xi_x(s-)^2} \ge u\Big) N^{\mathrm{br}}_{x,k}(ds\, dn\, du)\end{aligned} \tag{2.9}$$

for all $x \in S$ and $t \ge 0$.

Lemma 1 *a) For any $\eta \in E_{fin}$ there is a unique strong solution $(\xi(t))_{t\ge0}$ of (2.9) starting from $\xi(0) = \eta$. More precisely, there is a (measurable) function*

$$\Phi : E_{fin} \times \mathbb{R}_+ \times \big(\mathcal{N}(\mathbb{R}_+ \times \mathbb{N})^{S\times S} \times \mathcal{N}(\mathbb{R}_+ \times \mathbb{N} \times [0, 1])^{S\times(\mathbb{N}_0\setminus1)}\big) \to E_{fin}$$

such that for any $\eta \in E_{fin}$

$$\big(\Phi(\eta, t, \tilde{N})\big)_{t\ge0}$$

is an $(\mathcal{F}_t)$-adapted, E_{fin}-valued process with càdlàg paths that starts from η and is a solution to (2.9).

b) The function Φ provides a simultaneous monotone coupling for solutions to (2.9): Let $\eta \le \tilde{\eta} \in E_{fin}$ and

$$\xi(t) = \Phi(\eta, t, \tilde{N}),\ \tilde{\xi}(t) = \Phi(\tilde{\eta}, t, \tilde{N}), \quad t \ge 0.$$

Then

$$\xi_x(t) \le \tilde{\xi}_x(t) \quad \textit{for all } x \in S,\ t \ge 0 \quad \textit{almost surely.}$$

c) $(\xi(t))$ is a Markov process, we denote its semigroup by

$$S(t)f(\eta) := \mathbb{E}[f(\xi(t))|\xi(0) = \eta], \quad \eta \in E_{\mathit{fin}} \tag{2.10}$$

acting on $f \in \mathcal{B}((N_0)^S) \cup \mathrm{Lip}(E_1)$. $S(t)f \in \mathrm{Lip}(E_1)$ for $f \in \mathrm{Lip}(E_1)$, more precisely

$$|S(t)f(\eta) - S(t)f(\tilde{\eta})| \leq C_f \exp(t\kappa M)||\eta - \tilde{\eta}||_{\gamma,1} \quad \text{for } \eta, \tilde{\eta} \in E_{fin}, \tag{2.11}$$

where C_f is the Lipschitz constant of f. Furthermore, for any $f \in \mathrm{Lip}(E_1)$ and $\xi(0) = \eta \in E_{\mathit{fin}}$, the process

$$M^f(t) := f(\xi(t)) - f(\xi(0)) - \int_0^t Lf(\xi(s))\, ds \tag{2.12}$$

is a martingale.

d) For $m \in \mathbb{N}$ let

$$||\eta||_{\gamma,m} := \Big(\sum \gamma_x |\eta_x|^m\Big)^{1/m}$$

be the $L^m(\gamma)$-norm of η. If $\sum_k k^m \nu_k < \infty$ for some $m \in \mathbb{N}$, then there is a $C_m < \infty$ such that

$$\mathbb{E}\Big[||\xi(t)||^m_{\gamma,m}\Big] \leq \exp(C_m t)||\eta||^m_{\gamma,m} \tag{2.13}$$

for any $\eta \in E_{\mathit{fin}}$, where $(\xi(t))$ is the solution of (2.9) starting from η.

Proof. a) Fix $\eta \in E_{\mathit{fin}}$, define

$$T_1 := \inf\left\{ t \ : \ \begin{array}{l} N^{\mathrm{rw}}_{x,y}([0,t] \times \{1,2,\ldots,\eta_x\}) > 0 \text{ for some } x, y \in S \text{ or} \\ N^{\mathrm{br}}_{x,k}([0,t] \times \{1,2,\ldots,\eta_x^2\} \times [0,1]) > 0 \text{ for some } x \text{ and } k \end{array} \right\}.$$

Then $T_1 > 0$ almost surely, define $\xi(t) := \eta$ for $t \in [0, T_1)$. If $\Delta N^{\mathrm{rw}}_{x,y}(T_1, \{i\}) = 1$ for some x, y and i with $\xi_x(T_1-) \leq i$ put $\xi(T_1) := \xi(T_1-) + \delta_y - \delta_x$. Otherwise, there is some x, k and i, u such that $\Delta N^{\mathrm{br}}_{x,k}(T_1, \{i\}, \{u\}) = 1$ and $i \leq \xi_x(T_1-)^2$. In this case we put

$$\xi(T_1) := \xi(T_1-) + \mathbf{1}(u \leq \sigma(\xi_x(T_1-))/C_\sigma \xi_x(T_1-)^2)(k-1)\delta_x.$$

Note that independent Poisson processes almost surely have no jump times in common, so that this procedure is always well defined. Define similarly the jump time T_2 and $\xi(T_2)$, put $\xi(t) := \xi(T_1)$ for $t \in [T_1, T_2)$. Iterating this procedure we obtain a sequence of random times (in fact, $\mathcal{F}$-stopping times) $0 =: T_0 < T_1 < T_2 < T_3 < \cdots$ and a process $\xi(t)$ defined on $[0, \sup_n T_n)$ with paths in E_{fin} that are constant on each interval $[T_{i-1}, T_i)$. From $\sum_k (k-1)\nu_k = 0$ we get that $M_n := \sum_x \xi_x(T_n)$ is a non-negative martingale, hence $\sup_n M_n < \infty$ almost surely. This implies that $\sup_n T_n = \infty$, so that $(\xi(t))$, which is by construction a $\mathcal{F}$-adapted solution to (2.9), is well defined, has càdlàg paths and is E_{fin}-valued for all times. The function Φ is defined by the above procedure (we refrain from stating a lengthy formal definition).

In order to obtain (pathwise) uniqueness observe that for any other solution $(\tilde{\xi}(t))_{t \geq 0}$ starting from the same $\tilde{\xi}(0) = \eta$ we see from (2.9) that $\tilde{\xi}_x(t) = \xi_x(t)$ for all $x \in S$ and $t \in [0, T_1)$, and also that $\tilde{\xi}(T_1) = \xi(T_1)$. Now proceed inductively to see that ξ and $\tilde{\xi}$ agree on $[T_{n-1}, T_n)$ for all $n \in \mathbb{N}$.

b) Let $\eta \le \tilde\eta$, $\xi(\cdot)$ and $\tilde\xi(\cdot)$ be as in the statement. We have $\xi_\cdot(0) \le \tilde\xi_\cdot(0)$, and we have to prove that any jump of the joint process $\big(\xi,\tilde\xi\big)(t)$ again leads to an ordered pair. Assume for example that the Poisson process $N^{\rm br}_{x,k}$ has an atom $\{t\}\times\{n\}\times\{u\}$ (with $\xi_x(t-)\vee\tilde\xi_x(t-) \ge \sqrt{n}$), and that the pair $\big(\xi,\tilde\xi\big)(s)$ is ordered for all $s<t$. Now Assumption A comes into play: If (2.5) holds true, i.e. $\sigma(\cdot)$ is non-decreasing, then there are three possibilities for the joint process:

1. Depending on u, no jump might occur.

2. The (smaller) ξ-component jumps, replacing $\xi_x(t-)$ by $\xi_x(t-)+(k-1)$. But then because $\tilde\xi_x(t-)\ge\xi_x(t-)$, $\tilde\xi$ makes the same jump.

3. Only the $\tilde\xi$-component jumps. Note that this can only happen if $\tilde\xi_x(t-) > \xi_x(t-)$, and that then still $\tilde\xi_x(t-)+(k-1)\ge\xi_x(t-)$ because $k\ge 0$.

In either case, the ordering is preserved. The case of (2.4), i.e. binary branching, and the jumps of the "motion driving" processes $N^{\rm rw}_{x,y}$ can be treated similarly and are left to the reader.

c) and d) The fact that Poisson processes have independent and stationary increments implies that the distribution of $\xi(t+h)$, given $\mathcal{F}_t$, depends only on $\xi(t)$, thus $(\xi(t))$ is Markov. Now consider any measurable function $f:(\mathbb{N}_0)^S\to\mathbb{R}$. Then M^f defined by (2.12) need not be a martingale, but by compensating the driving Poisson processes in (2.9) we see that it is at least a local martingale: A natural localizing sequence of stopping times is given by

$$T_n := \inf\Big\{t\ge 0\ :\ \sum_x \xi_x(t) > n \text{ or there is } y\notin S_n \text{ such that } \xi_y(t)>0\Big\},$$

where $S_n \nearrow S$ is an exhausting sequence of finite subsets of S. Consider $m\in\mathbb{N}$ such that $\sum_k k^m\nu_k<\infty$ – note that our standing assumptions imply that this is always true for $m=1,2$ – and put $\psi_m(\eta) := ||\eta||^m_{\gamma,m}$. We have for η with $\psi_m(\eta)<\infty$

$$\begin{aligned}
L\psi_m(\eta) &= \kappa\sum_{x,y}\eta_x p_{xy}\Big\{\gamma_y\big((\eta_y+1)^m-\eta_y^m\big)+\gamma_x\big((\eta_x-1)^m-\eta_x^m\big)\Big\}\\
&\quad+\sum_x\sigma(\eta_x)\sum_k\nu_k\gamma_x\Big\{(\eta_x+k-1)^m-\eta_x^m\Big\}\\
&= \kappa\sum_x\gamma_x\eta_x\sum_{j=1}^m\binom{m}{j}(-1)^j\eta_x^{m-j}+\kappa\sum_{x,y}\eta_x p_{xy}\gamma_y\sum_{j=1}^m\binom{m}{j}\eta_y^{m-j}\\
&\quad+\sum_x\sigma(\eta_x)\gamma_x\sum_k\nu_k\sum_{j=1}^m\binom{m}{j}\eta_x^{m-j}(k-1)^j.
\end{aligned}$$

Observe that the term for $j=1$ in the last line disappears because $\sum(k-1)\nu_k=0$. Because of (2.2) we can estimate

$$|L\psi_m(\eta)| \le \kappa c_m\left(\psi_m(\eta)+\sum_{x,y}\eta_x p_{xy}\gamma_y\eta_y^{m-1}\right)+c'_m\psi_m(\eta) \qquad (2.14)$$

where $c_m := \sum_{j=1}^m \binom{m}{j} (= 2^m - 1)$, $c'_m := C_\sigma \sum_{j=2}^m \binom{m}{j} \sum_k \nu_k (k-1)^j < \infty$. Now the function $\tilde\eta$ given by $\tilde\eta_y := \sum_x \eta_x p_{xy}$ lies in $L^m(\gamma)$:

$$\begin{aligned}
||\tilde\eta||_{\gamma,m}^m &= \sum_y \gamma_y \Big(\sum\nolimits_z p_{zy}\Big)^m \Big| \sum_x \frac{p_{xy}}{\sum_z p_{zy}} \eta_x^m \Big|^m \\
&\le \sum_y \gamma_y \Big(\sum\nolimits_z p_{zy}\Big)^{m-1} \sum_x \eta_x^m p_{xy} \le A^{m-1} M \sum_x \gamma_x \eta_x^m = A^{m-1} M ||\eta||_{\gamma,m}^m
\end{aligned}$$

by Jensen's inequality, (2.3) and (2.6). This allows to use Hölder's inequality (with $p = m$, $q = m/(m-1)$) on the righthand side of (2.14) to estimate

$$|L\psi_m(\eta)| \le \underbrace{\Big(\kappa c_m (1 + (A^{m-1}M)^{(m-1)/m}) + c'_m\Big)}_{=: C_m} \psi_m(\eta). \tag{2.15}$$

Now for any n, $(M^{\psi_m}(t \wedge T_n))_{t \ge 0}$ *is* a martingale, thus

$$\begin{aligned}
\mathbb{E}[\psi_m(\xi(t \wedge T_n))] &= \psi_m(\eta) + \mathbb{E}\Big[\int_0^{t\wedge T_n} L\psi_m(\xi(s))\, ds\Big] \\
&\le \psi_m(\eta) + C_m \int_0^t \mathbb{E}\big[\mathbf{1}(s \le T_n)\psi_m(\xi(s))\big]\, ds \\
&\le \psi_m(\eta) + C_m \int_0^t \mathbb{E}\big[\psi_m(\xi(s \wedge T_n))\big]\, ds.
\end{aligned}$$

Thus we see from Gronwall's lemma that $\mathbb{E}[\psi_m(\xi(t \wedge T_n))] \le \exp(C_m t)\psi_m(\eta)$ uniformly in n. (2.13) follows from this and Fatou's lemma by letting $n \to \infty$.

Next we wish to prove that (2.12) is indeed a martingale for $f \in \mathrm{Lip}(E_1)$. A key observation is that for any such f there is a constant $C = C(\kappa, p, \sigma, \nu, f)$ such that

$$|Lf(\eta)| \le C||\eta||_{\gamma,2}^2 \quad \text{for all } \eta \in E_2. \tag{2.16}$$

To prove (2.16) note that

$$\begin{aligned}
|L_{rw} f(\eta)| &= \Big| \sum_{x,y} \eta_x p_{xy} (f(\eta^{(x,y)}) - f(\eta)) \Big| \le \sum_{x,y} \eta_x p_{xy} |f(\eta^{(x,y)}) - f(\eta)| \\
&\le C_f \sum_{x,y} \eta_x p_{xy} ||\eta^{(x,y)} - \eta||_{\gamma,1} \le C_f \sum_{x,y} \eta_x p_{xy} (\gamma_y + \gamma_x) \\
&\le C_f (M+1) ||\eta||_{\gamma,1} \le C_f (M+1) ||\eta||_{\gamma,2}^2,
\end{aligned}$$

where M is defined as in (2.6), and $||\eta||_{\gamma,1} \le ||\eta||_{\gamma,2}^2$ because configurations are integer-valued. Additionally, observe that

$$\begin{aligned}
|L_{br} f(\eta)| &= \Big| \sum_x \sigma(\eta_x) \sum_k \nu_k (f(\eta + (k-1)\delta_x) - f(\eta)) \Big| \\
&\le \sum_x \sigma(\eta_x) \sum_k \nu_k |f(\eta + (k-1)\delta_x) - f(\eta)| \\
&\le C_f \sum_x \sigma(\eta_x) \sum_k \nu_k ||(k-1)\delta_x||_{\gamma,1} = \Big(C_f \sum_k \nu_k |k-1|\Big) \sum_x \sigma(\eta_x)\gamma_x \\
&\le \Big(C_f C_\sigma \sum_k \nu_k |k-1|\Big) ||\eta||_{\gamma,2}^2
\end{aligned}$$

by assumption (2.2). Thus (2.16) holds true with

$$C := C_f\Big(M + 1 + C_\sigma \sum\nolimits_k \nu_k|k-1|\Big). \tag{2.17}$$

Now let us consider a bounded $f \in \mathrm{Lip}(E_1)$. As noted above $(M^f(t))$ is a local martingale, so we have for all $t, h \geq 0$

$$\mathbb{E}\big[M^f((t+h)\wedge T_n)\big|\mathcal{F}_t\big] = M^f(t \wedge T_n) \quad \text{a.s.}.$$

The righthand side tends to $M^f(t)$ as $n \to \infty$, to treat the lefthand side observe that

$$\mathbb{E}\big[f\big(\xi((t+h)\wedge T_n)\big) - f(\xi(0))\big|\mathcal{F}_t\big] \to \mathbb{E}\big[f\big(\xi(t+h)\big) - f(\xi(0))\big|\mathcal{F}_t\big]$$

by dominated convergence, and that

$$\mathbb{E}\bigg[\int_0^{(t+h)\wedge T_n} Lf(\xi(s))\,ds\bigg|\mathcal{F}_t\bigg] \to \mathbb{E}\bigg[\int_0^{t+h} Lf(\xi(s))\,ds\bigg|\mathcal{F}_t\bigg]$$

as $n \to \infty$ again by dominated convergence, as the term inside the conditional expectation is bounded by $C\int_0^{t+h}||\xi(s)||^2_{\gamma,2}\,ds$, which is integrable. Thus $(M^f(t))$ is indeed a martingale in this case.

Next, let $f \in \mathrm{Lip}(E_1)$ be non-negative, but not necessarily bounded. Put $f_n(\eta) := f(\eta)\wedge n$, note that f_n is bounded, $f_n \in \mathrm{Lip}(E_1)$ with Lipschitz constant $C_{f_n} \leq C_f$ (as the mapping $\mathbb{R}_+ \ni x \mapsto x \wedge n$ has Lipschitz constant 1). We have $M^{f_n}(t) \to M^f(t)$ a.s. as $n \to \infty$, $\mathbb{E}\big[f_n(\xi(t+h))\big|\mathcal{F}_t\big] \to \mathbb{E}\big[f(\xi(t+h))\big|\mathcal{F}_t\big]$ a.s. by monotone convergence, observing that $|Lf_n(\eta)| \leq C||\eta||^2_{\gamma,2}$ uniformly in n (note that by (2.17), C depends on f_n only through its Lipschitz constant) we obtain

$$\mathbb{E}\bigg[\int_0^{t+h} Lf_n(\xi(s))\,ds\bigg|\mathcal{F}_t\bigg] \to \mathbb{E}\bigg[\int_0^{t+h} Lf(\xi(s))\,ds\bigg|\mathcal{F}_t\bigg]$$

a.s. as $n \to \infty$ by dominated convergence. Thus $(M^f(t))$ is a martingale for non-negative Lipschitz f. To treat the general case we decompose $f \in \mathrm{Lip}(E_1)$ as $f = f^+ - f^-$ into its positive $f^+(\eta) := f(\eta)\vee 0$ and negative part $f^-(\eta) := (-f(\eta))\vee 0$.

It remains to prove that the semigroup $(S(t))$ corresponding to $(\xi(t))$ maps $\mathrm{Lip}(E_1)$ into itself. Fix $f \in \mathrm{Lip}(E_1)$, $\eta^{(1)}, \eta^{(2)} \in E_{\mathit{fin}}$, put $\eta^{(3)} := \eta^{(1)} \wedge \eta^{(2)}$, $\eta^{(4)} := \eta^{(1)} \vee \eta^{(2)}$. By solving (2.9) simultaneously for these four initial conditions (e.g. using the function Φ from part a)) we obtain $(\xi^{(j)}(t))_{t\geq 0}$, $j = 1, \ldots, 4$, where $\xi^{(j)}(0) = \eta^{(j)}$, on the same probability space such that $\xi^{(3)}_x(t) \leq \xi^{(1)}_x(t), \xi^{(2)}_x(t) \leq \xi^{(4)}_x(t)$ for all $x \in S$ and $t \geq 0$. Thus we see that

$$\mathbb{E}|\eta^{(1)}_x(t) - \eta^{(2)}_x(t)| \leq \mathbb{E}\left[\eta^{(4)}_x(t) - \eta^{(3)}_x(t)\right],$$

showing that

$$\begin{aligned}
\Big|S(t)f(\eta^{(1)}) - S(t)f(\eta^{(2)})\Big| &= \Big|\mathbb{E}\left[f(\eta^{(1)}(t)) - f(\eta^{(2)}(t))\right]\Big| \\
&\leq C_f\mathbb{E}||\eta^{(1)}(t) - \eta^{(2)}(t)||_{\gamma,1} = C_f\sum_x \gamma_x \mathbb{E}|\eta^{(1)}_x(t) - \eta^{(2)}_x(t)| \\
&\leq C_f\sum_x \gamma_x\mathbb{E}\left[\eta^{(4)}_x(t) - \eta^{(3)}_x(t)\right] = C_f\sum_{x,y}\gamma_x(\eta^{(4)}_y - \eta^{(3)}_y)p_{yx}(t) \\
&\leq C_f e^{t\kappa M}\sum_y \gamma_y(\eta^{(4)}_y - \eta^{(3)}_y) = C_f e^{t\kappa M}||\eta^{(1)} - \eta^{(2)}||_{\gamma,1}.
\end{aligned}$$

This proves (2.11). $\square$

Remark 2 By considering two solutions of (2.9) starting from η resp. $\tilde{\eta} \in E_{\mathit{fin}}$ simultaneously, we obtain a Markov process $(\xi(t), \tilde{\xi}(t))_{t \geq 0}$ on $E_{\mathit{fin}} \times E_{\mathit{fin}}$ starting from $(\eta, \tilde{\eta})$ such that each component is a system of (state dependent) branching random walks as above and the two systems move their particles together as much as possible. The generator of the joint process is given by

$$
\begin{aligned}
&L^{(2)} f(\xi, \tilde{\xi}) \\
&\quad = \kappa \sum_{x,y} (\xi_x \wedge \tilde{\xi}_x) p_{xy} \left(f(\xi^{(x,y)}, \tilde{\xi}^{(x,y)}) - f(\xi, \tilde{\xi}) \right) \\
&\quad + \kappa \sum_{x,y} (\xi_x - \xi_x \wedge \tilde{\xi}_x) p_{xy} \left(f(\xi^{(x,y)}, \tilde{\xi}) - f(\xi, \tilde{\xi}) \right) \\
&\quad + \kappa \sum_{x,y} (\tilde{\xi}_x - \xi_x \wedge \tilde{\xi}_x) p_{xy} \left(f(\xi, \tilde{\xi}^{(x,y)}) - f(\xi, \tilde{\xi}) \right) \\
&\quad + \sum_x \left(\sigma(\xi_x) \wedge \sigma(\tilde{\xi}_x) \right) \sum_k \nu_k \left(f(\xi + (k-1)\delta_x, \tilde{\xi} + (k-1)\delta_x) - f(\xi, \tilde{\xi}) \right) \\
&\quad + \sum_x \left(\sigma(\xi_x) - \sigma(\xi_x) \wedge \sigma(\tilde{\xi}_x) \right) \sum_k \nu_k \left(f(\xi + (k-1)\delta_x, \tilde{\xi}) - f(\xi, \tilde{\xi}) \right) \\
&\quad + \sum_x \left(\sigma(\tilde{\xi}_x) - \sigma(\xi_x) \wedge \sigma(\tilde{\xi}_x) \right) \sum_k \nu_k \left(f(\xi, \tilde{\xi} + (k-1)\delta_x) - f(\xi, \tilde{\xi}) \right),
\end{aligned}
\tag{2.18}
$$

the process

$$
f(\xi(t), \tilde{\xi}(t)) - f(\xi(0), \tilde{\xi}(0)) - \int_0^t L^{(2)} f(\xi(s), \tilde{\xi}(s))\, ds
$$

is a martingale for $f \in \mathrm{Lip}(E_1 \times E_1)$. This can be proved along the same lines as Lemma 1.

So far we have constructed the Markov process corresponding to the generator L given by (2.1) for finite initial conditions. Now we extend the definition to general initial states in E_1 by approximation. Let us denote by $\eta \mathbf{1}_{S'}$ the restriction of the configuration $\eta \in E_1$ to the subset $S' \subset S$, i.e.

$$
(\eta \mathbf{1}_{S'})_x = \eta_x \mathbf{1}(x \in S').
$$

Lemma 2 *Fix an initial condition $\eta \in E_1$, let $S_n \nearrow S$ be an exhausting sequence of finite subsets. For $n \in \mathbb{N}$, let $(\xi^{(n)}(t))_{t \geq 0}$ be the solution of (2.9) starting from $\xi^{(n)}(0) = \eta \mathbf{1}_{S_n}$, i.e. $\xi^{(n)} = \Phi(\eta \mathbf{1}_{S_n}, \cdot, \tilde{N})$, where Φ is as defined in Lemma 1, a). Then there is an E_1-valued $(\mathcal{F}_t)$-adapted process $(\xi(t))_{t \geq 0}$ with càdlàg paths starting from $\xi(0) = \eta$ such that for all $T > 0$, finite $B \subset S$*

$$
\mathbb{P}\Big(\bigcup_m \bigcap_{n \geq m} \big\{ \xi_x^{(n)}(t) = \xi_x(t) \quad \textit{for all } x \in B,\ t \in [0, T] \big\} \Big) = 1. \tag{2.19}
$$

$(\xi(t))$ is a (strong) solution of (2.9).

Furthermore, the limit process $\xi(\cdot)$ we obtain does not depend on the choice of the sequence (S_n) used in the approximation procedure.

Proof. Define $h(\xi,\tilde\xi) := ||\xi-\tilde\xi||_{\gamma,1}$. Then we have

$$\begin{aligned}
L^{(2)}h(\xi,\tilde\xi) &= \kappa\sum_{x,y}(\xi_x-\tilde\xi_x)^+p_{xy}\big(||\xi+\delta_y-\delta_x-\tilde\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big)\\
&\quad+\kappa\sum_{x,y}(\tilde\xi_x-\xi_x)^+p_{xy}\big(||\tilde\xi+\delta_y-\delta_x-\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big)\\
&\quad+\sum_x(\sigma(\xi_x)-\sigma(\tilde\xi_x))^+\sum_k\nu_k\big(||\xi+(k-1)\delta_x-\tilde\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big)\\
&\quad+\sum_x(\sigma(\tilde\xi_x)-\sigma(\xi_x))^+\sum_k\nu_k\big(||\tilde\xi+(k-1)\delta_x-\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big).
\end{aligned}$$

The sum of the "motion" terms yields $\kappa\sum_{x,y}|\xi_x-\tilde\xi_x|p_{xy}(\gamma_y-\gamma_x)$. For the "branching" terms, Assumption A comes into play: assume (2.4), i.e. critical binary branching. Then $\sum_k\nu_k\big(||\xi+(k-1)\delta_x-\tilde\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big)=(|\xi_x+1-\tilde\xi_x|+|\xi_x-1-\tilde\xi_x|-2|\xi_x-\tilde\xi_x|)/2=0$, if $\xi_x\neq\tilde\xi_x$, because $|a+1|+|a-1|-2|a|=0$ for $a\in\mathbb{Z}\setminus\{0\}$. Note that x with $\xi_x=\tilde\xi_x$ do not contribute to the sum. Analogously the second sum yields 0. Consequently we have in this case

$$L^{(2)}h(\xi,\tilde\xi)=\kappa\sum_{x,y}|\xi_x-\tilde\xi_x|p_{xy}(\gamma_y-\gamma_x)\le\kappa(M-1)||\xi-\tilde\xi||_{\gamma,1} \tag{2.20}$$

by (2.6). Now assume (2.5), i.e. $\sigma(\cdot)$ is non-decreasing. Then only x with $\xi_x>\tilde\xi_x$ contribute $\sum_k\nu_k\big(||\xi+(k-1)\delta_x-\tilde\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}\big)$ to the first "branching" sum. But in this case we have $||\xi+(k-1)\delta_x-\tilde\xi||_{\gamma,1}-||\xi-\tilde\xi||_{\gamma,1}=\gamma_x(|\xi_x+(k-1)-\tilde\xi_x|-|\xi_x-\tilde\xi_x|)=\gamma_x(k-1)$, and the term vanishes because $\sum_k k\nu_k=1$ by assumption. The second sum vanishes analogously, and hence (2.20) also holds in this case.

Let $g(t,\xi,\tilde\xi):=\exp(-t\kappa(M-1))||\xi-\tilde\xi||_{\gamma,1}$, then

$$M^{(n)}(t):=g(t,\xi^{(n)}(t),\xi^{(n+1)}(t))-\int_0^t\big(\frac{\partial}{\partial s}+L^{(2)}\big)g(s,\xi^{(n)}(s),\xi^{(n+1)}(s))\,ds$$

is a martingale starting from $M^{(n)}(0)=||\xi(\mathbf{1}_{S_{n+1}}-\mathbf{1}_{S_n})||_{\gamma,1}$. By (2.20), $M^{(n)}$ is non-negative, and $||\xi^{(n)}(t)-\xi^{(n+1)}(t)||_{\gamma,1}\le e^{t\kappa(M-1)}M^{(n)}(t)$ for all $t\ge0$. Thus we see that for any $\varepsilon>0$

$$\begin{aligned}
\mathbb{P}\big(\sup_{t\le T}||\xi^{(n)}(t)-\xi^{(n+1)}(t)||_{\gamma,1}\ge\varepsilon\big) &\le \mathbb{P}\big(\sup_{t\le T}M^{(n)}(t)\ge e^{-T\kappa(M-1)}\varepsilon\big)\\
&\le \frac{e^{T\kappa(M-1)}}{\varepsilon}||\eta(\mathbf{1}_{S_{n+1}}-\mathbf{1}_{S_n})||_{\gamma,1}\,,
\end{aligned}$$

where the second inequality comes from a Doob inequality for non-negative martingales, see e.g. [9], Prop. 2.2.16. As $\sum||\eta(\mathbf{1}_{S_{n+1}}-\mathbf{1}_{S_n})||_{\gamma,1}=||\eta||_{\gamma,1}<\infty$ we can conclude with the Borel-Cantelli lemma that for any $\varepsilon>0$ only finitely many of the events

$$\big\{\sup_{t\le T}||\xi^{(n)}(t)-\xi^{(n+1)}(t)||_{\gamma,1}\ge\varepsilon\big\}$$

occur. Observe that the opposite event implies that $\xi^{(n)}_x(t)=\xi^{(n+1)}_x(t)$ for all $t\in[0,T]$ and all $x\in S$ with $\gamma_x>\varepsilon$. This proves (2.19), which in turn implies that $\xi(t)$ is $\mathcal{F}_t$-adapted for any t and that $\xi(\cdot)$ has càdlàg paths. Observe that

$$\mathbb{E}\big[||\xi(t)||_{\gamma,1}\big]=\lim_n\mathbb{E}\big[||\xi^{(n)}(t)||_{\gamma,1}\big]\le\lim_n\exp(C_1t)||\eta\mathbf{1}_{S_n}||_{\gamma,1}=\exp(C_1t)||\eta||_{\gamma,1}<\infty$$

by monotone convergence and (2.13) to see that $\xi(t) \in E_1$ almost surely.

We wish to prove that $\xi(\cdot)$ solves (2.9). Fix $x \in S$ and $t \geq 0$ for the moment. For given $\varepsilon > 0$ we can find a finite $B \subset S \setminus \{x\}$ such that the event

$$A_1(B) := \big\{\text{some particle jumps from } B^c \text{ to } x \text{ during } [0,t]\big\}$$

has probability at most ε, because by the above $\mathbb{E}\big[||\xi(s)||_{\gamma,1}\big]$ is uniformly bounded for $s \leq t$. The above arguments also show that there is some n such that

$$A_2(B) := \big\{\xi_y(s) = \xi_y^{(n)}(s) \text{ for all } y \in B,\, s \in [0,t]\big\}$$

has probability at least $1-\varepsilon$. Note that $A_1(B)^c \cap A_2(B)$ imply that $\xi(\cdot)$ solves (2.9) at (x,t), because for any n, $\xi^{(n)}$ is a solution. Taking $\varepsilon \to 0$ we see that

$$\mathbb{P}\big(\xi(\cdot) \text{ solves (2.9) at } (x,t)\big) = 1.$$

Then the fact that $\xi(\cdot)$ has càdlàg paths shows that it is indeed a solution at all space-time points simultaneously.

Finally note that our arguments show that for finite subsets $S', S'' \subset S$, and $\xi(\cdot)$, $\xi'(\cdot)$ solutions of (2.9) starting from $\eta \mathbf{1}_{S'}$ resp. $\eta \mathbf{1}_{S''}$ we have

$$\mathbb{P}\big(\sup_{t\leq T} ||\xi'(t) - \xi''(t)||_{\gamma,1} \geq \varepsilon\big) \leq C(T,\varepsilon)||\eta\mathbf{1}_{S'} - \eta\mathbf{1}_{S''}||_{\gamma,1} \leq 2C(T,\varepsilon)||\eta\mathbf{1}_{(S'\cap S'')^c}||_{\gamma,1}.$$

This proves that $\xi(\cdot)$ does not depend on the choice of the sequence (S_n). $\square$

Proposition 1 *The process $(\xi(t))_{t\geq 0}$ constructed in Lemma 2 is a Markov process, we denote its semigroup again by*

$$S(t)f(\eta) := \mathbb{E}\big[f(\xi(t))\big|\xi(0) = \eta\big], \quad \eta \in E_1.$$

It is the unique extension of the semigroup constructed in Lemma 1 to inputs $\eta \in E_1$.

For all $f \in \mathrm{Lip}(E_1)$, $S(t)f$ is well defined and lies again in $\mathrm{Lip}(E_1)$, its Lipschitz constant is at most $C_f \times \exp(t\kappa M)$.

If f is either bounded, or in $\mathrm{Lip}(E_1)$, or monotone, then $S(t)f$ can be computed as

$$S(t)f(\eta) = lim_{n\to\infty} S(t)f(\eta\mathbf{1}_{S_n}) \quad \textit{for } \eta \in E_1, \tag{2.21}$$

where $S_n \nearrow S$ is any exhausting sequence of finite subsets of S. Furthermore $S(t)$ satisfies

$$S(t)f(\eta) = f(\eta) + \int_0^t \big(S(s)(Lf)\big)(\eta)\, ds \tag{2.22}$$

for all $\eta \in E_2$ and $f \in \mathrm{Lip}(E_1)$, also for $\eta \in E_1$ and any $f(\eta) = g(\eta_{x_1}, \ldots, \eta_{x_m})$ depending on finitely many coordinates that is constant on $\{\eta : \eta_{x_1} + \cdots + \eta_{x_m} \geq M\}$ for some $M > 0$. Under any of these conditions $t \mapsto S(t)f(\eta)$ is differentiable with

$$\frac{d}{dt} S(t)f(\eta) = \big(S(t)(Lf)\big)(\eta). \tag{2.23}$$

Proof. By Lemma 2, $(\xi(t))_{t\geq 0}$ is a solution to (2.9), hence the Markov property follows from the fact that Poisson processes have independent increments. Note that (2.21) is satisfied for bounded f by the dominated convergence theorem: in Lemma 2, we have constructed $\xi(\cdot)$ and the approximating $\xi^{(n)}(\cdot)$ on the same probability space.

Formula (2.11) in Lemma 1, c) shows that (2.21) is also true for $f \in \mathrm{Lip}(E_1)$, and gives the bound on the Lipschitz constant of $S(t)f$. For a monotone function $f : E_1 \to \mathbb{R}$, not necessarily Lipschitz, the limit in (2.21) exists because then by Lemma 1, b) we have $S(t)f(\eta\mathbf{1}_{S_n}) \le S(t)f(\eta\mathbf{1}_{S_{n'}})$ for $n \le n'$.

Furthermore note that any bounded function $g : E_1 \to \mathbb{R}$ which depends only on finitely many coordinates $x_1, \dots, x_m$ is automatically in $\mathrm{Lip}(E_1)$: $|g(\eta) - g(\tilde{\eta})| \le 2||g||_\infty(\min\{\gamma_{x_1}, \dots, \gamma_{x_m}\})^{-1}||\eta - \tilde{\eta}||_{\gamma,1}$ (where $||g||_\infty := \sup\{|g(\eta)| : \eta \in E_1\}$). Thus the extension of $S(t)$ is unique.

It remains to prove (2.22) Let us first consider $f \in \mathrm{Lip}(E_1)$ and $\eta \in E_2$. Let $S_n \nearrow S$ and $(\xi^{(n)}(t))_{t\ge0}$ be as above, $(\xi(t))_{t\ge0} := \lim_n(\xi^{(n)}(t)_{t\ge0})$. Taking expectation in (2.12) in Lemma 1, c) yields

$$\mathbb{E}[f(\xi^{(n)}(t))] = f(\eta\mathbf{1}_{S_n}) + \int_0^t \mathbb{E}\big[Lf(\xi^{(n)}(s))\big]\, ds. \tag{2.24}$$

The lefthand side equals $S(t)f(\eta\mathbf{1}_{S_n})$, which converges to $S(t)f(\eta)$ by the first part of the proof. For the righthand side note that

$$|Lf(\xi^{(n)}(s))| \le C||\xi^{(n)}(s)||^2_{\gamma,2} \le C||\xi(s)||^2_{\gamma,2}$$

by (2.16) and the monotonicity properties of solutions of (2.9). Furthermore note that

$$\mathbb{E}\big[||\xi(s)||^2_{\gamma,2}\big] \le \liminf_n \mathbb{E}\big[||\xi^{(n)}(s)||^2_{\gamma,2}\big] \le \exp(C_2 s)||\eta||^2_{\gamma,2}$$

by Fatou's lemma and by (2.13). This allows to conclude that

$$\mathbb{E}\big[Lf(\xi^{(n)}(s))\big] \to \mathbb{E}\big[Lf(\xi(s))\big] = \big(S(s)Lf\big)(\eta) \quad \text{as } n \to \infty$$

by dominated convergence. Another application of the dominated convergence theorem (on the integral $\int_0^t \dots ds$ in (2.24)) yields (2.22).

In order to prove (2.22) for a function $f(\xi) = g(\xi_{x_1}, \dots, \xi_{x_m})$ which is constant on $\{\xi \in E_1 : \xi_{x_1} + \dots + \xi_{x_m} \ge M\}$ note that such an f is bounded and satisfies

$$|Lf(\eta)| \le C||\eta||_{\gamma,1}, \quad \eta \in E_1$$

for some suitable C, then argue as before.

Finally, (2.23) follows from (2.22) by differentiation because under the given conditions, the mapping $t \mapsto \big(S(t)Lf\big)(\eta)$ is continuous and locally bounded. □

Remark 3 a) Incidentally, we have proved that (2.9) has a strong solution starting from any $\eta \in E_1$. Uniqueness in law follows from the Markov property of solutions and the fact that the semigroup is determined by its restriction to E_{fin}.

b) The semigroup $(S(t))$ has the following continuity property: Let $\mu_n \in \mathcal{M}_1(E_1)$, $n \in \mathbb{N}$ with $\sup_{n,x} \int \xi_x^2\, \mu_n(d\xi) < \infty$ and $\mu_n \Rightarrow \mu$ as $n \to \infty$. Then we also have $\mu_n S(t) \Rightarrow \mu S(t)$ for any $t \ge 0$:

Let $\xi^{(n)}(\cdot)$ be the solution of (2.9) constructed in Lemma 2 with $\mathcal{L}(\xi^{(n)}(0)) = \mu_n$, $\xi(\cdot)$ the corresponding solution with initial distribution $\mathcal{L}(\xi(0)) = \mu$. Note that by assumption, the family of random variables ξ_x under μ_n, $x \in S$, $n \in \mathbb{N}$ is uniformly integrable. This allows to choose $\xi^{(n)}(0)$, $n \in \mathbb{N}$ and $\xi(0)$ in such a way that $\mathbb{E}\,||\xi^{(n)}(0) - \xi(0)||_{\gamma,1} \to 0$. Arguing as in the proof of Lemma 2 we see that

$$\mathbb{E}\,||\xi^{(n)}(t) - \xi(t)||_{\gamma,1} \le \exp(t\kappa M)\mathbb{E}\,||\xi^{(n)}(0) - \xi(0)||_{\gamma,1} \to 0.$$

We will have occasion to use the coupled process $(\xi, \tilde{\xi})$ corresponding to the generator $L^{(2)}$ again in section 2.4. Then, the following proposition, which can be proved along the same lines, will be helpful:

Proposition 2 *The analogue of Proposition 1 holds true for the coupled process* $(\xi, \tilde{\xi})$.

Note that Assumption A guarantees that the model is *attractive*: $L^{(2)}$ provides a monotone coupling, that is starting from $\xi_\cdot(0) \leq \tilde{\xi}_\cdot(0)$ we have

$$\mathbb{P}(\xi_x(t) \leq \tilde{\xi}_x(t) \text{ for all } x \text{ and } t) = 1.$$

In order to see this we can either use Lemma 1, b) and an approximation argument, or observe that $f : E_1 \times E_1 \to \mathbb{R}_+$, $f(\xi, \tilde{\xi}) = \mathbf{1}(\xi_x \leq \tilde{\xi}_x \, \forall x)$ satisfies $L^{(2)} f(\xi, \tilde{\xi}) = 0$ for any pair $(\xi, \tilde{\xi})$ with $\xi \leq \tilde{\xi}$.

The next lemma implies that if the initial condition has uniformly bounded n-th moments and the offspring distribution has a finite n-th moment, the system will preserve finite n-th moments over any finite time horizon.

Lemma 3 *If* $\sum_k \nu_k k^m < \infty$ *for some* $m \in \mathbb{N}$, *then there is a* $C_m < \infty$ *such that*

$$\mathbb{E}\left[||\xi(t)||_{\gamma,m}^m\right] \leq \exp(C_m t)\mathbb{E}\left[||\xi(0)||_{\gamma,m}^m\right],$$

in particular $\mathbb{E}\left[||\xi(0)||_{\gamma,m}^m\right] < \infty$ *implies* $||\xi(t)||_{\gamma,m}^m < \infty$ *for all times.*

Proof. Consider first a deterministic initial condition η with $||\eta||_{\gamma,m} < \infty$, let $S_n \nearrow S$ be an exhausting sequence of finite subsets. Then

$$\begin{aligned} \mathbb{E}_\eta\left[||\xi(t)||_{\gamma,m}^m\right] &= \lim_{n\to\infty} \mathbb{E}_{\eta \mathbf{1}_{S_n}}\left[||\xi(t)||_{\gamma,m}^m\right] \\ &\leq \lim_{n\to\infty} \exp(C_m t)||\eta \mathbf{1}_{S_n}||_{\gamma,m}^m = \exp(C_m t)||\eta \mathbf{1}_{S_n}||_{\gamma,m}^m \end{aligned}$$

by (2.21) and Lemma 1, d). The general case follows by conditioning on the initial configuration. □

2.3 Moment computations

We assume further on that the transition matrix p has a strictly positive invariant measure α. Let $q_{xy} := (\alpha_y p_{yx})/\alpha_x$ be the dual transition matrix with respect to α. We denote the semigroup generated by $\kappa(p-I)$, that is the transition semigroup of a continuous time Markov chain X on S with jump rate matrix $\kappa(p-I)$, by $(p(t))_{t\geq 0} = \big((p_{xy}(t))_{x,y\in S}\big)_{t\geq 0}$, analogously we define $(q(t))_{t\geq 0}$.

Remark 4 If $S = \mathbb{Z}^d$ and $p_{xy} = p_{0,y-x}$ is shift invariant and irreducible then by the Choquet-Deny Theorem (see e.g. [23], Cor. II.7.2) any invariant measure α must be a constant multiple of counting measure. Thus in such a scenario q is just the transpose of p.

A system of independent p-random walks, that is an (E_1-valued) Markov process with generator L_{rw}, is dual with respect to the function given by (2.25) (in the sense of Liggett, [23], Definition II.3.1) to a system of independent q-random walks, see e.g. [13] and references in there:

For a configuration $\eta \in (\mathbb{N}_0)^S$ with $|\eta| := \sum_x \eta_x < \infty$ define the test function

$$F_\eta(\xi) := \prod_{x \in S} \alpha^{-\eta_x} [\xi_x]_{\eta_x}, \quad \xi \in (\mathbb{N}_0)^S, \tag{2.25}$$

where $[n]_k := n(n-1)\cdots(n-k+1)$, $[n]_0 := 1$. Observe that e.g. for $\alpha \equiv 1$ this factorial moment function has a combinatorial interpretation as the number of (ordered) $|\eta|$-tuples of particles one can pick from the ξ-configuration if one prescribes that η_x particles have to be chosen from site x. The reweighting with $\alpha^{-\eta_x}$ accounts for spatial inhomogeneity. A key observation is that

$$\big(L_{rw} F_\eta(\cdot)\big)(\xi) = \big(\widetilde{L}_{rw} F_\cdot(\xi)\big)(\eta) \qquad \text{for } \xi, \eta \in (\mathbb{N}_0)^S,\ |\eta| < \infty,\ \xi \in E_1, \tag{2.26}$$

where $\widetilde{L}_{rw}$ is is the generator of a system of independent q-random walks. In order to see this write

$$\begin{aligned}
\big(L_{rw} F_\eta(\cdot)\big)(\xi) &= \sum_{x,y} \xi_x p_{xy} \Big\{ F_\eta(\xi^{(x,y)}) - F_\eta(\xi) \Big\} \\
&= \sum_{x,y} \xi_x p_{xy} \Big\{ \prod_z \alpha^{-\eta_z} [\xi_z]_{\eta_z} \Big\} \left(\frac{\xi_x - \eta_x}{\xi_x} \frac{\xi_y + 1}{\xi_y - \eta_y + 1} - 1 \right) \\
&= \sum_{x,y} p_{xy} \Big\{ \prod_z \alpha^{-\eta_z} [\xi_z]_{\eta_z} \Big\} \left(\eta_y \frac{\xi_x - \eta_x}{\xi_y - \eta_y + 1} - \eta_x \right) \\
&= \sum_{x,y} p_{xy} \frac{\alpha_y}{\alpha_x} \eta_y \Big\{ \prod_z \alpha^{-\eta_z^{(y,x)}} [\xi_z]_{\eta_z^{(y,x)}} \Big\} - \Big\{ \prod_z \alpha^{-\eta_z} [\xi_z]_{\eta_z} \Big\} \sum_{x,y} p_{xy} \eta_x \\
&= \sum_{x,y} \eta_y q_{yx} F_{\eta^{(y,x)}}(\xi) - F_\eta(\xi) \sum_{x,y} q_{yx} \eta_y = \big(\widetilde{L}_{rw} F_\cdot(\xi)\big)(\eta).
\end{aligned}$$

Observe that the sums are all well defined because the condition $\xi \in E_1$ ensures that $\sum_x \xi_x \sum_{y \in B} p_{xy} < \infty$ for any finite set B, thus in particular for $B := \{z : \eta_z > 0\}$.

We use the duality relation (2.26) together with perturbation formulas for Markov semigroups to represent (the first and the second) moments of $\xi_x(t)$ as expected values of functionals of independent p-chains. This idea has been used before for other interacting particle systems and interacting systems of diffusions, see e.g. [24], [5], [13].

Lemma 4 *a) Assume that $\mathbb{E}||\xi(0)||_{\gamma,1} < \infty$. Then*

$$\mathbb{E}\, \xi_x(t) = \sum_z \frac{\alpha_x}{\alpha_z} q_{xz}(t) \mathbb{E}\, \xi_z(0) = \sum_y \mathbb{E}\, \xi_y(0) p_{yx}(t). \tag{2.27}$$

b) Assume that $\mathbb{E}||\xi(0)||^2_{\gamma,2} < \infty$. Then for all $x, y \in S$

$$\begin{aligned}
\mathbb{E}\big[\xi_x(t)(\xi_y(t) - \delta_{xy})\big] \;&=\; \sum_{u,v} \frac{\alpha_x}{\alpha_u} q_{xu}(t) \frac{\alpha_y}{\alpha_v} q_{yv}(t) \mathbb{E}\big[\xi_u(0)(\xi_v(0) - \delta_{uv})\big] \\
&\quad + \mathrm{Var}(\nu) \int_0^t \sum_u \frac{\alpha_x \alpha_y}{\alpha_u^2} q_{xu}(s) q_{yu}(s) \mathbb{E}\, \sigma(\xi_u(t-s))\, ds
\end{aligned} \tag{2.28}$$

$$= \sum_{u,v} \mathbb{E}\big[\xi_u(0)(\xi_v(0) - \delta_{uv})\big] p_{ux}(t) p_{vy}(t) + \sum_u \mathrm{Var}(\nu) \int_0^t \mathbb{E}\, \sigma(\xi_u(t-s)) p_{ux}(s) p_{uy}(s)\, ds$$

Remark 5 1) These formulas have an interpretation in terms of the genealogy of the branching particle system: In *a)* we ask about the expected number of particles at x at time t. As the expected number of offspring per branching event is always one, each summand $\mathbb{E}\,\xi_y(0)p_{yx}(t)$ represents the expected number of particles at x whose ancestor at time 0 lived at y.
In *b)* we compute the expected number of pairs of different particles we can build at time t by choosing one from x and one from y. The "+" on the righthand side combines two possibilities: Either we choose two unrelated particles, then we have to decompose according to the position of their respective ancestors at time 0. The integral gives the expected number of pairs of related particles.

2) Unlike the "classical" case $\sigma(k) = c \times k$ corresponding to independent branching there is in general no closed form solution for (2.28). As we have no special assumptions on the functional form of σ we cannot hope to obtain an equation that is expressed entirely in $\mathbb{E}\,F_\eta(\xi(t))$, where F_η is as in (2.25). Nonetheless together with assumption (2.2), equation (2.28) does give useful information that we will exploit in section 2.4.

3) For $n \geq 3$, if ν has an n-th moment and $\mathbb{E}||\xi(0)||^n_{\gamma,n} < \infty$, the n-th (mixed factorial) moments of $\xi(t)$ will be finite by Lemma 3. We can in principle use the same ideas to obtain expressions for them involving n independent p-chains. We have refrained from stating those here as they get more complex for increasing n and we have no further use for them in this thesis.

4) Observe that for $\sigma(k) = c \times k^2$, we obtain again systems of equations involving mixed moments. Then we can express n-th moments of $\xi(t)$ in terms of exponential functionals of collision times of n independent p-chains. This is considered by Greven and den Hollander in [15] for the parabolic Anderson model, a system of interacting diffusions which is a "superprocess" limit of ξ in the case of a quadratic function σ. They prove that there is non-classical behaviour: Even if a non-trivial limit exists, the limit system only has moments up to order $n_0(c)$ where the threshold depends on c, the variance of the driving Brownian motions. We expect analogous behaviour of the moments of ξ in this case, but we have not pursued this issue further.

Proof of Lemma 4. a) Let $F_x(\xi) := \alpha_x^{-1}\xi_x$, $f_x(t) := \mathbb{E}\,F_x(\xi(t))$. Using (2.1), (2.26) and the fact that $\sum k\nu_k = 1$ we see that

$$(L\,F_x(\cdot))(\xi) = \kappa(\widetilde{L}^{(1)}F_\cdot(\xi))_x + 0 \quad \text{for } x \in S,\, \xi \in E_1$$

where $\kappa\widetilde{L}^{(1)}$ is the generator of $(q(t))_{t\geq 0}$, the one-particle semigroup with jump rate matrix $\kappa(q - I)$. Hence

$$\frac{\partial}{\partial t} f_x(t) = \frac{\partial}{\partial t}\mathbb{E}F_x(\xi(t)) = \kappa(\widetilde{L}^{(1)}\,\mathbb{E}\,F_\cdot(\xi_t))_x = \kappa(\widetilde{L}^{(1)}f_\cdot(t))_x, \quad x \in S.$$

This system with given initial condition $f_\cdot(0) = \mathbb{E}\,\xi_\cdot(0)$ is (uniquely) solved by

$$f_x(t) = \sum_z q_{xz}(t)\mathbb{E}\,\xi_z(0),$$

yielding (2.27).

b) Defining $F_{x,y}(\xi) := \alpha_x^{-1}\alpha_y^{-1}\xi_x(\xi_y - \delta_{xy})$ and $f_{x,y}(t) := \mathbb{E}\,F_{x,y}(\xi(t))$ we similarly see that

$$(L\,F_{x,y}(\cdot))(\xi) \quad = \quad \kappa(\widetilde{L}^{(2)}F_{\cdot,\cdot}(\xi))_{xy} + \alpha_x^{-2}\mathrm{Var}(\nu)\delta_{xy}\sigma(\xi_x(t))$$

where $\kappa\widetilde{L}^{(2)}$ is the generator of a pair of independent q-chains (for the second term observe that $\sum_k \nu_k\big((\xi_x+k-1)(\xi_x+k-2)-\xi_x(\xi_x-1)\big) = \mathrm{Var}(\nu)$ because $\sum k\nu_k = 1$). Let us first consider a deterministic, finite initial condition $\xi(0)$. Then we can conclude from the above and Proposition 1 that

$$\frac{\partial}{\partial t} f_{xy}(t) = \kappa(\widetilde{L}^{(2)} f_{\cdot,\cdot}(t))_{xy} + \mathrm{Var}(\nu)\delta_{xy}\mathbb{E}\sigma(\xi_x(t)), \quad x, y \in S.$$

A perturbation formula (see e.g. Pazy, [28], Thm. 6.1.2) yields (2.28) for this initial condition. We approximate a general deterministic $\xi(0) \in E_2$ by a sequence $\xi\mathbf{1}_{S_n}$ and use monotone resp. dominated convergence (observe that $\mathbb{E}_\xi\xi_x(t)^2 < \infty$ for $\xi \in E_2$) on each of the terms in (2.28). Finally averaging over the distribution of $\xi(0)$ yields the general case. □

2.3.1 The shift-invariant case

Here we assume that $S = \mathbb{Z}^d$, $p_{xy} = p_{0,y-x}$, i.e. p is of random walk type, and $\mathcal{L}(\xi(0))$ is invariant under shifts. This assumption allows more explicit calculations. As remarked before, in this situation $q_{xy} = p_{yx}$ is just the transpose of p. We will also need the symmetrized transition rates $\hat{p}_{xy} := (p_{xy} + p_{yx})/2$ as well as the semigroup $(\hat{p}_{xy}(t))$ corresponding to $\hat{p}$-motion at rate κ. This is again a random walk. In fact, it can be represented as the difference of two p-random walks running at speed $1/2$. We will sometimes drop the first subscript and write $p_x := p_{0x}$, analogously $q_x, \hat{p}_x$.

If in this scenario $\mathbb{E}_\mu\xi_x(0) = \theta$ for some $\theta \in \mathbb{R}_+$ then we have $\mathbb{E}_\mu\xi_x(t) \equiv \theta$ for all t by Lemma 4. If furthermore the symmetrized motion $\hat{p}$ is transient the system will preserve globally bounded second moments, provided that the branching rate function does not grow too fast:

Lemma 5 *Let $p_{xy} = p_{0,y-x}$ and $\mathcal{L}(\xi(0))$ be shift invariant with $\mathbb{E}[\xi_0(0)] := \theta$, $\mathbb{E}[\xi_0(0)^2] < \infty$. If $\hat{p}$ is transient and*

$$\limsup_{k\to\infty} \frac{\sigma(k)}{k^2} < \frac{2}{\mathrm{Var}(\nu)\hat{G}_0(\infty)}, \tag{2.29}$$

where $\hat{G}_0(\infty) := \int_0^\infty \hat{p}_0(s)\,ds$ is the Green function of $\hat{p}$. Then

$$\sup_{t\geq 0} \mathbb{E}[\xi_x(t)^2] < \infty.$$

Proof. By (2.29) there exist $C_1 > 0$ and $0 \leq C_2 < 2/(\mathrm{Var}(\nu)\hat{G}_0(\infty))$ such that $\sigma(k) \leq C_1 + C_2k^2$. Denote $h(t) := \mathbb{E}[\xi_0(t)^2]$. Using (2.28) and shift-invariance we can estimate

$$\begin{aligned} h(t) &= \theta + \mathbb{E}\,\xi_0(t)(\xi_0(t)-1) \\ &\leq \theta + \sup_{u,v}\mathbb{E}\,\xi_u(0)(\xi_v(0)-\delta_{uv}) + \mathrm{Var}(\nu)\int_0^t \sum_u q_{0u}(s)^2\mathbb{E}\,\sigma(\xi_u(t-s))\,ds \\ &\leq C + \int_0^t h(t-s)\,\mathrm{Var}(\nu)C_2\hat{p}_0(2s)ds = C + \int_0^t h(t-s)\,F(ds), \end{aligned} \tag{2.30}$$

where $F(ds) := \mathrm{Var}(\nu)C_2\hat{p}_0(2s)ds$ and C is a suitable constant. Now e.g. by [10], Thm. VI.6.1, the solution $z(\cdot)$ of the (renewal) equation corresponding to (2.30) is given by $z(t) = C\cdot U([0,t])$, where $U = \sum_{n=0}^\infty F^{*n}$. Observe that $F([0,\infty)) < 1$ by (2.29), hence $U([0,\infty)) = 1/(1-F([0,\infty))) < \infty$ and $\sup z(t) < \infty$. Finally observe that $z(t) \geq h(t)$ for all t, as the difference $z-h$ starts non-negative at 0 and satisfies an obvious renewal inequality, forcing it to remain non-negative for all times. □

Remark 6 Condition (2.29) is sharp in the following sense: If $\sigma(k) = C_\sigma k^2$ with $C_\sigma \geq 2/(\mathrm{Var}(\nu)\hat{G}_0(\infty))$ we have $\mathbb{E}[\xi_x(t)^2] \to \infty$ as $t \to \infty$. In this case we have

$$h(t) \geq \theta + C_\sigma \mathrm{Var}(\nu) \int_0^t \sum_u q_{0u}(s)^2 h(t-s)\, ds = \theta + \int_0^t h(t-s)\, F(ds)$$

with $F(ds) = C_\sigma \mathrm{Var}(\nu)\hat{p}_0(2s)ds$, which satisfies $F((0,\infty)) \geq 1$ by the assumption on C_σ. The second moment diverges exponentially if C_σ is strictly larger than $2/(\mathrm{Var}(\nu)\hat{G}_0(\infty))$, as then $F((0,\infty)) > 1$. It is unclear to me if the same holds if we only require $\sigma(k) \sim ck^2$ as $k \to \infty$ where $c > 2/(\mathrm{Var}(\nu)\hat{G}_0(\infty))$.

Definition We denote by $\mathcal{R}_\theta$ the set of all shift-invariant probability measures μ on $(\mathbb{N}_0)^S$ with $\int \xi_0\, d\mu = \theta$ and $\int (\xi_0)^2 d\mu < \infty$ which are "asymptotically uncorrelated" in the following sense:

$$\lim_{t\to\infty} \sum_{x,y\in\mathbb{Z}^d} q_{0x}(t) q_{0y}(t) \int \xi_x \xi_y d\mu = \theta^2. \tag{2.31}$$

Observe that this amounts to $\sum_x q_{0x}(t)\xi_x \to \theta$ as $t \to \infty$ in $L^2(\mu)$. Any shift-ergodic μ with $\int \xi_0\, d\mu = \theta$ and $\int (\xi_0)^2 d\mu < \infty$ lies in $\mathcal{R}_\theta$, see e.g. [24], Lemma 5.2.

In the situation of Lemma 5, $\mathcal{R}_\theta$ will be a good set for limits to concentrate on. This will be very helpful in the next section where we combine it with coupling arguments to find equilibrium states for $(\xi(t))_{t\geq 0}$.

Lemma 6 *Let $p_{xy} = p_{0,y-x}$ be of random walk type, $\mu \in \mathcal{R}_\theta$ and assume that (2.29) holds. Then*
a) $\mu S(t) \in \mathcal{R}_\theta$ for all $t \geq 0$.
b) If $\hat{p}$ is transient, σ satisfies (2.29) and $\mu S(t_n) \Rightarrow \mu_\infty$ for some sequence $t_n \uparrow \infty$ then also $\mu_\infty \in \mathcal{R}_\theta$.

Proof. Fix $t \geq 0$ for the moment. Using Lemma 4 we can compute (with $F_{x,y}(\xi)$ as in the proof of Lemma 4 and $\alpha_x \equiv 1$)

$$\begin{aligned}
\mathbb{E}_\mu\Big(\sum_x q_{0x}(T)(\xi_x(t)-\theta)\Big)^2 &= \sum_{x,y\in\mathbb{Z}^d} q_{0x}(T)q_{0y}(T)\left\{\mathbb{E}\, F_{x,y}(\xi(t)) + \delta_{xy}\theta - \theta^2\right\} \\
&= \theta\hat{p}_{00}(2T) + \sum_{x,y} q_{0x}(T)q_{0y}(T)\Big\{\sum_{u,v} q_{xu}(t)q_{yv}(t)\left[\mathbb{E}_\mu F_{u,v}(\xi(0)) - \theta^2\right] \\
&\qquad + \gamma^2 \int_0^t \sum_u q_{xu}(s)q_{yu}(s)\mathbb{E}_\mu\, \sigma(\xi(t-s))\, ds\Big\} \\
&= \theta\hat{p}_{00}(2T) + \sum_{u,v} q_{0u}(T+t)q_{0v}(T+t)\left[\mathbb{E}_\mu F_{u,v}(\xi(0)) - \theta^2\right] \qquad (2.32) \\
&\quad + \gamma^2 \int_0^t \sum_u q_{0u}(T+s)^2 \mathbb{E}_\mu\, \sigma(\xi(t-s))\, ds.
\end{aligned}$$

The first term on the righthand side of (2.32) tends to 0 as $T \to \infty$ because $\hat{p}$ is transient or null recurrent, and the second term tends to 0 because $\mu \in \mathcal{R}_\theta$ by assumption. The third term is at most

$$\gamma^2 \left(C_1 + C_2 \sup_{0 \le r \le t} \mathbb{E}[\xi_0(r)^2] \right) \int_0^t \hat{p}_{00}(2T + 2s)\, ds,$$

which is finite by Lemma 3 and tends to 0 as $T \to \infty$ because $\lim_{s \to \infty} \hat{p}_{00}(s) = 0$. This proves part a).

Now assume additionally that $\mu S(t_n) \to \mu_\infty$ in finite dimensional distributions along some sequence $t_n \uparrow \infty$. Fatou's lemma shows that for any fixed $T > 0$ we have

$$\int \left(\sum_x q_{0x}(T)(\xi_x - \theta) \right)^2 \mu_\infty(d\xi) \le \liminf_n \mathbb{E}_\mu \left(\sum_x q_{0x}(T)(\xi_x(t_n) - \theta) \right)^2.$$

The proof of part a) shows that the righthand side is at most

$$\theta \hat{p}_{00}(2T) + 0 + \gamma^2 \left(C_1 + C_2 \sup_{0 \le r} \mathbb{E}[\xi_0(r)^2] \right) \int_0^\infty \hat{p}_{00}(2T + 2s)\, ds$$

(note that $\lim_n \sum_{u,v} q_{0u}(T + t_n) q_{0v}(T + t_n) \left[\mathbb{E}_\mu F_{u,v}(\xi(0)) - \theta^2 \right] = 0$ because $\mu \in \mathcal{R}_\theta$) which is finite by Lemma 5 and tends to 0 as $T \to \infty$ by the transience of $\hat{p}$. □

2.4 Coupling in the regime of globally bounded second moments

In our setting a sucessful coupling – if available – is a useful tool to prove convergence to equilibrium and uniqueness of equilibria with given intensities. Let us consider a process $(\xi(t), \tilde{\xi}(t))_t$ on $E_1 \times E_1$ such that each component is a system of (state dependent) branching random walks as above and the two systems move their particles "in unison as much as possible". The coupled systems have already been considered in section 2.2, see (2.18) for the form of their (formal) generator $L^{(2)}$ and Proposition 2 for basic properties. We wish to prove that the coupling works, that is that the two systems really become similar in the long run. A well-established strategy, see e.g. [24] or [5], works as follows: attractivity can be used to show that in the long run, either ξ is above $\tilde{\xi}$ or vice versa. Then we exploit the fact that both have the same constant intensity θ to conclude that $\xi(\infty) = \tilde{\xi}(\infty)$.

Let $\Delta_x := \xi_x - \tilde{\xi}_x$, $F_x(\xi, \tilde{\xi}) := |\Delta_x|$. In the case that (2.4) holds we use the fact that $|a + 1| + |a - 1| - 2|a| = 0$ for all $a \in \mathbb{Z} \setminus \{0\}$, respectively in case of (2.5) we observe that $\sum_k \nu_k(|a+k-1|-|a|)=0$ for $a>0$. This allows to compute

$$\begin{aligned}
L^{(2)}F_z(\xi,\tilde\xi) &= \kappa\sum_{x,y}(\xi_x-\xi_x\wedge\tilde\xi_x)p_{xy}\big(F_z(\xi^{(x,y)},\tilde\xi)-F_z(\xi,\tilde\xi)\big)\\
&\quad+\kappa\sum_{x,y}(\tilde\xi_x-\xi_x\wedge\tilde\xi_x)p_{xy}\big(F_z(\xi,\tilde\xi^{(x,y)})-F_z(\xi,\tilde\xi)\big)\\
&= \kappa\sum_{x,y}\Delta_x\mathbf{1}(\Delta_x>0)p_{xy}\big(|\Delta_z+\delta_{yz}-\delta_{xz}|-|\Delta_z|\big)\\
&\quad+\kappa\sum_{x,y}(-\Delta_x)\mathbf{1}(\Delta_x<0)p_{xy}\big(|\Delta_z-\delta_{yz}+\delta_{xz}|-|\Delta_z|\big)\\
&= \kappa\sum_x\Big\{\Delta_x p_{xz}\mathrm{sgn}(\Delta_z)-\Delta_z p_{zx}\mathrm{sgn}(\Delta_z)+|\Delta_x|p_{xz}\mathbf{1}(\Delta_z=0)\Big\}\\
&= \kappa\sum_x p_{xz}(|\Delta_x|-|\Delta_z|)+\kappa\sum_x p_{xz}|\Delta_x|\big(\mathbf{1}(\Delta_z=0)-\mathbf{1}(\mathrm{sgn}(\Delta_x)\neq\mathrm{sgn}(\Delta_z))\big)\\
&\quad-\kappa\sum_x p_{xz}|\Delta_x|\mathbf{1}(\mathrm{sgn}(\Delta_x)\neq\mathrm{sgn}(\Delta_z)).
\end{aligned}$$

We used that $|a+\delta|-|a|=\delta\mathrm{sgn}(a)+|\delta|\mathbf{1}(a=0)$ for $a\in\mathbb{Z}$, $\delta\in\{-1,0,1\}$, where $\mathrm{sgn}(a):=\mathbf{1}(a>0)-\mathbf{1}(a<0)$.

Lemma 7 *Let $\mathcal{L}(\xi_\cdot(0),\tilde\xi_\cdot(0))$ be shift-invariant, $\mathbb{E}\,\xi_0(0)^2$, $\mathbb{E}\,\tilde\xi_0(0)^2<\infty$ and assume that $\hat p$ is transient and σ satisfies (2.29).*
a) $\mathbb{E}|\Delta_x(t)|$ is non-increasing in t and for all $x,y\in\mathbb{Z}^d$ we have

$$\mathbb{P}(\xi_x(t)>\tilde\xi_x(t),\xi_y(t)<\tilde\xi_y(t))\to 0\quad \textit{as } t\to\infty.$$

b) If $\mathcal{L}(\xi(0)),\mathcal{L}(\tilde\xi(0))\in\mathcal{R}_\theta$ for some $\theta\in\mathbb{R}_+$ the coupling is successful, i.e.

$$\mathbb{E}\,|\Delta_x(t)|\to 0 \textit{ as } t\to\infty\quad \textit{for all } x\in S.$$

Proof. a) From the above computation and shift-invariance we have

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\,|\Delta_z(t)| &= \kappa\sum_x p_{xz}\mathbb{E}\left[|\Delta_x(t)|\big(\mathbf{1}(\Delta_z(t)=0)-\mathbf{1}(\mathrm{sgn}(\Delta_x(t))\neq\mathrm{sgn}(\Delta_z(t)))\big)\right]\\
&\quad-\kappa\sum_x p_{xz}\mathbb{E}\left[|\Delta_x(t)|\mathbf{1}(\mathrm{sgn}(\Delta_x(t))\neq\mathrm{sgn}(\Delta_z(t)))\right]\\
&\le -\kappa\sum_x p_{xz}\mathbb{E}\left[|\Delta_x(t)|\mathbf{1}(\mathrm{sgn}(\Delta_x(t))\neq\mathrm{sgn}(\Delta_z(t)))\right]\le 0.
\end{aligned}$$

In particular, $c:=\lim_{t\to\infty}\mathbb{E}\,|\Delta_z(t)|\le\mathbb{E}[\xi_0(0)+\tilde\xi_0(0)]$ exists (and is independent of z by shift-invariance). If $c=0$ we are done, so let us assume $c>0$. Let $H_{xz}(\xi,\tilde\xi):=\mathbf{1}(\Delta_x>0,\Delta_z<0)$. As $(d/dt)\mathbb{E}\,|\Delta_z(t)|\le-\kappa p_{xz}\mathbb{E}\,H_{xz}(\xi(t),\tilde\xi(t))$ we conclude that

$$\int_0^\infty \mathbb{P}(\Delta_x(t)>0,\Delta_z(t)<0)\,dt<\infty \tag{2.33}$$

for any x,z with $p_{xz}>0$. One easily checks that

$$\begin{aligned}
\left|L^{(2)}H_{xz}(\xi,\tilde\xi)\right| &\le \kappa\sum_y\{|\Delta_x|p_{xy}+|\Delta_z|p_{zy}+|\Delta_y|(p_{yx}+p_{yz})\}\\
&\quad+\sigma(\xi_x)+\sigma(\tilde\xi_x)+\sigma(\xi_z)+\sigma(\tilde\xi_z),
\end{aligned}$$

hence (observe that $\sup_{t\geq 0} \mathbb{E}\,\sigma(\xi_x(t)) < \infty$ for all x by our assumptions on $\hat{p}$, $\sigma(\cdot)$, $\mathbb{E}\,\xi_x(0)^2 < \infty$ and Lemma 5)

$$\sup_{t\geq 0}\left|\frac{d}{dt}\mathbb{P}(\Delta_x(t) > 0, \Delta_z(t) < 0)\right| = \sup_{t\geq 0}\left|\mathbb{E}\,L^{(2)} H_{xz}(\xi(t), \tilde{\xi}(t))\right| < \infty,$$

which together with (2.33) implies our claim for this pair x, z. Now consider $x, y \in \mathbb{Z}^d$ with $p_{xy} = 0$, but assume that there exists some z satisfying $p_{xz}p_{zy} > 0$. Consider the events

$$A(t) := \left\{\xi_x(t) > \tilde{\xi}_x(t), \xi_y(t) < \tilde{\xi}_y(t), \xi_z(t) = \tilde{\xi}_z(t)\right\}. \tag{2.34}$$

If we can show that

$$\lim_{t\to\infty} \mathbb{P}(A(t)) = 0 \tag{2.35}$$

then we have shown that the claim is also true for x and y, as we already know from the above that it holds for $\{x, z\}$ and $\{y, z\}$. Then we use irreducibility of p and induction on the number of p-steps required to connect a given pair $x, y \in \mathbb{Z}^d$ to complete the proof of a).

It remains to prove (2.35), so let us assume that on the contrary $a := \limsup_t \mathbb{P}(A(t)) > 0$ was true. Consider also the events

$$B(t) := \left\{\xi_z(t) > \tilde{\xi}_z(t), \xi_y(t) < \tilde{\xi}_y(t)\right\}.$$

If (2.35) failed then we would also have

$$\limsup_{t\to\infty} \mathbb{P}(B(t)) > 0 \tag{2.36}$$

in contradiction to what we already know about the pair $\{y, z\}$. The intuitive idea behind (2.36) is as follows: Note that the bounded intensity of $(\xi, \tilde{\xi})$ guarantees that it is unlikely that a particle from outside jumps to $\{x, y, z\}$ during a very short time interval. Given that $A(t)$ has occured at time t there is a certain probability that the only transition we observe at the sites x, y and z in the time interval $[t, t+h]$ is that a ξ-surplus particle jumps from x to z, which implies that $B(t+h)$ occurs.

To give a formal argument let us denote the event we have just described by $C(t, h)$, and define for $K > 0$

$$D(t, K) := \left\{\sum_u (\xi_u(t) + \tilde{\xi}_u(t))(p_{ux} + p_{uy} + p_{uz}) \leq K,\ \xi_v(t) + \tilde{\xi}_v(t) \leq K \text{ for } v = x, y, z\right\}.$$

We have $\mathbb{P}(D(t, K)) \to 1$ as $K \to \infty$ uniformly in t (use e.g. the fact that $\mathbb{E}[\xi_x(t) + \tilde{\xi}_x(t)] \equiv 2\theta$ and the Markov inequality). Furthermore we have $C(t, h) \subset B(t+h)$ and $\mathbb{P}(C(t, h)|A(t) \cap D(t, K)) \geq \varepsilon(K, h) > 0$. Choose K large enough that $\mathbb{P}(D(t, K)^c) \leq a/4$ for all $t \geq 0$. Then we have for any t_* with $\mathbb{P}(A(t_*)) \geq a/2$

$$\begin{aligned}\mathbb{P}(A(t_*) \cap D(t_*, K)) &= \mathbb{P}(A(t_*)) - \mathbb{P}(A(t_*) \cap D(t_*, K)^c)\\ &\geq \mathbb{P}(A(t_*)) - \mathbb{P}(D(t_*, K)^c) \geq \mathbb{P}(A(t_*))/2,\end{aligned}$$

which implies

$$\begin{aligned}\mathbb{P}(C(t_*, h)|A(t_*)) &\geq \frac{\mathbb{P}(C(t_*, h) \cap A(t_*) \cap D(t_*, K))}{\mathbb{P}(A(t_*))}\\ &\geq 2^{-1}\frac{\mathbb{P}(C(t_*, h) \cap A(t_*) \cap D(t_*, K))}{\mathbb{P}(A(t_*) \cap D(t_*, K))} \geq \varepsilon(K, h)/2 > 0,\end{aligned}$$

yielding $\limsup_t \mathbb{P}(B(t+h)) \geq a\varepsilon(K,h)/4 > 0$. This is a contradiction, proving that (2.35) holds true.

b) Note that the distributions of $(\xi(t), \tilde{\xi}(t))$ are tight (as probability measures on $E_1 \times E_1$, with respect to the vague topology) because $\sup_{t,x} \mathbb{E}[\xi_x(t) + \tilde{\xi}_x(t)] = 2\theta < \infty$. In order to see this define for $K > 0$

$$\Lambda(K) := \left\{(\xi, \tilde{\xi}) \in E_1 \times E_1 : \xi_{z_i}, \tilde{\xi}_{z_i} \leq K2^i \text{ for } i = 1, 2, \dots \right\},$$

where $z_1, z_2, \dots$ is any enumeration of the state space. $\Lambda(K)$ is a compact subset of $E_1 \times E_1$ with respect to the product topology, and by the Markov inequality we have

$$\begin{aligned}
\mathbb{P}((\xi(t), \tilde{\xi}(t)) \notin \Lambda(K)) \quad &\leq \quad \sum_{i=1}^{\infty} \mathbb{P}(\xi_{z_i} > K2^i) + \mathbb{P}(\tilde{\xi}_{z_i} > K2^i) \\
&\leq \quad \sum_{i=1}^{\infty} 2\frac{\theta}{K2^i} = \frac{2\theta}{K}
\end{aligned}$$

Thus can choose a sequence $t_n \uparrow \infty$ such that $(\xi(t_n), \tilde{\xi}(t_n)) \to (\xi(\infty), \tilde{\xi}(\infty))$ in the sense of finite dimensional distributions. By Lemma 6 we have $\mathcal{L}(\xi(\infty)), \mathcal{L}(\tilde{\xi}(\infty)) \in \mathcal{R}_\theta$. Using shift-invariance and part a) we can compute

$$\begin{aligned}
\mathbb{E}\left|\xi_0(\infty) - \tilde{\xi}_0(\infty)\right| \quad &= \quad \mathbb{E}\left[\sum_x q_{0x}(T)\left|\xi_x(\infty) - \tilde{\xi}_x(\infty)\right|\right] \\
&= \quad \mathbb{E}\left|\sum_x q_{0x}(T)\left(\xi_x(\infty) - \tilde{\xi}_x(\infty)\right)\right| \\
&\leq \quad \mathbb{E}\left|\sum_x q_{0x}(T)\big(\xi_x(\infty) - \theta\big)\right| + \mathbb{E}\left|\sum_x q_{0x}(T)\big(\tilde{\xi}_x(\infty) - \theta\big)\right|.
\end{aligned}$$

Letting $T \to \infty$ we obtain $\mathbb{E}|\xi_0(\infty) - \tilde{\xi}_0(\infty)| = 0$ because $\mathcal{L}(\xi(\infty)), \mathcal{L}(\tilde{\xi}(\infty)) \in \mathcal{R}_\theta$ both have asymptotic density θ (in an L^2-sense, which implies the same in L^1). Finally observe that $\{|\Delta_0(t)| : t \geq 0\}$ is uniformly integrable (we are in a regime with globally bounded second moments by Lemma 5), proving $\lim_{n\to\infty} \mathbb{E}|\Delta_0(t_n)| = \mathbb{E}|\Delta_0(\infty)| = 0$. $\square$

These ingredients allow us to clarify the long-time behaviour of $\xi(t)$ in the regime of globally bounded second moments:

Proposition 3 *Assume that $\hat{p}$ is transient and σ satisfies (2.29). For any $\theta \geq 0$*

$$\nu_\theta = \lim_{t\to\infty} \mathrm{Poi}(\underline{\theta})S(t) \in \mathcal{R}_\theta \tag{2.37}$$

exists. ν_θ is a shift-invariant equilibrium and satisfies $\int \xi_x \, d\nu_\theta = \theta$, $\int (\xi_x)^2 \, d\nu_\theta < \infty$. ν_θ is stochastically smaller than $\nu_{\theta'}$ for $\theta \leq \theta'$. We have

$$\lim_{t\to\infty} \mu S(t) = \nu_\theta \quad \textit{for all } \mu \in \mathcal{R}_\theta.$$

Proof. Observing that $\int \xi_x \mathrm{Poi}(\underline{\theta})S(t)(d\xi) \equiv \theta$ we see that the family $\mathrm{Poi}(\underline{\theta})S(t)$, $t \in \mathbb{R}_+$ is tight, so there is a sequence $t_n \nearrow \infty$ such that $\nu_\theta = \lim_{n\to\infty} \mathrm{Poi}(\underline{\theta})S(t_n)$ exists. $\nu_\theta \in \mathcal{R}_\theta$ by Lemma 6 b). In order to prove that ν_θ is invariant fix $h > 0$ and put $\mu := \mathrm{Poi}(\underline{\theta})S(h)$ ($\in \mathcal{R}_\theta$ by Lemma 6 a)). Lemma 7 b) with $(\xi, \tilde{\xi})$ starting from $\mu \otimes \mathrm{Poi}(\underline{\theta})$ shows that

$\mu S(t_n) \Rightarrow \nu_\theta$ as $n \to \infty$, on the other hand we see from Lemma 5 and Remark 3 b) that $\mu S(t_n) = (\mathrm{Poi}(\underline{\theta})S(t_n))S(h) \Rightarrow \nu_\theta S(h)$, proving that ν_θ is invariant.

Now assume that $\nu'_\theta := \lim_n \mathrm{Poi}(\underline{\theta}S(t'_n)$ along some other sequence $t'_n \nearrow \infty$. Arguing as above we see that $\nu'_\theta \in \mathcal{R}_\theta$ is $S(t)$-invariant. Another application of Lemma 7 b), starting from $\nu_\theta \otimes \nu'_\theta$ shows that $\nu_\theta = \nu'_\theta$ and hence that the limit in (2.37) exists.

We have $\sup_t \int \xi_x^2 \,\mathrm{Poi}(\underline{\theta}S(t)(d\xi) < \infty$ by Lemma 5, which implies that ν_θ has full intensity θ and bounded second moments. Finally, another application of Lemma 7 shows that $\mu S(t) \Rightarrow \nu_\theta$ for any $\mu \in \mathcal{R}_\theta$. $\square$

Let us remark that we strongly expect ν_θ, which is a limit of branching particle systems, to have positive correlations in the sense of [23], chap. 2. If this was the case, we could proceed as in the proof of Corollary 2 in chapter 4 and show that ν_θ is spatially mixing. This would show that any shift-ergodic equilibrium ν of $(\xi(t))$ with finite intensity is necessarily some ν_θ.

2.5 Local size-biasing

2.5.1 Discrete generations

Here we look at a discrete-time analogue of the branching systems considered so far. Apart from our standing belief that the big picture should not depend on implementation details, discrete-time systems will be used to approximate the continuous-time system in order to transfer the structural result about local size-biasing to the continuous-time world, see Propositions 4 and 5 below.

Thus, in this subsection we play the following game: Given the configuration $\eta(n) \in E_1$ at time n, the next configuration $\eta(n+1)$ arises as follows: At site x a branching event occurs with probability $s(\eta_x(n))$. In this case, one particle – uniformly chosen among those at x – is replaced by K offspring, otherwise the number remains unchanged. Then all particles take an independent step according to transition matrix r_{xy} to form the $(n+1)$-th generation. Here $s : \mathbb{N}_0 \to [0,1]$ is the branching probability function (observe that at most one particle per site and time-step can branch in this model), $\mathcal{L}(K) = \nu$ is the offspring law as before, and (r_{xy}) is a stochastic matrix on S. Formally define

$$M(n,x,k) := k + \mathbf{1}(U_n(x) \le s(k))(K_x(n) - 1),$$

the number of offspring of k particles located at x in generation n,

$$\varkappa^{(n)}(x,k) := \sum_{i=1}^{M(n,x,k)} \delta_{Y_x(n,i)}, \quad \text{and let}$$

$$\eta(n+1) = \sum_x \varkappa^{(n)}(x, \eta_x(n)), \tag{2.38}$$

with $U_x(n)$, $x \in S$, $n = 0,1,\dots$ i.i.d. uniform$([0,1])$, $K_x(n)$ independent with law ν, $Y_x(n,i)$ independent with $\mathbb{P}(Y_x(\cdot,\cdot) = y) = r_{xy}$. The random population $\eta(n+1)$ is well defined and lies in E_1 almost surely for $\eta(n) \in E_1$, if r is compatible with γ from the definition of E_1 in the sense that (2.6) also holds if we replace p by r. As $\sum_k k\nu_k = 1$ we have $\mathbb{E}\,\eta_x(n) = \sum_y \mathbb{E}\,\eta_y(m) r_{yx}^{(n-m)}$ for $0 \le m \le n$.

We are interested in the behaviour of $\hat{\eta}^{(x)}(n)$, the locally size-biased $\eta(n)$, with distribution given by $\mathbb{E}\,F(\hat{\eta}^{(x)}(n)) = \mathbb{E}[\eta_x(n)F(\eta(n))]/\mathbb{E}[\eta_x(n)]$ – technically speaking,

$\mathcal{L}(\hat\eta^{(x)}(n))$ is the Palm measure (with respect to $x \in S$) of $\mathcal{L}(\eta(n))$. Note that the following considerations assume a fixed n, we do not consider $\hat\eta^{(x)}(n)$ as a process in n (in forward time). The following lemma is the key step for a stochastic representation of $\hat\eta^{(x)}(n)$.

Lemma 8 *Write*

$$\begin{aligned}\widehat{\varkappa}^{x,(n)}(y,k) \quad &:= \quad \delta_x + \mathbf{1}(V_y(n) \le \frac{1}{k}) \sum_{i=1}^{\tilde M(n,y,k)} \delta_{Y_y(n,i)} \\ &\qquad\qquad + \mathbf{1}(V_y(n) > \frac{1}{k}) \sum_{i=1}^{M(n,y,k)} \delta_{Y_y(n,i)}\end{aligned}$$

for the offspring of k particles at y in generation n, locally size-biased in x. Here, $M(n,y,k)$ is as above and

$$\tilde M(n,x,k) := k + \mathbf{1}(U_n(x) \le s(k))(\hat K_x(n) - 1)$$

where $V_y(n)$ is uniform$([0,1])$, $\hat K_x(n)$ is distributed according to the size-biasing of K, and the remaining ingredients are as above (and all independent). Fix $x \in S$. Then for any bounded, measurable function F we have

$$\begin{aligned}&\frac{1}{\mathbb{E}\,\eta_x(n+1)}\mathbb{E}[\eta_x(n+1)F(\eta(n+1))] \\ &\quad = \sum_y \frac{\mathbb{E}\,\eta_y(n) r_{yx}}{\mathbb{E}\,\eta_x(n+1)} \times \mathbb{E}\left[\frac{\eta_y(n)}{\mathbb{E}\,\eta_y(n)} F\Big(\widehat{\varkappa}^{x,(n)}(y,\eta_y(n)) + \sum_{z\neq y} \varkappa^{(n)}(z,\eta_z(n))\Big)\right].\end{aligned}$$

In words, the configuration $\hat\eta^{(x)}(n+1)$ can be constructed in the following manner: pick a site y with probability $(\mathbb{E}\,\eta_y(n) r_{yx})/\mathbb{E}\,\eta_x(n+1)$. This will be the site the ancestor of the selected particle in x came from. Given y choose a configuration with the law of $\hat\eta^{(y)}(n)$. Its particles in $z \neq y$ branch (with probability depending on the local number) and move independently as above. One of the, say, k particles at y is the "selected particle". If a branching occurs at y, it is with probability $1/k$ the selected particle that produces offspring. In this case, there is a size-biased number of children. With probability $1-1/k$ the branching happens to one of the remaining individuals who use the usual offspring law. Finally, everybody (except the "selected one", who automatically moves to x) has to make their independent step according to $r_{y\cdot}$.

Proof of Lemma 8. It suffices to consider $F(\eta) = \exp(-\langle g, \eta\rangle)$ for some $g : S \to \mathbb{R}_+$. We have

$$\begin{aligned}&\mathbb{E}\left[\eta_x(n+1)F(\eta(n+1)\right] \\ &\quad = \sum_y \mathbb{E}\left[\left(\varkappa^{(n)}(y,\eta_y(n))\right)_x \times \prod_z \exp\left(-\langle g, \varkappa^{(n)}(z,\eta_z(n))\rangle\right)\right] \\ &\quad = \sum_y \mathbb{E}\Bigg[\mathbb{E}\left[\left(\varkappa^{(n)}(y,\eta_y(n))\right)_x \times \exp\left(-\langle g, \varkappa^{(n)}(y,\eta_y(n))\rangle\right)\Big|\, \eta(n)\right] \\ &\qquad\qquad \times \prod_{z\neq y} \mathbb{E}\left[\exp\left(-\langle g, \varkappa^{(n)}(z,\eta_z(n))\rangle\right)\Big|\, \eta(n)\right]\Bigg] \\ &\quad = \sum_y \mathbb{E}\left[F_y(\eta_y(n)) \prod_{z\neq y} \mathbb{E}\left[\exp\left(-\langle g, \varkappa^{(n)}(z,\eta_z(n))\rangle\right)\Big|\, \eta(n)\right]\right],\end{aligned}$$

where $F_y(k) := \mathbb{E}\left[\left(\varkappa^{(n)}(y,k)\right)_x \times \exp\left(-\langle g, \varkappa^{(n)}(y,k)\rangle\right)\right]$. Let $L := k + \mathbf{1}(U_n(y) \le s(k))(K_y(n)-1)$, then

$$\begin{aligned}
F_y(k) &= \sum_{l=1}^{\infty} \mathbb{P}(L=l)\mathbb{E}\left[\sum_{j=1}^{l} \mathbf{1}(Y_y(n,j)=x)\exp\Big(-\sum_{m=1}^{l} g_{Y_y(n,m)}\Big)\right] \\
&= \sum_{l=1}^{\infty} \mathbb{P}(L=l) l \times r_{yx} e^{-g_x}\Big(\sum_z r_{yz} e^{-g_z}\Big)^{l-1} \\
&= k r_{yx} \sum_{l=1}^{\infty} \mathbb{P}(\hat{L}=l)\mathbb{E}\left[\exp\Big(-\langle g, \delta_x + \sum_{m=1}^{l-1} \delta_{Y_y(n,m)}\rangle\Big)\right] \\
&= k r_{yx} \mathbb{E}\exp\Big(-\langle g, \widehat{\varkappa}^{x,(n)}(y,k)\Big).
\end{aligned}$$

Observe that $\mathbb{E}\, L = k$ and the size-biased $\hat{L}$ satisfies

$$\mathbb{P}(\hat{L} = k+m) = \begin{cases} 1 - s(k) + s(k)\nu_1 & \text{if } m = 0 \\ s(k)\frac{k+m}{k}\nu_{m+1} & \text{if } m \in \{-1, 1, 2, \ldots\}, \end{cases}$$

hence

$$\hat{L} =_d k + \mathbf{1}(U \le s(k))\big\{\mathbf{1}(V \le \frac{1}{k})(\hat{K}-1) + \mathbf{1}(V > \frac{1}{k})(K-1)\big\},$$

where U, V, K and $\hat{K}$ are as in the definition of $\widehat{\varkappa}^{x,(n)}(y,k)$. Thus we see that

$$\begin{aligned}
&\mathbb{E}\left[\eta_x(n+1)F(\eta(n+1)\right] \\
&\quad = \sum_y r_{yx}\mathbb{E}\left[\eta_y(n)\exp\Big(-\big\langle g, \widehat{\varkappa}^{x,(n)}(y,\eta_y(n)) + \sum_{z\neq y}\varkappa^{(n)}(z,\eta_z(n))\big\rangle\Big)\right],
\end{aligned}$$

which completes the proof. □

The following alternative representation of $\hat{\eta}^{(x)}(N)$ is the main result of this subsection:

Proposition 4 *Fix $N \in \mathbb{N}$, $x \in S$. Assume that $\mathbb{E}\,\eta_x(N) = \sum_y \mathbb{E}\,\eta_y(0) r_{yx}^{(N)} < \infty$. Choose $Y(0)$ in S with $\mathbb{P}(Y(0) = y) = \mathbb{E}\,\eta_y(0) r_{yx}^{(N)}/\mathbb{E}\,\eta_x(N)$. Given $Y(0) = y$ let $(Y(0), Y(1), \ldots, Y(N) = x)$ be distributed like an r-bridge from y to x and choose $\widetilde{\eta}(0)$ with distribution $\mathcal{L}(\hat{\eta}^{(y)}(0))$. Pick one particle uniformly from those at y in $\widetilde{\eta}(0)$, this is the (ancestor of the) "selected particle". The system $\widetilde{\eta}(0)$, $i = 1, 2, \ldots, N$ evolves as follows: The selected particle is immortal, its spatial path is $Y(0), Y(1), \ldots, Y(N)$. Particles* not *at the site of the selected one branch and move with the same dynamics as in the original η-system. If there are k particles (including the selected one) at site $Y(i)$ in generation i, a branching event occurs – as in the unmodified system – with probability $s(k)$. In the case of a branching event at a site with currently k particles including the selected one we add a random number of offspring with distribution $\frac{1}{k}\mathcal{L}(\hat{K}-1) + \frac{k-1}{k}\mathcal{L}(K-1)$. Then*

$$\widetilde{\eta}(N) \text{ and } \hat{\eta}^{(x)}(N) \text{ have the same distribution.}$$

Proof. We use induction on N: For $N = 0$, and given $x \in S$, then distribution of $\widetilde{\eta}(0)$ is by definition the the same as that of $\hat{\eta}^{(x)}(0)$. If the representation is correct for $N-1$, we see from Lemma 8 that it also holds for N. □

2.5.2 Approximating the continuous-time system

The discrete-time results of the previous subsection suggest the following approach to the Palm measures of $\xi(T)$: Fix $T > 0$, $x \in S$. Let $\xi(0)$ satisfy $\sum_y \mathbb{E}\,\xi_y(0)p_{yx}(T) < \infty$. Pick $X(0)$ with distribution $\mathbb{P}(X(0) = y) = \mathbb{E}\,\xi_y(0)p_{yx}(T)/\mathbb{E}\,\xi_x(T)$. Given $X(0) = y$ let $(X(t))_{0\leq t\leq T}$ be a p-bridge from y to x and let $\widetilde{\xi}^{x,T}(0)$ have the law of $\hat{\xi}^{(x)}(0)$. Given the path $(X(t))$ the system $(\widetilde{\xi}^{x,T}(t))_{0\leq t\leq T}$ evolves according to the dynamics of a branching system from section 2.1 except that one of the particles at $X(0)$ at time 0 becomes the "selected particle". This particle follows the path X. Whenever a branching event refers to the selected particle (which happens with probability $1/\widetilde{\xi}^{x,T}_{X(t)}(t)$ if a branching event occurs at site $X(t)$ at time t), the system uses the size-biased offspring law $\hat{\nu}$ instead of the usual offspring law ν. (We remark that we could also view $\widetilde{\xi}^{x,T}$ as a time-inhomogeneous Markov process on configurations together with one marked particle. The time-inhomogeneity comes from the condition that the marker has to be at position x at time T.) For fixed x and T, $\widetilde{\xi}^{x,T}$ is a stochastic representation of the locally size-biased law of $\xi(T)$:

Proposition 5 *Assume that $\mathbb{E}\,||\xi(0)||_{\gamma,1} < \infty$. $\widetilde{\xi}^{x,T}(T)$ is distributed according to the Palm measure of $\xi(T)$ with respect to x, that is*

$$\mathbb{E}\Big[F(\widetilde{\xi}^{x,T}(T))\Big] = \frac{1}{\mathbb{E}\,\xi_x(T)}\mathbb{E}\Big[\xi_x(T)F(\xi(T))\Big] \tag{2.39}$$

for all bounded measurable functions F.

We only give a sketch of the proof. Let us first consider the case that the branching rate function $\sigma(\cdot)$ is bounded. Then we can approximate the system ξ by a sequence of discrete-time branching systems η^N as defined in section 2.5.1 in the following way: For $N \in \mathbb{N}$ ($N \geq \sup_k \sigma(k)$) let η^N be a discrete-time branching system as considered in the previous subsection with offspring law ν, branching probability function $s^N(k) = \sigma(k)/N$ and individual motion transition matrix $r^N_{xy} = p_{xy}(1/N)$ starting from $\eta^N(0) = \xi(0)$. Then we have for all $t \geq 0$

$$\eta^N([Nt]) \to \xi(t) \quad \text{as } N \to \infty,$$

at least in the sense of finite-dimensional distributions. Furthermore we have for fixed $T \geq 0$, $x \in S$ (here, we always think of $\widetilde{\eta}^N$ as being constructed with respect to this x, but we do not give it another index)

$$\widetilde{\eta}^N([NT]) \to \widetilde{\xi}^{x,T}(T) \quad \text{as } N \to \infty.$$

Since $\mathcal{L}(\widetilde{\eta}^N([NT]) = \mathcal{L}(\widehat{\eta^N}^{(x)}([NT]))$ for all N, we obtain (2.39) by taking $N \to \infty$.

In order to treat the case of an unbounded σ we approximate ξ by a sequence of continuous-time state dependent branching systems ξ^M with the same specifications as ξ and $\xi^M(0) = \xi(0)$, except that ξ^M uses the modified branching rate function $\sigma^M(k) := \sigma(k \wedge M)$. We conclude from the above that (2.39) holds true for all ξ^M and the respective $(\widetilde{\xi^M})^{x,T}$. Furthermore we can couple ξ and ξ^M in such a way that motion and branching of particles is done exactly in parallel in both systems at all sites with at most M particles. As $||\xi(t)||_{\gamma,1}$ has bounded paths we see that for any fixed finite $A \subset S$, the probability that ξ and ξ^M agree on $A \times [0, T]$ tends to one as $M \to \infty$. Thus (2.39) holds in general.

Remark 7 $\widetilde{\xi}^{x,T}$ can be continued beyond time T in the obvious manner by letting all particles including the "selected one", which gives up its special status at time T, simply follow the ξ-dynamics. This allows to study also $\mathbb{E}[\xi_x(T)F(\xi(T+t))]$ and might thus be interesting for questions about temporal or space-time correlations of ξ. We do not pursue this issue here.

2.6 The case of a quadratic branching rate function

Here, we use the representation of $\hat{\xi}^{(x)}(T)$ via $\widetilde{\xi}^{x,T}$ from the previous section to give a sufficient criterion for persistence of $\xi(t)$ in the situation of a quadratic branching rate function $\sigma(k) = C_\sigma k^2$ and shift-invariant $p_{xy} = p_{0,y-x}$. Technically, the link between ξ and its locally size-biased law is provided by the following well-known, elementary lemma.

Lemma 9 *Let $(X_t)_{t\in\mathbb{T}}$, where $\mathbb{T} \subset \mathbb{R}_+$, be a family of non-negative random variables with finite constant mean $\theta > 0$. Let $(\widehat{X}_t)_{t\in\mathbb{T}}$ have the respective* size-biased *distributions, that is $\mathbb{E}\big[f(\widehat{X}_t)\big] = \theta^{-1}\mathbb{E}\big[X_t f(X_t)\big]$ for $t \in \mathbb{T}$, $f : \mathbb{R}_+ \to \mathbb{R}_+$. Then*

$$(X_t)_{t\in\mathbb{T}} \text{ is uniformly integrable} \iff (\widehat{X}_t)_{t\in\mathbb{T}} \text{ is tight},$$

$$X_t \to 0 \text{ stochastically as } t \to \infty \iff \widehat{X}_t \to \infty \text{ stochastically as } t \to \infty.$$

We include the proof for convenience: For the first equivalence note that for all $t, c \geq 0$ we have $\mathbb{E}[X_t \mathbf{1}(X_t \geq c)] = \theta\mathbb{P}(\widehat{X}_t \geq c)$. Now let us assume that $X_t \to 0$ stochastically as $t \to \infty$. Fix $K > 0$. Then

$$\mathbb{P}(\widehat{X}_t \leq K) = \frac{1}{\theta}\int_0^K x\mathbb{P}(X_t \in dx) \leq \frac{\varepsilon}{\theta}\mathbb{P}(X_t \leq \varepsilon) + \frac{K}{\theta}\mathbb{P}(X_t \in (\varepsilon, K]),$$

showing that $\limsup_{t\to\infty}\mathbb{P}(\widehat{X}_t \leq K) \leq \varepsilon/\theta$. Let $\varepsilon \searrow 0$ to obtain "$\Rightarrow$" in the second line.

Finally assume that $\widehat{X}_t \to \infty$ stochastically. Fix $\delta > 0$, let $\varepsilon > 0$. We have

$$\mathbb{P}(X_t \geq \delta) = \mathbb{P}(X_t \in [\delta, \theta/\varepsilon]) + \mathbb{P}(X_t > \theta/\varepsilon) \leq \mathbb{P}(X_t \in [\delta, \theta/\varepsilon]) + \varepsilon$$

by the Markov inequality. Observing that $\mathbb{P}(X_t \in [\delta, \theta/\varepsilon]) \leq (\theta/\delta)\mathbb{P}(\widehat{X}_t \in [\delta, \theta/\varepsilon])$ we see that $\limsup_{t\to\infty}\mathbb{P}(X_t \geq \delta) \leq \varepsilon$. Let $\varepsilon \searrow 0$ to obtain "$\Leftarrow$". $\square$

We remark that the criterion given by Proposition 6 is parallel to Proposition 9 in section 4.4.1, which treats the corresponding situation for branching random walk in random environment.

Proposition 6 *Let $\xi_x(0)$, $x \in \mathbb{Z}^d$ be shift-invariant with $\mathbb{E}\,\xi_0(0) = \theta > 0$ and $\mathbb{E}[\xi_0(0)^2] < \infty$, and assume that the difference of two independent q-random walks is transient. Let the branching rate function be given by $\sigma(k) = C_\sigma k^2$ and assume that $C_\sigma \mathrm{Var}(\nu) < \beta_*$, where*

$$\beta_* = \sup\left\{\beta \geq 0 : \mathbb{E}\Big[\exp(\beta\int_0^\infty \mathbf{1}(X'(s) = X(s))\,ds \,\Big|\, X\Big] < \infty \;\; \mathcal{L}(X)\text{-almost surely}\right\}$$

with X and X' independent random walks with rate matrix $\kappa(q-I)$ starting from $X(0) = X'(0) = 0$. Then the family $(\xi_0(t))_{t\geq 0}$ is uniformly integrable. In particular, whenever $\xi(t_n) \to \xi(\infty)$ in finite dimensional distributions for some sequence $t_n \to \infty$, then $\xi(\infty)$ has full intensity θ.

Remark 8 Note that by Theorem 7 and Jensen's inequality, we have $\beta_* > 2/\hat{G}_0(\infty)$ whenever p satiesfies (5.18) and $\hat{p}$ is transient, see section 5.5. Combining Remark 6 and Proposition 6 we see that in the case of transient $\hat{p}$ and a quadratic branching rate function $\sigma(k) = C_\sigma k^2$, there is a regime in which all equilibria have infinite second moments.

Proof of Proposition 6. We start with a general observation about branching particle systems. Consider a system ζ consisting of particles that move independently with rate κp_{xy} and that at each site x branch at rate $\tilde{\sigma}(\zeta_x(t-)$ into a random number of offspring with mean $m_x(t)$, where $m_x(t)$ can depend on x, t and $\zeta_x(t-)$. Then we have

$$\frac{\partial}{\partial t}\mathbb{E}\,\zeta_x(t) \quad = \quad \kappa \sum_y p_{yx}\big(\mathbb{E}\,\zeta_y(t) - \mathbb{E}\,\zeta_x(t)\big) + \mathbb{E}[\tilde{\sigma}(\zeta_x(t))(m_x(t)-1)].$$

Now fix $T > 0$ and consider the population in the construction of $\widetilde{\xi}^{0,T}$ with the exception of the selected particle, i.e. define

$$\zeta(t) := \widetilde{\xi}^{0,T}(t) - \delta_{X(t)}, \quad 0 \le t \le T.$$

where $(X(t))_{0\le t\le T}$ is the path of the selected particle in the construction from section 2.5.2. It will be more convenient to work with the time reversal Y defined by $Y(t) := X(T-t)$, $0 \le t \le T$, which is by construction and the shift-invariance of $\mathcal{L}(\xi(0))$ a (unrestricted) q-random walk starting from 0.

Conditional on the path Y, ζ is a branching system with space-time dependent branching rates and offspring laws as considered above. By construction we have $\tilde{\sigma}(\zeta_x(t)) = C_\sigma(\zeta_x(t) + \mathbf{1}(Y(T-t)=x))^2$ and

$$m_x(t) = \begin{cases} \dfrac{1}{1+\zeta_x(t)}\sum k\hat{\nu}_k + \dfrac{\zeta_x(t)}{1+\zeta_x(t)} & \\ \quad = \mathrm{Var}(\nu)/(1+\zeta_x(t)) + 1 & \text{if } Y(T-t) = x \\ \qquad 1 & \text{otherwise.} \end{cases}$$

This allows to compute

$$\begin{aligned} \frac{\partial}{\partial t}\mathbb{E}[\zeta_x(t)|Y] \quad &= \quad \kappa \sum_y p_{yx}\big(\mathbb{E}[\zeta_y(t)|Y] - \mathbb{E}[\zeta_x(t)|Y]\big) \\ &\qquad + C_\sigma \mathrm{Var}(\nu)\mathbf{1}(Y(T-t)=x)\mathbb{E}[\zeta_x(t)+1|Y]. \end{aligned}$$

The Feynman-Kac formula (see e.g. [18], chapter 2.17, Thm. 3) shows that

$$\begin{aligned} \mathbb{E}[\zeta_x(T)|Y] \quad &= \quad \mathbb{E}\Big[\exp\Big(C_\sigma \mathrm{Var}(\nu)\int_0^T \mathbf{1}(X'(t)=Y(t))\,dt\Big)\zeta_{X'(T)}(0)\,\Big|\,Y\Big] \\ &\qquad + \int_0^T C_\sigma \mathrm{Var}(\nu)\mathbb{E}\Big[\mathbf{1}(X'(t)=Y(t) \\ &\qquad\qquad \times \exp\Big(C_\sigma \mathrm{Var}(\nu)\int_0^t \mathbf{1}(X'(s)=Y(s))\,ds\Big)\Big|\,Y\Big]\,dt, \end{aligned}$$

where X' is a q-random walk starting from x, independent of Y. Note that this formula has a natural interpretation in genealogical terms: the first summand on the righthand side gives the expected number of particles in $\widetilde{\xi}^{0,T}(T)$ at site 0 which are not related to the selected particle, while the second term is the expected number of relatives. Our assumption $\mathbb{E}[\xi_0(0)^2] < \infty$ ensures that $\sup_x \mathbb{E}\zeta_x(0) < \infty$, yielding

$$\begin{aligned}
\mathbb{E}[\zeta_x(T)|Y] \quad \leq \quad & c \times \mathbb{E}\left[\exp\left(C_\sigma \mathrm{Var}(\nu)\int_0^\infty \mathbf{1}(X'(t) = Y(t))\,dt\right)\Big|\, Y\right] \\
& + \int_0^\infty C_\sigma \mathrm{Var}(\nu)\mathbb{E}\Big[\mathbf{1}(X'(t) = Y(t) \\
& \qquad \times \exp\left(C_\sigma \mathrm{Var}(\nu)\int_0^t \mathbf{1}(X'(s) = Y(s))\,ds\right)\Big|\, Y\Big]\,dt.
\end{aligned}$$

The righthand side is finite for $C_\sigma < \beta_*/\mathrm{Var}(\nu)$, showing that the family $\widetilde{\xi}^{0,T}(T)$, $T \geq 0$, and thus also the family $\hat{\xi}^{(0)}(T)$ is tight. The claim follows from Lemma 9. □

2.7 If branching is bad for you more branching is worse

In this section we prove a comparison result for state dependent critical binary branching random walks. Consider two such systems $(\xi^{(1)}(t))_{t\geq 0}$ and $(\xi^{(2)}(t))_{t\geq 0}$ starting from the same initial condition with the same specifications except for the branching rate functions σ_i corresponding to $\xi^{(i)}$, and assume that $\sigma_1(\cdot) \geq \sigma_2(\cdot)$. Then we intuitively expect the first system to "branch more frequently" than the second, thus there should be more variability in it and it should "more easily" reach the boundary (namely the absorbing state $\underline{0}$) than the second system. This idea, which technically means that $\mathbb{E}\, f(\xi^{(1)}(t)) \geq \mathbb{E}\, f(\xi^{(2)}(t))$ for certain convex functions, is made precise in Theorem 1. It may be viewed as a "particle companion" of the main result in [4], where a corresponding theorem for interacting diffusions is proved.

Denote by $\mathbb{F}$ the set of all bounded Lipschitz continuous functions $F : E_1 \to \mathbb{R}$ such that

$$\forall\, \xi \in E_1, x, y \in S \quad F(\xi + \delta_x + \delta_y) - F(\xi + \delta_x) - F(\xi + \delta_y) + F(\xi) \geq 0. \tag{2.40}$$

This is a discrete analogue of the requirement that all second derivatives should be positive: If $F(\xi) = f(\xi_{y_1}, \ldots, \xi_{y_n})$ for some $y_1, \ldots, y_n \in S$ and $f : \mathbb{R}^n \to \mathbb{R}_+$ with $\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f \geq 0$, $i, j = 1, \ldots, n$, then $F \in \mathbb{F}$. We remark that $F \in \mathbb{F}$ is equivalent to the requirement that

$$F(\xi + h_1\delta_x + h_2\delta_y) - F(\xi + h_1\delta_x) - F(\xi + h_2\delta_y) + F(\xi) \geq 0$$

for all ξ, x, y and all $h_1, h_2 \in \mathbb{N}_0$, as can be checked by induction. An important example is of course $F(\xi) := \exp(-\lambda_1\xi_{y_1} - \cdots - \lambda_n\xi_{y_n})$ for any $\lambda_1, \ldots, \lambda_n \geq 0$ and $y_1, \ldots, y_n \in S$.

Let $L^{(\sigma_i)}$ denote the generator, $(S^{(\sigma_i)}(t))_{t\geq 0}$ the semigroup corresponding to $\xi^{(i)}$, $i = 1, 2$.

Theorem 1 *Assume (2.4), that is the branching law is critical binary. Then for all $F \in \mathbb{F}$, $\xi \in E_1, t \geq 0$: $S^{(\sigma_1)}(t)F(\xi) \geq S^{(\sigma_2)}(t)F(\xi)$.*

Before giving a proof let us explain the title of this section by the following

Corollary 1 *Assume (2.4), and let $\xi^{(1)}$ and $\xi^{(2)}$ be as above. If $\xi^{(2)}$ suffers (weak) local extinction, i.e. if*

$$(\xi^{(2)}_{x_1}(t), \ldots, \xi^{(2)}_{x_n}(t)) \to (0, \ldots, 0), \quad t \to \infty$$

in distribution for all $x_1, \ldots x_n \in S$, then so does $\xi^{(1)}$.

Proof. Let $F(\xi) := \exp(-\xi_{x_1} - \cdots - \xi_{x_n})$ (≤ 1). By Theorem 1 we have $S^{(\sigma_1)}(t)F(\xi) \geq S^{(\sigma_2)}(t)F(\xi)$ and the rhs converges to 1 as $t \to \infty$ because $\xi^{(2)}$ suffers local extinction. □

Remark 9 We work in this chapter only under assumption (2.4) of critical binary branching. We do not consider more general offspring distributions because the analogue of the crucial Lemma 10 will be false in the general case. The present proof, which relies on the preservation of convexity properties by the semigroup, can therefore not be adapted to the general setting. On the other hand, the intuitive arguments given at the beginning of this section apply whenever the expected number of children equals one. Thus one might suspect that Corollary 1 should be true for any critical offspring distribution. Alas, it is unclear how to prove this.

Modulo technicalities – which we leave aside for a second – the proof of Theorem 1 rests on two pillars:

Lemma 10 *Assume (2.4). The semigroup $(S^{(\sigma_i)}(t))$ preserves $\mathbb{F}$, i.e. $S^{(\sigma_i)}(t)F \in \mathbb{F}$ for all $F \in \mathbb{F}$, $i = 1, 2$.*

and an application of the integration by parts formula for semigroups (cf. e.g. [23], p. 367)

$$S^{(\sigma_1)}(t) - S^{(\sigma_2)}(t) = \int_0^t S^{(\sigma_2)}(t-s)(L^{(\sigma_1)} - L^{(\sigma_2)})S^{(\sigma_1)}(s)\, ds. \tag{2.41}$$

Because the motion parts of $L^{(\sigma_1)}$ and $L^{(\sigma_2)}$ are the same we have

$$(L^{(\sigma_1)} - L^{(\sigma_2)})F(\xi) = \sum_{x \in S} \big(\sigma_1(\xi_x) - \sigma_2(\xi_x)\big) \sum_{k \geq 0} \nu_k \left(F(\xi + (k-1)\delta_x) - F(\xi)\right).$$

This is (summand-wise) non-negative for $F \in \mathbb{F}$ since $\sigma_1(\xi_x) \geq \sigma_2(\xi_x)$ by assumption and by Jensen's inequality as the function $\mathbb{N}_0 \ni n \mapsto F(\xi + n\delta_x) \in \mathbb{R}$ is convex for any $F \in \mathbb{F}$, $\xi \in E_1$, $x \in S$.

Finally, there remain some technicalities to be settled before the theorem follows from (2.41) and Lemma 10. Unfortunately, (2.41) does in general not hold in our situation, $\big(S^{(\sigma_i)}(t)L^{(\sigma_i)}f\big)(\xi)$ need not even be defined for all $\xi \in E_1$ when $\sigma^{(i)}$ grows faster than linearly. This can be remedied by an approximation argument: We indeed have

$$S^{(\sigma_1)}(t)f(\xi) - S^{(\sigma_2)}(t)f(\xi) = \int_0^t S^{(\sigma_2)}(t-s)(L^{(\sigma_1)} - L^{(\sigma_2)})S^{(\sigma_1)}(s)f(\xi)\, ds$$

for $f \in \mathrm{Lip}(E_1)$ and any *finite* ξ, so the theorem is true for all finite ξ. It remains to observe that by (2.21) we have $S^{(\sigma_i)}(t)f(\xi) = \lim_n S^{(\sigma_i)}(t)f(\mathbf{1}_{S_n}\xi)$ for all $f \in \mathrm{Lip}(E_1)$ and $\xi \in E_1$, where $S_n \nearrow S$ is an exhausting sequence of finite subsets. □

Proof of Lemma 10. The idea is to first show that both the motion dynamics and the branching dynamics preserve $\mathbb{F}$ and then glue the two parts together via Trotter's product formula. For technical reasons we first consider E_{fin} as the underlying configuration space because L given by (2.1) does generate a strongly continuous contraction semigroup $S(t)$ on $C_b(E_{fin}, \mathbb{R})$. Observe that $Lf(\xi) = L_{rw}f(\xi) + L_{br}f(\xi)$, where

$$\begin{aligned} L_{rw}f(\xi) &= \kappa \sum_{x,y \in S} \xi_x p_{xy} \Big(f(\xi^{(x,y)}) - f(\xi)\Big), \text{ and} \\ L_{br}f(\xi) &= \sum_{x \in S} \sigma(\xi_x) \sum_{k \geq 0} \nu_k \big(F(\xi + (k-1)\delta_x) - F(\xi)\big) \end{aligned}$$

are the motion and branching part respectively. We denote the corresponding semigroups by $S_{rw}(t)$ and $S_{br}(t)$.

Claim 1 $S_{rw}(t)$ preserves $\mathbb{F}$.

We observe that the Markov process $(\xi^{rw}(t))$ with semigroup $(S_{rw}(t))$ consists of independent p-chains jumping at rate κ. This allows a simple coupling argument: Fix $\xi \in E_1$, $x, y \in S$. Let $(\xi^{rw}(t))$ be such a system with initial condition ξ and (X_t) and (Y_t) be two independent p-Markov chains with $X_0 = x$, $Y_0 = y$. Thus

$$\begin{aligned} &S_{rw}(t)F(\xi + \delta_x + \delta_y) - S_{rw}(t)F(\xi + \delta_x) - S_{rw}(t)F(\xi + \delta_y) + S_{rw}(t)F(\xi) \\ &\quad = \mathbb{E}\left[F(\xi^{rw}(t) + \delta_{X_t} + \delta_{Y_t}) - F(\xi^{rw}(t) + \delta_{X_t}) - F(\xi^{rw}(t) + \delta_{Y_t}) + F(\xi^{rw}(t))\right] \geq 0 \end{aligned}$$

for $F \in \mathbb{F}$, proving Claim 1.

Claim 2 $S_{br}(t)$ preserves $\mathbb{F}$.

The Markov process $(\xi^{br}(t))$ corresponding to $(S_{br}(t))$ consists of $|S|$ independent copies of a continuous-time Markov process $(Y(t))$ on $\mathbb{Z}_+$ with Q-matrix $r(m,n) = \sigma(m)\mathbf{1}(|m-n| = 1)/2$ for $m \neq n$. Let $F \in \mathbb{F}$. Fix $x, y \in S$ and an initial state $\eta \in E_1$. We wish to prove that

$$\begin{aligned} &\mathbb{E}[F(\xi^{br}(t))|\xi^{br}(0) = \eta + \delta_x + \delta_y] - \mathbb{E}[F(\xi^{br}(t))|\xi^{br}(0) = \eta + \delta_x] \\ &\qquad - \mathbb{E}[F(\xi^{br}(t))|\xi^{br}(0) = \eta + \delta_y] + \mathbb{E}[F(\xi^{br}(t))|\xi^{br}(0) = \eta] \quad \geq \quad 0. \end{aligned} \tag{2.42}$$

Let us denote by $X(t)$ the number of particles at x at time t, by $Y(t)$ the corresponding number at y, and let $\xi'(t) = (\xi'_z(t))_{z \neq x,y}$ be the vector of the remaining coordinates. Conditioning on $\xi'(t)$ and observing that the coordinate processes are independent, we can rewrite (2.42) as follows:

$$\begin{aligned} \mathbb{E}\Bigg[&\mathbb{E}[F(\xi^{br}(t))|X(0) = \eta_x + 1, Y(0) = \eta_y + 1, \xi'(t)] \\ &- \mathbb{E}[F(\xi^{br}(t))|X(0) = \eta_x + 1, Y(0) = \eta_y, \xi'(t)] \\ &- \mathbb{E}[F(\xi^{br}(t))|X(0) = \eta_x, Y(0) = \eta_y + 1, \xi'(t)] \\ &\quad + \mathbb{E}[F(\xi^{br}(t))|X(0) = \eta_x, Y(0) = \eta_y, \xi'(t)] \Bigg| \xi'(0) = (\eta_z)_{z \neq x,y}\Bigg]. \end{aligned}$$

Now for fixed $\xi'(t)$, the terms inside the big expectation are functions of only two coordinates, namely η_x and η_y. Consequently it suffices to consider $F \in \mathbb{F}$ of the form $F(\xi) = g(\xi_x, \xi_y)$ for some suitable $g : \mathbb{N}_0^2 \to \mathbb{R}$.

Case $x = y$, i.e. $F(\xi) = g(\xi(x))$ for some (convex) $g : \mathbb{N}_0 \to \mathbb{R}$. Define

$$\varphi(n,t) := \mathbb{E}_n g(Y(t)), \quad n \in \mathbb{N}_0, \ t \geq 0.$$

φ solves

$$\begin{aligned} \frac{\partial}{\partial t}\varphi(n,t) &= \frac{1}{2}\sigma(n)(\varphi(n+1,t) + \varphi(n-1,t) - 2\varphi(n,t)), \quad t \geq 0, \ n \in \mathbb{N}_0 \\ \varphi(n,0) &= g(n) \end{aligned}$$

and because $(Y(t))$ is a martingale and g convex, $g(Y(t))$ is a submartingale obeying $\frac{\partial}{\partial t}\mathbb{E}_n g(Y(t)) \geq 0$. This proves

$$\mathbb{E}_{\xi+2\delta_x}F(\xi^{br}(t)) + \mathbb{E}_{\xi}F(\xi^{br}(t)) - 2\mathbb{E}_{\xi+\delta_x}F(\xi^{br}(t)) = \frac{2}{\sigma(\xi_x+1)}\frac{\partial}{\partial t}\varphi(\xi_x+1,t) \geq 0 \tag{2.43}$$

as desired if $\sigma(\xi_x+1) > 0$. But on the other hand if $\sigma(\xi_x+1) = 0$ we have $\mathbb{E}_{\xi+\delta_x} F(\xi_x^{br}(t)) \equiv F(\xi_x+\delta_x)$ so that the lhs of (2.43) is still non-negative for all $t \geq 0$ by the submartingale property.

Case $x \neq y$, i.e. $F(\xi) = g(\xi_x, \xi_y)$ for some g satisfying $g(m+h, n+k) - g(m+h, n) - g(m, n+k) + g(m,n) \geq 0$ for all $m, n, h, k \in \mathbb{N}_0$. Here we use an obvious coupling: Let $\xi_x = n, \xi_y = m$ and $(Y^1(t))$ and $(\widetilde{Y}^1(t))$ be two continuous time Markov chains with rate matrix $r(a,b) = \sigma(a)\mathbf{1}(|a-b| = 1)/2$ as above with $Y^1(0) = m, \widetilde{Y}^1(0) = m+1$ coupled such that $Y_t^1 \leq \widetilde{Y}_t^1$ for all t, and similarly $(Y^2(t))$ and $(\widetilde{Y}^2(t))$ be two such chains as above (but independent of $(Y^1, \widetilde{Y}^1)$) with $Y^2(0) = n, \widetilde{Y}^2(0) = n+1$ and $Y^2(t) \leq \widetilde{Y}^2(t)$ for all t. Observe that our assumption of attractiveness guarantees that we can choose such order-preserving couplings. This allows to estimate

$$\begin{aligned}
&\mathbb{E}_{\xi+\delta_x+\delta_y} F(\xi^{br}(t)) - \mathbb{E}_{\xi+\delta_x} F(\xi^{br}(t)) - \mathbb{E}_{\xi+\delta_y} F(\xi^{br}(t)) + \mathbb{E}_{\xi} F(\xi^{br}(t)) \\
&\quad = \mathbb{E}\big[g(\widetilde{Y}^1(t), \widetilde{Y}^2(t)) - g(\widetilde{Y}^1(t), Y^2(t)) - g(Y^1(t), \widetilde{Y}^2(t)) + g(Y^1(t), Y^2(t))\big] \geq 0
\end{aligned}$$

by the properties of g.

To complete the proof of Lemma 10 we split the interval $[0,t]$ into n pieces in which we first run the ξ^{rw}-dynamics for time t/n and then the ξ^{br}-dynamics for time t/n. This corresponds to the operator

$$F \mapsto (S_{br}(t/n) S_{rw}(t/n))^n \, F$$

which by Claims 1 and 2 preserves $\mathbb{F}$. But by Trotter's product formula (see e.g. [9], Cor. 1.6.7) this converges to $S_t F$ (as an element of $C(E_{fin}, \mathbb{R})$). Then approximate $\xi \in E_1$ by $\xi \mathbf{1}_{S_n}$ to obtain the general statement. □

2.8 With recurrent motion, critical branching is presumably always bad for you

In view of the comparison result Corollary 1 it is natural to ask about the long-time behaviour of the so-called lonely branching system using the branching rate function $\sigma(k) = \mathbf{1}(k = 0)$. By Lemma 9 and Proposition 5, the lonely branchers will become locally extinct if and only if the corresponding $\widetilde{\xi}^{0,T}(T)$ become infinitely dense. Unfortunately, we have not been able to decide whether this is the case. This has lead us to consider a caricature, where the "trunk" does not move and "side-lines" do not branch. This model, which is a (special case of a) system of independent Markov chains with a density-regulated immigration mechanism, is studied in the next chapter. Here we have been able to decide the longtime behaviour, see Theorem 2. It turns out that (as in the case of the "classical Kallenberg tree") the system becomes infinitely dense if and only if the underlying motion is recurrent.

These results and our faith that the truth is captured by the caricature step lead us to

Conjecture 1 *Consider a state dependent branching system $\xi(t)$ with critical binary branching, a non-trivial branching rate function $\sigma(\cdot)$ and underlying motion given by a symmetric recurrent random walk, starting from a shift invariant initial condition (with finite intensity $\mathbb{E}\,\xi_x(0) < \infty$). Then $\xi(t)$ dies out locally as $t \to \infty$. In particular, there are no non-trivial equilibria.*

Chapter 3

Systems with self-blocking immigration

We consider systems of independent Markov chains on a countable space S. New particles try to immigrate at $x_0 \in S$ at rate 1 but are only allowed to do so if the source x_0 is not occupied by another (older) particle. We show that – starting from the empty configuration – such a system grows locally without bound iff the motion of individual particles is recurrent. We present some heuristic arguments for the growth rate in certain special cases, namely positive recurrent motion and simple random walk in $d = 1$, and give a proof for a resulting conjecture under rather restrictive assumptions, namely if the individual motion is uniformly exponentially ergodic. We find some partial results about the quantitative long-time behaviour in the random walk case.

3.1 Ingredients and construction

S .... "basic space", some countable set
$\eta_x(t)$ .... no. of particles in $x \in S$ at time t
(p_{xy}) .... irreducible stochastic matrix on S,
particles follow independent p-motion with rate 1.
$x_0 \in S$ .... "immigration source", at rate $\mathbf{1}(\eta_{x_0}(t-) = 0)$
a new particle enters at x_0.

We construct an interacting particle system $(\eta(t))$. Particles move around independently in a Markovian manner, at rate 1 new particles enter the system at "source" x_0 unless there is already one (or more) particle(s) at x_0 at this time: In that case no immigration occurs. This is the "self-blocking" the name refers to.

We denote the corresponding transition semigroup of the continuous-time chain with jump rates p by $p_{xy}(t)$. For (fixed) initial condition $\eta(0)$ in

$$\mathbb{S} := \{\xi : S \to \mathbb{N}_0 \,|\, \forall t \geq 0, x \in S : \textstyle\sum_{y \in S} \xi_y p_{yx}(t) < \infty\} \subset {\mathbb{N}_0}^S \tag{3.1}$$

Note that $E_1 \subset \mathbb{S}$ if p is compatible with the weight function γ from the previous chapter in the sense that (2.6) is satisfied.

We can construct the process η in the straightforward manner:
Let $N = (\tau_1, \tau_2, \ldots)$ $(\tau_1 < \tau_2 < \cdots)$ be a homogeneous rate 1 Poisson point process on $\mathbb{R}_+$ and given N let $(X^{(i)}(t))_{t \geq \tau_i}$ be independent p-chains, $X^{(i)}(\tau_i) = x_0$. Furthermore

let $(X^{(x,i)}(t))_{t\geq 0}$, $x \in S$, $i = 1,2,\dots$ be independent p-chains with $X^{(x,i)}(0) = x$, independent of N and $\{X^{(i)}\}$. Observe that for $\eta(0) \in \mathbb{S}$, $x \in S$, $t \geq 0$

$$\mathbb{E}\left[\sum_y \sum_{i=1}^{\eta_y(0)} \mathbf{1}(X^{(y,i)}(t) = x)\right] = \sum_y \eta_y(0) p_{yx}(t) < \infty \tag{3.2}$$

saving the system from "instantaneous explosions". We can therefore safely define

$$\eta(t) := \sum_x \sum_{i=1}^{\eta_x(0)} \delta_{X^{(x,i)}(t)} + \sum_n \mathbf{1}(\tau_n \leq t)\mathbf{1}(A_n)\delta_{X^{(n)}(t)}$$

where the A_n are recursively defined as $A_n := \{\eta_{x_0}(\tau_n-) = 0\}$.

Because in finite time only a finite number of immigrants enters the system a.s. (the number of immigrants up to time t is stochastically bounded by Poisson(t)) we have for all $\eta(0) \in \mathbb{S}$

$$\eta(t) \in \mathbb{S} \quad \text{for all } t \geq 0 \text{ almost surely.}$$

By the independence of the ingredients and standard properties of Poisson processes on $\mathbb{R}$, $(\eta(t))$ is a Markov process. We denote its semigroup (acting, say, on bounded measurable functions f on $\mathbb{S}$) by

$$T(t)f(\xi) := \mathbb{E}_\xi\big[f(\eta(t))\big], \quad \xi \in \mathbb{S}.$$

Remark 10 a) The set $\mathbb{S}$ is the largest possible state space for such a system: An easy argument using Laplace transforms shows that if for some t, x we have $\sum_y \eta_y(0)p_{yx}(t) = \infty$ then the system explodes by time t in the sense that $\eta_x(t) = \infty$ almost surely.

b) A simple coupling shows that $T(t)$ preserves monotonicity, that is $T(t)f$ is monotone whenever $f : \mathbb{S} \to \mathbb{R}$ is: Let $\eta.(0) \leq \tilde{\eta}.(0)$. We let the particles in η move in unison with "partners" in $\tilde{\eta}$, possible overshoot particles in the larger system simply move independently. We use a common Poisson process of potential immigrations: Immigration events at times t when $\eta_0(t-) = \tilde{\eta}_0(t-) = 0$ are realised in both systems; while if $0 = \eta_0(t-) < \tilde{\eta}_0(t-)$, the immigration occurs only in the smaller system. Note that the ordering is preserved in any case.

3.2 The backward picture

Here we give an alternative construction of the historical process of $(\eta(u))_{0\leq u\leq v}$ for a fixed v. This construction works "backwards" in time. It shows easily that $\mathcal{L}(\eta(u))$ is increasing in u.

Let $\mathbb{M} := \{(f,t) \,:\, t \leq 0, f : [t,0] \to S \text{ càdlàg}, f(t) = x_0\}$ be the set of càdlàg paths on S that start at some time $t \leq 0$ at x_0 and are marked by their "birth time". We construct a family $(\Psi_t)_{t\leq 0}$ of random simple counting measures on $\mathbb{M}$ (or equivalently, of finite random subsets of $\mathbb{M}$), such that for $s < t \leq 0$ "the restriction of Ψ_s to 'what happens to paths in $[t,0]$' is equal to Ψ_t".

Instead of (or at least in addition to) reading the rest of this rather notation-loaded section the reader might want to look at the cartoon in section 3.7.

Let $N = \sum_i \delta_{\tau_i}$, $0 \geq \tau_1 > \tau_2 > \dots$ be a rate 1 Poisson process on $\mathbb{R}_-$, given N let $X^{(1)}, X^{(2)}, \dots$ be independent continuous-time p-chains, $(X^{(i)}(t))_{t\geq\tau_i}$ starts at time τ_i at x_0. With these ingredients we define iteratively $\Psi_t := \emptyset$ for $0 \geq t > \tau_1$, $\Psi_{\tau_1} := \delta_{(X^{(1)},\tau_1)}$.

If Ψ_t has been constructed for $t \geq \tau_{n-1}$ we define $\Psi_t := \Psi_{\tau_{n-1}}$ for $t \in (\tau_n, \tau_{n-1})$. Now if $\Psi_{\tau_{n-1}}(\{\text{all paths}\} \times \{r : X^{(n)}(r) = x_0\}) = 0$, that is if $X^{(n)}$ does not hit any "birth point" of a path currently "in Ψ", we define $\Psi_{\tau_n} := \Psi_{\tau_{n-1}} + \delta_{(X^{(n)}, \tau_n)}$. Otherwise there is some $R_n \in [\tau_{n-1}, 0]$ and $\widetilde{Y}^{(n)}(\cdot)$ such that $\Psi_{\tau_{n-1}}(\{(\widetilde{Y}^{(n)}, R_n)\}) = 1$ and $X^{(n)}(R_n) = x_0$ (and R_n is the smallest r with these properties). By construction, R_n will be some τ_{k_n} with $k_n < n$. In this case we define

$$Y^{(n)}(r) := \begin{cases} X^{(n)}(r), & r \in [\tau_n, R_n) \\ \widetilde{Y}^{(n)}(r), & r \in [R_n, 0] \end{cases}$$

and set $\Psi_{\tau_n} := \Psi_{\tau_{n-1}} - \delta_{(\widetilde{Y}^{(n)}, R_n)} + \delta_{(Y^{(n)}, \tau_n)}$. Observe that $Y^{(n)}$ is distributed like a p-chain starting at time τ_n at x_0.

For a fixed $t < 0$ we can also look at $\Psi_t = \sum \delta_{(Y^{(i)}, S_i)}$ in "reversed" (with respect to the above construction), i.e. forward direction, time:
Let $\tau^1_{N_t}$ be the smallest jump point of N in $[t, 0]$ then by construction Ψ_t has an atom $\delta_{(Y^{(1)}, \tau^1_{N_t})}$ where $Y^{(1)}$ is a continuous-time p-chain starting at x_0 at time $\tau^1_{N_t}$ and Ψ_t cannot have an atom of the form $\delta_{(Y,r)}$ for $r \in \{s : Y^{(1)}(s) = x_0\}$. Now let $\tau^2_{N_t}$ be the smallest jump point of N in $[\tau^1_{N_t}, 0] \setminus \{s : Y^{(1)}(s) = x_0\}$. If there is any such $\tau^2_{N_t}$ by construction Ψ_t also has an atom $\delta_{(Y^{(2)}, \tau^2_{N_t})}$ where $Y^{(2)}$ is a p-chain starting at x_0 at time $\tau^2_{N_t}$ independent of $(Y^{(1)}, \tau^1_{N_t})$. Now we look for the smallest jump point of N in $[\tau^2_{N_t}, 0] \setminus \{s : Y^{(1)}(s) = x_0 \text{ or } Y^{(2)}(s) = x_0\}$, etc.

Thus we see that for $t \leq 0$ fixed Ψ_t is the historical empirical measure (at time 0) of a system of independent p-chains with self-blocking immigration that starts with the empty configuration at time t (For convenience we may think of a shifted time-axis and speak rather of $[0, |t|]$ than of $[t, 0]$).

3.3 Qualitative longtime behaviour

In this section we consider the case $\eta(0) \equiv 0$ and ask whether the system will overflow locally, i.e. $\eta_x(t) \to \infty$ in probability?

Lemma 11 *Let $\eta(0) \equiv 0$. The family $\{\mathcal{L}(\eta(t)) : t \geq 0\}$ is stochastically monotone. Furthermore exactly one of the following holds: Either*

i) $\{\mathcal{L}(\eta(t)) : t \geq 0\}$ is tight, or

ii) $\eta_{x_0}(t) \to \infty$ stochastically as $t \to \infty$.

In case i) $\eta(t) \to \eta^{(\infty)}$ (in the sense of finite-dimensional distributions), where $\eta^{(\infty)}$ is an ($\mathbb{S}$-valued) equilibrium for the process.

Proof. To show monotonicity we use the coupling provided by the backward construction of the previous section. Denoting for $t \geq 0$

$$\widetilde{\eta}(t) := \int \delta_{f(0)} \Psi_{-t}(d(f, s))$$

(that is we record only the terminal positions of the particles in the backward construction) we have by construction

$$\widetilde{\eta}.(s) \leq \widetilde{\eta}.(t) \quad \text{for } s \leq t$$

(because once a path has reached the 'top' in the backward construction its tip is never removed) and for all $t \geq 0$

$$\mathcal{L}(\eta(t)) = \mathcal{L}(\widetilde{\eta}(t)).$$

This gives the monotonicity. In particular there exists a family $(\eta_x^{(\infty)})_{x\in S}$ of $(\mathbb{N}_0 \cup \{\infty\})$-valued random variables such that $\eta(t) \to \eta^{(\infty)}$ in f.d.d. (functions of the type $f(\eta) = \mathbf{1}(\eta_{x_1} \geq k_1, \ldots, \eta_{x_n} \geq k_n)$ are monotone).

Observe that the backward construction can be viewed as a deterministic construction using independent random ingredients: $Z_1, Z_2, \ldots$ Exp(1)-rv's for the waiting times between jumps of the Poisson process N and $(X_s^{(1)}), (X_s^{(2)}), \ldots$ paths of p-chains starting from x_0. The event

$$\{\widetilde{\eta}_{x_0}(t) \to \infty\} = \{\text{infinitely many paths arrive at } x_0 \text{ at time } 0\}$$

remains unchanged if we permute a finite number of these ingredients: Any such permutation can only create or destroy a finite number of "birth points" and hence only affects the behaviour of a finite number of paths in the construction. By Hewitt-Savage's 0-1 law we obtain $\mathbb{P}(\widetilde{\eta}_{x_0}(t) \to \infty) \in \{0,1\}$. This gives the dichotomy we claimed above because $\mathbb{P}(\widetilde{\eta}_{x_0}(t) \to \infty) = \mathbb{P}(\eta_{x_0}^{(\infty)} = \infty)$. So if the family $\mathcal{L}(\eta(t)), t \geq 0$ is not tight, i.e. $\mathbb{P}(\eta_{x_0}^{(\infty)} = \infty) > 0$, we see that indeed the limit law puts *all* of its mass on ∞.

Assume now case *i)*. To conclude the proof we want to convince ourselves that $\mathbb{P}(\eta^{(\infty)} \in \mathbb{S}) = 1$. Hence we check that

$$\eta_{x_0}(t), t \geq 0 \text{ is tight} \iff \forall h, x : \Big(\sum\nolimits_y \eta_y(t) p_{yx}(h), t \geq 0 \text{ is tight}\Big).$$

"$\Leftarrow$" is obvious because $\eta_{x_0}(t) \leq (1/p_{x_0,x_0}(t)) \sum_y \eta_y(t) p_{y,x_0}(1)$. For the other direction assume that there exist h' and x' for which the righthand side above is not tight. Because the distribution of $\eta_{x'}(t+h')$ given $\eta(t)$ is stochastically larger than that of $\sum_y B_y$ where the B_y are independent, $\mathrm{Bin}(\eta_y(t), p_{yx'}(h'))$-distributed we can compute for (small) $\lambda > 0$:

$$\begin{aligned}
\mathbb{E}[e^{-\lambda \eta_{x'}(t+s)}] &= \mathbb{E}\big[\mathbb{E}[e^{-\lambda \eta_{x'}(t+s)} | \eta(t)]\big] \\
&\leq \mathbb{E}\Big[\prod\nolimits_y (1 - p_{yx'}(s)(1 - e^{-\lambda}))^{\eta_y(t)}\Big] \\
&= \mathbb{E}\Big[\exp\Big(\sum \eta_y(t) \log(1 - p_{yx'}(s)(1 - e^{-\lambda}))\Big)\Big] \\
&\leq \mathbb{E}\Big[\exp\big(-(1 - e^{-\lambda}) \sum \eta_y(t) p_{yx'}(s)\big)\Big] \\
&\leq \mathbb{E}\Big[\exp\big((-\lambda/2) \sum \eta_y(t) p_{yx'}(s)\big)\Big]
\end{aligned}$$

From this and the assumed non-tightness of the family appearing inside the exp in the last line we get

$$\lim_{\lambda \to 0+} \inf_t E(e^{-\lambda \eta_{x'}(t)}) < 1$$

so that $(\eta_{x'}(t), t \geq 0)$ is also not tight. Because $p_{x',x_0}(1) > 0$ the same holds with x' replaced by x_0. Consequently in case *i)*:

$$1 = \mathbb{P}\big(\forall x \in S, t \in \mathbb{Q}_+ : \textstyle\sum_y \eta_y^{(\infty)} p_{yx}(t) < \infty\big) = \mathbb{P}(\eta^{(\infty)} \in \mathbb{S}).$$

To see that $\mathcal{L}(\eta^{(\infty)}$ is invariant fix $h > 0$ and a bounded monotone function $f : \mathbb{S} \to \mathbb{R}_+$ and compute ($\underline{0} := (0, 0, \ldots) \in \mathbb{S}$ denotes the empty configuration)

$$\mathbb{E}\big[T(h)f(\eta_\infty)\big] = \lim_{t\to\infty} T(t)(T(h)f)(\underline{0}) = \lim_{t\to\infty} T(t+h)f(\underline{0}) = \mathbb{E}\big[f(\eta^{(\infty)})\big].$$

□

Remark 11 It is obvious that $\eta(t)$ converges to some equilibrium if p-motion is transient. A simple coupling with a system $(\eta'(t))$ in which immigration occurs unrestrictedly at constant rate shows that

$$\mathbb{E}\,\eta_{x_0}(t) \le \mathbb{E}\,\eta'_{x_0}(t) = \int_0^t p_{x_0x_0}(t-s)ds \le \int_0^\infty p_{x_0x_0}(u)du < \infty.$$

On the other hand if p is recurrent there exist no equilibria at all. Heuristically the existence of an equilibrium leads to a contradiction because started from it the system would produce new particles at constant rate but a recurrent motion is too slow to move them away so that they pile up higher and higher "around x_0": Assume $(\eta^{(\infty)}(t))_{t\ge0}$ was such a stationary process with $\eta^{(\infty)}(t) \in \mathbb{S}$ for all t almost surely, whence $\beta := \mathbb{P}(\eta^{(\infty)}_{x_0}(0) = 0) > 0$. Fix a $T > 0$. By stationarity for all $t \le T$ the probability that a new immigrant enters the system (at x_0) in $(t, t+dt)$ is $\beta\,dt$ and because once born the immigrant performs independent p-motion the probability that she is back in x_0 at time T is then $p_{x_0x_0}(T-t)$. Thus

$$\mathbb{E}\,\eta^{(\infty)}_{x_0}(0) = \mathbb{E}\,\eta^{(\infty)}_{x_0}(T) \ge \beta \int_0^T p_{x_0x_0}(T-t)dt \longrightarrow \infty \quad \text{as } T \to \infty. \tag{3.3}$$

While this simple argument only shows that necessarily $\mathbb{E}\,\eta^{(\infty)}_{x_0}(t) \equiv \infty$ it can be be strengthened to

Lemma 12 *For recurrent p there exists no equilibrium for η concentrated on $\mathbb{S}$.*

Lemma 12 and Lemma 11 together yield

Theorem 2 *Let $\eta(0) \equiv 0$, assume $|S| > 1$. Then*

$$\eta_{x_0}(t) \to \infty \textit{ in probability} \iff p \textit{ is recurrent.}$$

Proof of Lemma 12. We argue by contradiction. Assume that $\eta(0)$ were distributed according to an ($\mathbb{S}$-valued) equilibrium for $(T(t))$. We denote $\beta := \mathbb{P}(\eta_{x_0}(0) = 0) > 0$. Then we can use the construction of section 3.1 to construct a stationary process η_t, $t \ge 0$. Recall the ingredients $N = (\tau_1, \tau_2, \ldots)$, a homogeneous rate 1 Poisson process of "potential immigrations", $(X^{(n)}(t))_{t\ge\tau_n}$, $n = 1, 2, \ldots$, the paths of these (potential) immigrants, and $(X^{(x,i)}(t))_{t\ge0}$, $x \in S$, $i = 1, 2, \ldots$, the paths of all individuals (and possibly many more) originally in the population. We denote as before by A_n the event that the n-th potential immigration is realized in the system. Fix a $T > 0$ for the moment and let

$$B_T := \sum_{n\ge1} \mathbf{1}(\tau_n \le T, A_n, X^{(n)}(T) = x_0)$$

be the number of immigrants that enter the system in $[0,T]$ and are back in x_0 at time T. By stationarity we can compute

$$\mathbb{E}\, B_T = \int_0^T \beta p_{x_0x_0}(T-t)\, dt = \beta \int_0^T p_{x_0x_0}(t)\, dt. \tag{3.4}$$

On the other hand

$$\begin{aligned} \mathbb{E}\,(B_T)^2 &\leq \sum_{m,n\geq 1} \mathbb{P}(\tau_m \leq T, \tau_n \leq T, X_T^{(m)} = x_0 = X_T^{(n)}) \\ &= \int_0^T p_{x_0x_0}(t)\, dt + \left(\int_0^T p_{x_0x_0}(t)\, dt\right)^2 \end{aligned} \tag{3.5}$$

(observe that the rhs in the first line above is just the expectation of the square of the number of marked points in $[0,T]$ of a Poisson process where a point at t is marked with probability $p_{x_0x_0}(T-t)$ independently of the others). This allows to estimate (using Cauchy-Schwarz for the first inequality)

$$\begin{aligned} \mathbb{P}(B_T \geq (\mathbb{E}\, B_T)/2) &= \mathbb{E}\left[\mathbf{1}(B_T \geq (\mathbb{E}\, B_T)/2)^2\right] \\ &\geq \frac{\left(\mathbb{E}\left[\mathbf{1}(B_T \geq (\mathbb{E}\, B_T)/2)\mathbf{1}(B_T \geq (\mathbb{E}\, B_T)/2)B_T\right]\right)^2}{\mathbb{E}\left[\mathbf{1}(B_T \geq (\mathbb{E}\, B_T)/2)^2 B_T^2\right]} \\ &\geq \frac{\left(\mathbb{E}\, B_T - \mathbb{E}\left[B_T\mathbf{1}(B_T < (\mathbb{E}\, B_T)/2)\right]\right)^2}{\mathbb{E}\, B_T^2} \geq \frac{1}{4}\frac{(\mathbb{E}\, B_T)^2}{\mathbb{E}\, B_T^2} \end{aligned}$$

But by construction $\eta_{x_0}(T) \geq B_T$, hence using (3.4) and (3.5)

$$\mathbb{P}\left(\eta_{x_0}(0) \geq (\beta/2)\int_0^T p_{x_0x_0}(t)\, dt\right) \geq \frac{\beta^2}{4}\frac{\left(\int_0^T p_{x_0x_0}(t)\, dt\right)^2}{\int_0^T p_{x_0x_0}(t)\, dt + \left(\int_0^T p_{x_0x_0}(t)\, dt\right)^2}.$$

By assumption $\int_0^T p_{x_0x_0}(t)dt \to \infty$ as $T \to \infty$ so that the rhs of the above tends to $\beta^2/4 > 0$ showing that $\mathbb{P}(\eta_{x_0}(0) = \infty) > 0$ which is in contradiction to $\eta(0)$ being in (a $\mathbb{S}$-valued) equilibrium.

3.4 Quantitative longtime behaviour I: The positive recurrent case

We have seen in the previous section that a system (η_t) of independent p-chains with self-blocking immigration (started off from $\eta(0) = 0$, say) experiences local overflow as $t \to \infty$ whenever p is recurrent. Thus a (possibly) natural question is the speed of divergence, i.e. how many particles do we typically see at some late time t? Here we consider the case that p is *positive recurrent* with stationary distribution π and present some evidence (and prove under a uniform exponential ergodicity assumption) that this number grows logarithmically.

3.4.1 Heuristics

Let η be as above, $\varphi(t) := \mathbb{P}(\eta_{x_0}(t) = 0)$. The expected number of particles at time t is then $\Phi(t) := \mathbb{E}\sum_x \eta_x(t) = \int_0^t \varphi(s)\, ds$. If at time t there are $\Phi(t)$ particles which perform independent p-motions the probability that all of them avoid x_0 at this instant might plausibly be something like

$$\varphi(t) \approx (1-\pi_{x_0})^{\Phi(t)} = \exp(\log(1-\pi_{x_0}) \int_0^t \varphi(s)\, ds). \tag{3.6}$$

This ansatz makes the implicit assumption that the immigration events are rare so that the particles' locations have had a lot of time to be close to equilibrium. Taking equality in (3.6) leads to the ordinary boundary value problem

$$\frac{d}{dt}\varphi(t) = \log(1-\pi_{x_0})\, \varphi(t)^2,\ t \geq 0, \quad \varphi(0) = 1$$

with solution $\varphi(t) = \big(1 + t \cdot \log(1/(1-\pi_{x_0}))\big)^{-1}$, leading to the conjecture

$$\Phi(t) \approx \frac{\log(t)}{\log(1/(1-\pi_{x_0}))}. \tag{3.7}$$

Under stronger assumptions this can indeed be proved, see the next subsection.

3.4.2 The "finite" case

From now on we make the stronger assumption (something like "uniform exponential ergodicity") that there exist constants $\kappa, \lambda > 0$ such that

$$\forall\, t: \sup_{x \in S} \mathrm{d}(P_x(X(t) \in \cdot), \pi) \leq \kappa\, e^{-\lambda t} \tag{3.8}$$

where P_x denotes the distribution of a p-chain X starting at $X(0) = x$ and d is the total variation distance. Recall that for probability measures μ, ν on a measurable space M the total variation distance is defined as $\mathrm{d}(\mu,\nu) := \sup |\mu(A) - \nu(A)|$ where the sup ranges over all measurable subsets $A \subset M$. Observe that (3.8) is quite stringent but is fulfilled whenever S is finite and p is irreducible. This is what the apostrophes in the title of this subsection refer to.

Theorem 3 *Let $N(t) := \sum_x \eta_t(x)$ be the overall number of particles at time t. Under assumption (3.8) starting from $\eta_0 \equiv 0$ we have*

$$\frac{N(t)}{\log_{1/a} t} \longrightarrow 1 \quad \text{in probability as } t \to \infty$$

where $a := \pi(S \setminus \{x_0\})$ and $\log_{1/a} t = (\log t)/(\log 1/a)$ is the logarithm with basis $1/a$.

Proof. We denote the time until the n-th immigrant enters by

$$T_n := \inf\{t \,:\, N(t) \geq n\}. \tag{3.9}$$

For $0 \leq x < 1/\log(1/a)$ we have

$$\begin{aligned} \mathbb{P}(N(t) \leq x \log t) &= \mathbb{P}(T_{[x \log t]} > t) \\ &= \mathbb{P}(T_{[x \log t]} a^{x \log t} > t a^{x \log t}) \to 0 \end{aligned}$$

as $t \to \infty$ because by Proposition 7 a) the family $(T_{[x \log t]} a^{x \log t})_{t>1}$ is tight and $t a^{x \log t} = t^{1 - x \log(1/a)} \to \infty$ by the choice of x. On the other hand for $x > 1/\log(1/a)$ we can choose $\tilde{a} > a$ such that still $x > 1/\log(1/\tilde{a})$ and then use Proposition 7 b) to see that as $t \to \infty$

$$\begin{aligned} \mathbb{P}(N(t) \geq x \log t) &= \mathbb{P}(T_{[x \log t]} \leq t) = \mathbb{P}\left(T_{[x \log t]} \tilde{a}^{[x \log t]} \leq t \tilde{a}^{[x \log t]}\right) \\ &\leq \mathbb{P}\left(T_{[x \log t]} \tilde{a}^{[x \log t]} \leq t \tilde{a}^{x \log t - 1}\right) \to 0. \end{aligned}$$

□

Lemma 13 *For $n \in \mathbb{N}$ let $\eta^{(n)}$ be a self-blocking system that starts with n particles independently distributed according to π, i.e. $\eta^{(n)}(0) = \sum_{i=1}^n \delta_{X_i}$ where $X_1, X_2, \dots$ are independent, π-distributed. Let $\tau^{(n)}$ be the time until in system $\eta^{(n)}$ the next, i.e. the $(n+1)$-th immigrant enters. We have*

$$\mathcal{L}((1-\pi_{x_0})^n \tau^{(n)}) \longrightarrow \mathrm{Exp}(1) \quad \text{as } n \to \infty.$$

Proof. Let $X^{(1)}, X^{(2)}, \dots$ be independent p-chains, N an independent homogeneous rate 1 Poisson point process on $\mathbb{R}_+$. As it will appear frequently throughout the proof we abbreviate (again) $a := \pi(S \setminus \{x_0\})$. Define

$$F^{(n)}(t) := \int_0^{a^{-n}t} \prod_{i=1}^n \mathbf{1}(X^{(i)}(s) \neq x_0)\, ds, \tag{3.10}$$

the "free time of x_0" (measured on a sped-up time scale). We have for $t \geq 0$

$$\begin{aligned} \mathbb{P}(a^n \tau^{(n)} > t) &= \mathbb{P}\big(N(\{s \leq a^{-n}t \,:\, \textstyle\prod_{i=1}^n \mathbf{1}(X_s^{(i)} \neq x_0)\}) = 0\big) \\ &= \mathbb{E}\big[\exp(-F^{(n)}(t))\big], \end{aligned}$$

so all we have to show is that

$$F^{(n)}(t) \to t \text{ in distribution as } n \to \infty. \tag{3.11}$$

We have $\mathbb{E}\, F^{(n)}(t) = t$ for all t and n, furthermore

$$\begin{aligned} \operatorname{Var} F^{(n)}(t) &= \mathbb{E}\left(\int_0^{a^{-n}t} \textstyle\prod_{i=1}^n \mathbf{1}(X^{(i)}(s) \neq x_0)\right)^2 - t^2 \\ &= 2\int\!\!\int_{0 \leq s_1 \leq s_2 \leq a^{-n}t} \mathbb{E}\Big[\textstyle\prod_{i=1}^n \mathbf{1}(X^{(i)}(s_1) \neq x_0 \neq X^{(i)}(s_2))\Big] - a^{2n}\, ds_1\, ds_2 \\ &= 2\int\!\!\int_{0 \leq s_1 \leq s_2 \leq a^{-n}t} \Big(\textstyle\sum_{x,y \neq x_0} \pi_x p_{xy}(s_2 - s_1)\Big)^n - a^{2n}\, ds_1\, ds_2 \\ &= 2\int\!\!\int_{0 \leq u_1 \leq u_2 \leq t} (f(a^{-n}(u_2 - u_1))^n - 1)\, du_1\, du_2 \end{aligned} \tag{3.12}$$

with

$$f(u) := \frac{1}{a^2} \sum_{x,y \neq x_0} \pi_x p_{xy}(u) = \frac{1}{a} - \frac{1}{a^2} \sum_{x \neq x_0} \pi_x p_{xx_0}(u). \tag{3.13}$$

Note that $0 \leq f(u) \leq 1/a$ for all $u \geq 0$, and that

$$f'(0) = -\frac{1}{a^2} \sum_{x \neq x_0} \pi_x r_{xx_0} < 0, \tag{3.14}$$

where $(r_{xy})_{x,y \in S}$ is the Q-Matrix of the p-chain X. As $p_{xx_0}(u)$ is continuous and strictly positive for any x and $u > 0$ we see from (3.13) that $\sup_{u \geq \epsilon} f(u) < 1/a$ for any $\epsilon > 0$. This together with (3.14) allows us to choose some $b > a$ such that

$$f(u) \leq \Big(\big(\frac{1}{b} - \frac{1}{a}\big)u + \frac{1}{a}\Big)\mathbf{1}(u \leq 1) + \frac{1}{b}\mathbf{1}(u > 1). \tag{3.15}$$

Additionally, (3.8) implies

$$|f(u)-1| \le \frac{1}{a^2}\sum_{x\neq x_0}\pi_x \sum_{y\neq x_0}|p_{xy}(u)-\pi(y)| \le \frac{\kappa}{a}\exp(-\lambda u). \tag{3.16}$$

Finally, we split the integral on the righthand side in (3.12) into three parts

$$\begin{aligned}
\operatorname{Var} F^{(n)}(t) &= 2 \iint\limits_{0\le u_1\le u_2\le t,\, u_2-u_1<a^n} f(a^{-n}(u_2-u_1))^n - 1\, du_1\, du_2 \\
&\quad + 2 \iint\limits_{\dots,\, a^n\le u_2-u_1<na^n} \dots\, du_1\, du_2 + 2 \iint\limits_{\dots,\, na^n\le u_2-u_1} \dots\, du_1\, du_2 \\
&=: 2\big(I_1+I_2+I_3\big)
\end{aligned}$$

and estimate them e.g. in the following way:

$$\begin{aligned}
I_1 &\le t\int_0^{a^n} \Big(\big(\frac{1}{b}-\frac{1}{a}\big)a^{-n}u+\frac{1}{a}\Big)^n du = ta^n \int_0^1 \Big(\big(\frac{1}{b}-\frac{1}{a}\big)v+\frac{1}{a}\Big)^n dv \\
&= ta^n\Big[\frac{1}{(n+1)}\frac{1}{1/b-1/a}\Big(\big(\frac{1}{b}-\frac{1}{a}\big)v+\frac{1}{a}\Big)^{n+1}\Big]_{v=0}^{v=1} \to 0
\end{aligned}$$

by (3.15), which also shows that

$$I_2 \le t\int_{a^n}^{na^n}\Big(\frac{1}{b}\Big)^n du = (n-1)\Big(\frac{a}{b}\Big)^n \to 0.$$

Using (3.16) we see that

$$\begin{aligned}
I_3 &\le t\int_{na^n}^{t}\Big(1+\frac{\kappa}{a}e^{-\lambda a^{-n}u}\Big)^n - 1\, du = tna^n\int_1^{ta^{-n}/n}\Big(1+\frac{\kappa}{a}e^{-\lambda n w}\Big)^n - 1\, dw \\
&\le t^2\Big\{\Big(1+\frac{\kappa}{a}e^{-\lambda n}\Big)^n - 1\Big\} = t^2\Big\{\exp\Big(n\log\big(1+\frac{\kappa}{a}e^{-\lambda n}\big)\Big)-1\Big\} \\
&\le t^2\Big\{\exp\Big(n\frac{\kappa}{a}e^{-\lambda n}\Big)-1\Big\} \to 0.
\end{aligned}$$

Thus $\operatorname{Var} F^{(n)}(t)\to 0$ as $n\to\infty$ and we conclude, e.g. with Chebyshev's inequality, that (3.11) holds true. □

Remark. Of course assumption (3.8) is a bit overkill here when all we need is that $\operatorname{Var} F^{(n)}(t)$ converges to 0. However, (3.8) will be needed later on.

Now consider a system η that starts with the empty configuration $\eta(0)=0$, and let

$$T_n := \inf\big\{t : \textstyle\sum_x \eta_x(t) \ge n\big\}$$

be the time until the n-th immigrant enters. In view of the above we might expect T_n to be of the order $(1-\pi_{x_0})^{-n}$. Indeed we can show

Proposition 7 *a) The sequence* $((1-\pi_{x_0})^n T_n), n\in\mathbb{N}$ *is tight.*
b) For each $\gamma > 1-\pi_{x_0}$ *we have*

$$\gamma^n T_n \longrightarrow \infty \quad \textit{in probability as } n\to\infty.$$

Proof. a) We compare η with a system η' of "censored" self-blocking immigrants which behaves like η except that immediately after the n-th particle has entered the system η', any further immigration is prohibited for a time $K \log n$ irrespective of the number of particles at x_0 (K will be chosen suitably later on). We can couple η and η' such that $\eta'(0) = \eta(0)$ and $\eta'_\cdot(t) \leq \eta_\cdot(t)$ for all $t \geq 0$. Define T'_n analogously to T_n, obviously $T_n \leq T'_n$. Thus it suffices to show that there exists a real-valued r.v. Z such that

$$(1-\pi_{x_0})^n T'_n \longrightarrow Z \quad \text{in distribution as } t \to \infty. \tag{3.17}$$

In order to do this we write

$$T'_n = \theta_0 + K \log 1 + \theta_1 + K \log 2 + \theta_2 + \cdots + K \log(n-1) + \theta_{n-1}$$

where the time between the immigration of the $(n-1)$-th and the n-th particle is $K \log(n-1) + \theta_{n-1}$. Let $\tau^{(1)}, \tau^{(2)}, \cdots$ be a sequence of independent r.v.s where $\tau^{(n)}$ is distributed as in Lemma 13. The key step is to show that for all $m \leq n$

$$\mathrm{d}(\mathcal{L}(\theta_m, \theta_{m+1}, \ldots, \theta_n), \mathcal{L}(\tau^{(m)}, \tau^{(m+1)}, \ldots, \tau^{(n)})) \leq \sum_{j=m}^{n} \frac{\kappa}{j^{\lambda K - 1}}. \tag{3.18}$$

To see this observe first that by assumption (3.8)

$$\mathrm{d}(\mathcal{L}(\theta_m), \mathcal{L}(\tau^{(m)})) \leq m\kappa e^{-\lambda K \log m} = \frac{\kappa}{m^{\lambda K - 1}}.$$

Now fix $n > m$ and a (measurable) $B \subset \mathbb{R}^{n-m+1}$. Using the strong Markov property at the stopping time $T'_{n-1} + K \log(n-1)$ we see that

$$\begin{aligned} \mathbb{P}((\theta_m, \ldots, \theta_n) \in B) &= \mathbb{E}\left(\mathbb{E}\left(\mathbf{1}_B(\theta_m, \ldots, \theta_{n-1}) \,\middle|\, \mathcal{F}_{T'_{n-1} + K \log(n-1)}\right)\right) \\ &= \mathbb{E}\left(f_{m,n,B}(\eta'_{T'_{n-1} + K \log(n-1)}, \theta_m, \ldots, \theta_{n-1})\right), \end{aligned}$$

where

$$f_{m,n,B}(\eta, s_1, \ldots, s_{n-m}) := \mathbb{P}_\eta((s_1, \ldots, s_{n-m}, T_{\text{next}}) \in B)$$

with T_{next} being the time until the first new immigrant appears. Using assumption (3.8) again we see that

$$\left| f_{m,n,B}(\eta, s_1, \ldots, s_{n-m}) - \int_{\mathbb{R}^+} \mathbf{1}_B(s_1, \ldots, s_{n-m}, t) \mathbb{P}(\tau_n \in dt) \right| \leq n\, \kappa e^{-\lambda K \log n}.$$

Hence we can see inductively

$$\begin{aligned} &\left| \mathbb{P}((\theta_m, \ldots, \theta_n) \in B) - \mathbb{P}((\tau^{(m)}, \ldots, \tau^{(n)}) \in B) \right| = \\ &\quad \left| \int f_{m,n,B}(\eta, s_m, \ldots, s_{n-1}) \mathbb{P}(\eta'_{T'_{n-1} + K \log(n-1)} \in d\eta, \theta_m \in ds_m, \ldots, \theta_{n-1} \in ds_{n-1}) \right. \\ &\qquad \left. - \int \mathbb{P}((s_m, \ldots, s_{n-1}, \tau^{(n)}) \in B) \mathbb{P}(\tau^{(m)} \in ds_m, \ldots, \tau^{(n-1)} \in ds_{n-1}) \right| \\ &\quad \leq \frac{\kappa}{n^{\lambda K - 1}} + \left| \int \mathbb{P}((s_m, \ldots, s_{n-1}, \tau^{(n)}) \in B) \big(\mathbb{P}((\theta_m, \ldots, \theta_{n-1}) \in d(s_m, \ldots, s_{n-1})) \right. \\ &\qquad \left. - \mathbb{P}((\tau^{(m)}, \ldots, \tau^{(n-1)}) \in d(s_m, \ldots, s_{n-1})) \big) \right| \\ &\quad \leq \frac{\kappa}{n^{\lambda K - 1}} + \mathrm{d}\left(\mathcal{L}(\theta_m, \ldots \theta_{n-1}), \mathcal{L}(\tau^{(m)}, \ldots, \tau^{(n-1)})\right) \end{aligned}$$

(Remember that $|\int f(d\mu - d\nu)| \le \mathrm{d}(\mu,\nu)$ for any (measurable) f with values in $[0,1]$). Finally Lemma 13 and estimate (3.18) with a $K \ge 3/\lambda$ imply (3.17).

b) It suffices to show that for each $\gamma > 1 - \pi_{x_0}$ $(=: a)$

$$\mathbb{P}(T_n \ge \gamma^{-n}) \longrightarrow 1. \tag{3.19}$$

We initially start the system with 0 particles. At time $T_{n-[\sqrt{n}]-1}$ we observe the first $n - [\sqrt{n}] - 1$ particles. Denote the positions of these particles at time $T'_n := T_{n-[\sqrt{n}]-1} + n^{1/4}$ by $Y_1, \dots, Y_{n-[\sqrt{n}]-1}$ (there might be more particles around at time T'_n but here we keep only track of those that are older than $n^{1/4}$). By assumption (3.8) we have as $n \to \infty$

$$d_n := \mathrm{d}\left(\mathcal{L}(Y_1, \dots, Y_{n-[\sqrt{n}]-1}), \pi^{\otimes(n-[\sqrt{n}]-1)}\right) \le (n - [\sqrt{n}] - 1)\kappa e^{-\lambda n^{1/4}} \to 0.$$

Observe that $\mathbb{P}(T_n \le T'_n) \to 0$ because otherwise there would have to be $[\sqrt{n}]+1$ immigrations over a time period of length $n^{1/4}$ — this becomes more and more improbable even without any blocking of the immigration mechanism (to be precise, one could bound it above by the probability that a Poisson$(n^{1/4})$-r.v. exceeds $[\sqrt{n}]+1$). Now on the event $\{T_n > T'_n\}$ there are at most $n-1$ particles in the system at time T'_n and $T_n - T'_n$ is stochastically larger than the time that it takes until in a system η' started off from $\sum_{i=1}^{n-[\sqrt{n}]-1} \delta_{Y_i}$ the next immigrant appears — we can arrange a coupling by simply ignoring blockings in η caused by particles that were born after $T_{n-[\sqrt{n}]-1}$. Hence (using the notation of Lemma 13)

$$\begin{aligned}
\mathbb{P}(T_n \ge \gamma^{-n}) &\ge \mathbb{P}(T_n \ge \gamma^{-n}, T_n > T'_n) \\
&\ge \mathbb{P}\left(\inf\{s:\ \textstyle\sum_x \eta'_x(s) \ge n - [\sqrt{n}]\} \ge \gamma^{-n}\right) - P(T_n \le T'_n) \\
&\ge \mathbb{P}\left(\tau^{(n-[\sqrt{n}]-1)} \ge \gamma^{-n}\right) - d_n - \mathbb{P}(T_n \le T'_n) \\
&= \mathbb{P}\left(a^{(n-[\sqrt{n}]-1)}\tau^{(n-[\sqrt{n}]-1)} \ge a^{(n-[\sqrt{n}]-1)}\gamma^{-n}\right) - d_n - \mathbb{P}(T_n \le T'_n) \to 1
\end{aligned}$$

as $n \to \infty$ by Lemma 13 because $a^{(n-[\sqrt{n}]-1)}\gamma^{-n} \to 0$. □

Remark 12 1) The proof of Proposition 7 shows that the Z appearing on the rhs of (3.17) can be represented as $\sum_{i=0}^{\infty}(1-\pi_{x_0})^i Z_i$ where the Z_i are i.i.d. Exp(1).

2) It is unclear if $(1-\pi_{x_0})^n T_n$ also converges in distribution. At least for the case $S = \{0,1\}$ and symmetric p I have a proof that $2^{-n}T_n$ indeed has a limit (of the above form).

3.5 Quantitative longtime behaviour II: 1-d symmetric simple random walk

In view of section 2.8, the most interesting case is that of recurrent random walk. Unfortunately, this case is also more difficult, and we have only partial answers, even for the case of symmetric simple random walk on $\mathbb{Z}$ on which we concentrate in the following.

Let η be a system of self-blocking immigrants on $\mathbb{Z}$ with $\eta_\cdot(0) = 0$, denote $\varphi(t) := \mathbb{P}(\eta_0(t) = 0)$ and let $N(t) := \sum_x \eta_x(t)$ be the number of particles that immigrated up to time t. Heuristics suggests that $\varphi(t)$ should decay a little slower than $1/\sqrt{t}$ or equivalently $\mathbb{E}\,N(t) = \int_0^t \varphi(s)\,ds$ should grow a little faster than $\sqrt{t}$. The intuitive idea is the following: At time t there are about $\mathbb{E}\,N(t)$ particles, most of which have an age comparable to t (because as the birth rate is trivially ≤ 1 there can be no "explosion of newly-borns") and are thus spread out on an interval roughly of size $\sqrt{t}$. Now if $\mathbb{E}\,N(t) \le t^{1/2-\varepsilon}$ was eventually true, the expected number of particles at the origin would be $\approx \mathbb{E}\,N(t)/\sqrt{t} \to 0$ as $t \to \infty$ showing that $\varphi(t)$ would not converge to 0.

On the other hand if $N(t) \ge t^{1/2+\varepsilon}$ held true for all sufficiently large t, we would have at least a density of about t^ε on the interval $[-\sqrt{t}, \sqrt{t}]$. If the "$t^{1/2+\varepsilon}$ growth regime" did hold at time $2t$ there would have to be $\approx 2^{1/2+\varepsilon}$ times more particles than at time t. But we expect the dense mass of particles existing at time t to become similar to a Poisson process with constant intensity t^ε over the time interval $(t, 2t]$, at least if we look at a small window around the origin. Thus the system could up to time $2t$ let at most $\approx t\exp(-t^\varepsilon)$ new immigrants enter even if we ignore blocking by particles born in $(t, 2t]$, which is by far too few to sustain the desired growth.

The heuristics that a system of many independent particles should be approximately Poisson leads to the ansatz

$$\varphi(t) = \mathbb{P}(\eta_0(t) = 0) \approx \exp(-\mathbb{E}\,\eta_0(t)) = \exp\left(-\int_0^t \varphi(s)p_0(t-s)\,ds\right), \tag{3.20}$$

where $(p_x(t))$ is the transition semigroup of symmetric simple random walk on $\mathbb{Z}$. Let $\rho_0(t)$ be the solution of the integral equation $\rho_0(t) = \int_0^t p_0(t-s)\exp(-\rho_0(s))\,ds$ we obtain by putting $\rho_0(t) := -\log\varphi(t)$ and assuming equality in (3.20), set $R(t) := \int_0^t \exp(-\rho_0(s))\,ds$.

The Poisson ansatz, which amounts to replacing the immigration by an "effective immigration", leads to the conjecture that

$$N(t)\big/R(t) \to 1 \quad \text{(in probability as } t \to \infty). \tag{3.21}$$

Furthermore, if (3.21) holds it is also natural to guess that the shape of $\eta(t)$ for a large t should look like a mixture of the transition probabilities weighted with the effective immigration probabilities, i.e. $\eta(t) \approx \sum_x \left(\int_0^t p_x(t-s)e^{-\rho_0(s)}\,ds\right)\delta_x$.

We shall see below that $\rho_0(t) \sim \frac{1}{2}\log t$ and $R(t) \sim (2/\pi)^{1/2} \times \sqrt{t}\log t$ as $t \to \infty$, see Lemma 14 with $\beta = \infty$, but we have only been able to prove a rather weak result in the direction of (3.21), namely

Proposition 8 ("very weak law") *Let η be a system of random walks on $\mathbb{Z}$ with self-blocking immigration at the origin, $\eta(0) \equiv 0$. Then for any $\epsilon > 0$*

$$\mathbb{P}\left(\textstyle\sum_x \eta_x(t) \ge t^{1/2+\epsilon}\right) \longrightarrow 0 \quad \textit{as } t \to \infty.$$

We prove this proposition in the next section using an adaptation of the so-called relative entropy method, which is well-known in the field of hydrodynamic limits, see e.g. [22], chapter 6.

Remark 13 The techniques in section 3.6 are in principle not restricted to $d = 1$, nor do they rely on strict nearest-neighbour motion. In fact, they can be applied to any random walk. The analogue of Lemma 14 can be proved along the same lines for simple random walk in $d = 2$, yielding a prediction of $\rho_0(t) = \log\log t - \log\log\log t - \log 2\pi + o(1)$ for the local density and $R(t) = (2\pi t\log\log t)/\log t$ for the total number of particles at time t in a self-blocking system on $\mathbb{Z}^2$. But observe that then the analogue of Proposition 8 for $d = 2$ (namely that the number of particles grows more slowly than $t^{1+\epsilon}$) is trivial. Unfortunately, the present method seems unable to capture the finer correction terms.

3.6 1-d soft self-blocking, the relative entropy method and the proof of a very weak law

Consider a system η of independent symmetric simple random walkers on $\mathbb{Z}$ where at rate $\exp(-\beta\eta_0(t-))$ a new particle enters at the origin. Here $\eta_x(t)$ denotes the number of particles at $x \in \mathbb{Z}$ at time t; $\beta \in \overline{\mathbb{R}}_+$ is a (fixed) parameter. We interpret $e^{-\infty n} := \mathbf{1}(n = 0)$. Observe that for $\beta = \infty$ this is the notorious self-blocking system.

η is a Markov process with generator given by

$$\begin{aligned} Lf(\eta) &= \frac{1}{2} \sum_{x\in\mathbb{Z},e=\pm1} \eta_x \left(f(\eta^{x,x+e}) - f(\eta)\right) + \exp(-\beta\eta_0)\left(f(\eta^{0,+}) - f(\eta)\right) \\ &=: L_{rw}f(\eta) + L_c f(\eta), \end{aligned} \tag{3.22}$$

where $\eta^{0,+} := \eta + \delta_0$ is the configuration we obtain by adding to η a particle at the origin. The "c" in L_c stands for "creation".

Here, we pursue the program indicated in the previous section: In order to obtain information on the long-time behaviour of $\eta(t)$, we want to compare it with a Poisson system, which is much easier to analyse. Observe that if η_0 really was Poisson(ρ)-distributed, we would have $\mathbb{E}_{\mathrm{Poi}(\rho)} \exp(-\beta\eta_0) = \exp(-\rho(1-e^{-\beta}))$. Let ρ be the solution of the following inhomogeneous (lattice) heat equation

$$\begin{aligned} \partial_t\rho_x(t) &= \frac{1}{2}\Delta\rho_x(t) + \delta_0(x)\exp(-(1-e^{-\beta})\rho_0(t)), \quad t > 0, x \in \mathbb{Z} \\ \rho_x(0) &\equiv 0 \end{aligned} \tag{3.23}$$

and for $t \geq 0$ let π_t be a probability measure on (finite) particle configurations on $\mathbb{Z}$ given by

$$\pi_t(\{\eta : \eta_{x_i} = k_i\}) = \prod_i \exp(-\rho_{x_i}(t))\frac{\rho_{x_i}(t)^{k_i}}{k_i!},$$

i.e. under π_t the configuration is product Poisson with local intensities given by $\rho(t)$. The long-time behaviour of ρ is given by the following lemma, which is a special case of Lemma 17 in subsection 3.6.1 (put $\alpha := 1$, $\gamma := 1 - e^{-\beta}$).

Lemma 14 *As t tends to infinity*

$$\rho_0(t) = \frac{1}{1-e^{-\beta}}\left(\frac{1}{2}\log t - \log\log t + \log\left((1-e^{-\beta})\sqrt{2\pi}\right)\right) + o(1), \tag{3.24}$$

$$\sum_x \rho_x(t) = \int_0^t e^{-(1-e^{-\beta})\rho_0(s)}ds \sim (1-e^{-\beta})^{-1}\left(\frac{2}{\pi}\right)^{1/2}\sqrt{t}\log t. \tag{3.25}$$

Our aim is to compare $\mathcal{L}(\eta(t))$ (when started off from $\eta(0) = \emptyset$) and π_t. We use the relative entropy method, which consists in computing (and suitably estimating) the relative entropy of $\mathcal{L}(\eta(t))$ with respect to π_t. If we can show that (3.28), the specific relative entropy, is small, then we see by the entropy inequality (cf. (3.36)) that the two systems are "macroscopically similar" in the sense that the distributions of functionals involving a positive fraction of all the particles are close. This method is well-known in the field of scaling limits of interacting particle systems, see e.g. [22], chapter 6 and the references there.

It is convenient to introduce a (product) reference measure μ on (finite) particle configurations. Choose any summable[1] $m : \mathbb{Z} \to (0, \infty)$ and let μ be the product Poisson measure with $\int \eta_x \mu(d\eta) = m_x$, $x \in \mathbb{Z}$.

Let q_t be the density of $\mathcal{L}(\eta(t))$ with respect to μ. By general theory for continuous-time Markov chains, it solves

$$\partial_t q_t(\eta) = L^* q_t(\eta), \tag{3.26}$$

where L^* is the adjoint in $L^2(\mu)$ of the generator L of $(\eta(t))$ given in (3.22).

Furthermore let p_t be the density of π_t with respect to μ. As π_t is a product measure one easily computes

$$p_t(\eta) = \exp\left(\sum m_x\right) \exp\left(\sum -\rho_x(t) + \eta_x(\log \rho_x(t) - \log m_x)\right)$$

and checks that

$$\begin{aligned}
\partial_t p_t(\eta) &= p_t(\eta) \sum_x \left\{ -\partial_t \rho_x(t) + \eta_x \frac{\partial_t \rho_x(t)}{\rho_x(t)} \right\} \\
&= p_t(\eta) \Bigg\{ - \exp(-(1 - e^{-\beta})\rho_0(t)) \\
&\qquad + \sum_x \eta_x \frac{1}{\rho_x(t)} \left(\frac{1}{2}\rho_{x+1}(t) + \frac{1}{2}\rho_{x+1}(t) - \rho_x(t) \right) \\
&\qquad + \eta_0 \frac{1}{\rho_0(t)} \exp(-(1 - e^{-\beta})\rho_0(t)) \Bigg\} \\
&= \frac{1}{2} \sum_x \eta_x \left(\frac{m_{x+1}}{m_x} p_t(\eta^{x,x+1}) + \frac{m_{x-1}}{m_x} p_t(\eta^{x,x-1}) - 2p_t(\eta) \right) \\
&\qquad + \exp(-(1 - e^{-\beta})\rho_0(t)) \left(\frac{\eta_0}{\rho_0(t)} - 1 \right) p_t(\eta)
\end{aligned} \tag{3.27}$$

Note that the first term in the last equation is the adjoint of L_{rw} with respect to $L^2(\mu)$.

The relative entropy of $\mathcal{L}(\eta(t))$ with respect to π_t is

$$H(\mathcal{L}(\eta(t))|\pi_t) = \sum_\eta \mu(\eta) q_t(\eta) \log(q_t(\eta)/p_t(\eta)),$$

where the sum ranges over all configurations of finitely many particles. Our ultimate goal will be to estimate

$$H(\mathcal{L}(\eta(t))|\pi_t) \Big/ \sum_x \rho_x(t) \tag{3.28}$$

i.e. to control the per particle relative entropy. Since $\mathcal{L}(\eta(0)) = \pi_0 = \delta_\emptyset$ we have

$$H(\mathcal{L}(\eta(t))|\pi_t) = \int_0^t \partial_s H(\mathcal{L}(\eta_s)|\pi_s)\, ds.$$

Thus we have to compute

$$\partial_t H(\mathcal{L}(\eta(t))|\pi_t) = \sum_\eta \mu(\eta) \left\{ \partial_t q_t(\eta) \log \frac{q_t(\eta)}{p_t(\eta)} + q_t(\eta) \frac{\partial_t q_t(\eta)}{q_t(\eta)} - q_t(\eta) \frac{\partial_t p_t(\eta)}{p_t(\eta)} \right\}.$$

[1]It would seem natural to choose $m \equiv 1$ but then there would be infinitely many particles under μ so that it would be singular to $\mathcal{L}(\eta(t))$ and π_t. Another way out would be to first consider everything on $\mathbb{Z}/(n\mathbb{Z})$ and then let $n \to \infty$.

The summation over the second term yields 0 because $\sum_\eta \mu(\eta)q_t(\eta) \equiv 1$. Using (3.26) in the first term we get

$$\partial_t H(\mathcal{L}(\eta(t))|\pi_t) = \sum_\eta \mu(\eta) \left\{ q_t(\eta) L \left(\log \frac{q_t}{p_t} \right)(\eta) - \frac{q_t(\eta)}{q_t(\eta)} \partial_t p_t(\eta) \right\}.$$

Using the elementary estimate $\log b - \log a \le (b-a)/a$ and the fact that L_{rw} is the generator of a (pure jump) Markov process we see that for any positive function f (such that $\log f$ is in the domain of L_{rw}) on particle configurations

$$L \log f(\eta) \le \frac{1}{f(\eta)} L f(\eta).$$

Hence we can estimate

$$\begin{aligned} \partial_t H(\mathcal{L}(\eta(t))|\pi_t) &\le \sum_\eta \mu(\eta) p_t(\eta) L \left(\frac{q_t}{p_t} \right)(\eta) - \sum_\eta \mu(\eta) \frac{q_t}{p_t}(\eta) \partial_t p_t(\eta) \\ &= \sum_\eta \mu(\eta) \frac{q_t(\eta)}{p_t(\eta)} \left\{ (L^* - \partial_t) p_t(\eta) \right\}. \end{aligned}$$

One easily checks that L_c^*, the adjoint of L_c with respect to $L^2(\mu)$, is given by $L_c^* f(\eta) = \exp(-\beta\eta_0) \left(\frac{e^\beta}{m_0} \eta_0 f(\eta^{0,-}) - f(\eta) \right)$, where $\eta^{0,-} = \eta - \delta_0$ is obtained from η by removing one particle at 0. This allows to compute

$$\begin{aligned} (L^* - \partial_t) p_t(\eta) &= L_c^* p_t(\eta) - \exp(-(1-e^{-\beta})\rho_0(t)) \left(\frac{\eta_0}{\rho_0(t)} - 1 \right) p_t(\eta) \\ &= p_t(\eta) \left\{ e^{-\beta\eta_0} \left(\frac{e^\beta \eta_0}{\rho_0(t)} - 1 \right) - e^{-\rho_0(t)(1-e^{-\beta})} \left(\frac{\eta_0}{\rho_0(t)} - 1 \right) \right\} \\ &=: p_t(\eta) V_t(\eta_0). \end{aligned}$$

(observe that $p_t(\eta^{0,-}) = p_t(\eta) m_0/\rho_0(t)$). Hence we arrive at the following bound on the increase of relative entropy

$$\partial_t H(\mathcal{L}(\eta(t))|\pi_t) \le \int \frac{q_t}{p_t}(\eta) V_t(\eta_0) d\pi_t(\eta) \tag{3.29}$$

Remark that $\int V_t(\eta_0) d\pi_t(\eta) = 0$. Observing that q_t/p_t is a probability density with respect to π_t, (3.29) immediately yields

$$\partial_t H(\mathcal{L}(\eta(t))|\pi_t) \le \sup_{n \in \mathbb{Z}_+} V_t(n). \tag{3.30}$$

The proof of the following lemma is a lengthy, but straightforward calculus exercise, and we omit it.

Lemma 15 *Let $f_r : \mathbb{R}_+ \to \mathbb{R}$ for $r > 0$ be given by*

$$f_r(x) = e^{-\beta x} \left(\frac{e^\beta x}{r} - 1 \right) - e^{-r(1-e^{-\beta})} \left(\frac{x}{r} - 1 \right). \tag{3.31}$$

f_r assumes its maximal value

$$\begin{aligned} \sup_{\mathbb{R}_+} f_r &= \left\{ \frac{1}{\beta r} \left(\mathrm{W}(e^{r(e^{-\beta}(\beta+1)-1)-\beta+1}) + \mathrm{W}(e^{r(e^{-\beta}(\beta+1)-1)-\beta+1})^{-1} - 2 \right) + 1 - e^{-\beta} \right\} \\ &\qquad \times \exp(-r(1-e^{-\beta})) \end{aligned} \tag{3.32}$$

at
$$x_*(r) = \frac{1}{\beta} + re^{-\beta} - \frac{1}{\beta}\mathrm{W}(\exp(r(e^{-\beta}(\beta+1)-1)-\beta+1)).$$

Here, W denotes (the principal branch of) Lambert's W function, i.e. $xe^x = y \iff x = \mathrm{W}(y)$.

Observe that for any $\beta > 0$ we have $e^{-\beta}(\beta+1)-1 < 0$, hence for large r the biggest term inside the curly brackets in (3.32) will be $\mathrm{W}(\dots)^{-1}$. Furthermore $\mathrm{W}(0) = 0$, $\mathrm{W}'(0) = 1$, thus

$$\lim_{r\to\infty} \frac{1}{r} \log f_r(x_*(r)) = 1 - e^{-\beta}(\beta+1) - (1-e^{-\beta}) = -\beta e^{-\beta} \tag{3.33}$$

Denote the system with immigration rate $\exp(-\beta\#\{\text{particles at } 0\})$ by $(\eta^{(\beta)}(t))_{t\geq 0}$, the corresponding product Poisson laws by $(\pi_t^{(\beta)})_t$. We conclude from (3.33) and Lemma 17 below that for any $\epsilon > 0$

$$\begin{aligned} H(\mathcal{L}(\eta^{(\beta)}(t))|\pi_t^{(\beta)}) &= O\left(\int_0^t \exp(-(1-\epsilon/2)\beta e^{-\beta}\rho_0^{(\beta)}(s))\,ds\right) \\ &= O\left(\int_0^t \exp(-(1-\epsilon)\frac{\beta e^{-\beta}}{1-e^{-\beta}}\frac{1}{2}\log s)\,ds\right) \\ &= O\left(\exp\left\{\left(1-\frac{1}{2}(1-\epsilon)\frac{\beta e^{-\beta}}{1-e^{-\beta}}\right)\log t\right\}\right). \end{aligned} \tag{3.34}$$

Under $\pi_t^{(\beta)}$ the total number of particles $\sum_x \eta_x$ is Poisson distributed with mean $\sum_x \rho_x^{(\beta)}(t)$. For a Poisson($y$)-distributed random variable Z and $a > 1$ we have

$$\mathbb{P}_{\mathrm{Poi}(y)}(Z > y^a) = \exp\left(-(a-1)y^a \log y + O(y^a)\right),$$

which can be easily checked by the form of the Poisson weights and Stirling's formula. Lemma 17 shows that $\sum_x \rho_x^{(\beta)}(t) \sim C_\beta\sqrt{t}\log t$, hence for any $\delta > 0$

$$\pi_t^{(\beta)}\left(\sum_x \eta_x \geq t^{1/2+\delta}\right) \geq \pi_t^{(\beta)}\left(\sum_x \eta_x \geq (C_\beta\sqrt{t}\log t)^{1+2\delta}\right) = \exp\left(-\Theta(t^{1/2+\delta})\right). \tag{3.35}$$

Combining (3.34), (3.35) and the entropy inequality (see e.g. [22], Prop. A1.8.2)

$$\mathbb{P}(\eta^{(\beta)}(t) \in A) \leq \frac{\log 2 + H(\mathcal{L}(\eta^{(\beta)}(t))|\pi_t^{(\beta)})}{\log(1+1/\pi_t^{(\beta)}(A))} \quad \text{for all measurable } A \subset (\mathbb{N}_0)^{\mathbb{Z}} \tag{3.36}$$

we obtain the following weak (and β-dependent) upper bounds on the total number of particles by choosing $A := \left\{\eta : \sum_x \eta_x \geq t^{1/2+\delta}\right\}$ in (3.36):

Lemma 16
$$\mathbb{P}\left(\sum_x \eta_x^{(\beta)}(t) \geq t^{1/2+\delta}\right) \longrightarrow 0 \quad \text{as } t\to\infty$$

provided
$$\delta > \frac{1}{2}\left(1 - \frac{\beta e^{-\beta}}{1-e^{-\beta}}\right). \tag{3.37}$$

We do not expect the bound on δ given in the statement of Lemma 16 to be sharp at all: Observe that the righthand side of (3.37) becomes smaller as $\beta \searrow 0$, even though smaller β means less blocking of immigrations. In view of (3.25), we would expect the total number of particles in $\eta^{(\beta)}(t)$ to grow like $\mathrm{const}_\beta \times \sqrt{t}\log t$ for any $\beta \in (0,\infty]$, where only the constant depends on β. The bound in (3.37) is an artefact of our method, note that e.g. by passing from (3.29) to (3.30), we give away information (alas, it is unclear how to proceed with (3.29) directly). Note also that we can never obtain a lower bound on the number of particles in this way because $\pi_t^{(\beta)}(\sum_x \eta_x = 0) = \exp(-O(\sqrt{t}\log t))$, which shows that we can not even make use of (3.34) in combination with (3.36) for $A = \{\eta \equiv 0\}$.

The method will also fail in dimension 2 where we expect about $(t \log\log t)/\log t$ particles at time t, so that there will be no room to play with the exponent.

Finally observe that the results of this section still tell us a little bit about "real" self-blocking immigration:

Proof of Proposition 8. For $\beta \le \beta' \le \infty$ we can couple $\eta^{(\beta)}$ and $\eta^{(\beta')}$ starting from the same initial condition in such a way that $\eta_x^{(\beta)}(t) \ge \eta_x^{(\beta')}(t)$ for all x and t. In particular, a "real" self-blocking system $\eta\,(=\eta^{(\infty)})$ is below any soft-blocking $\eta^{(\beta)}$. As the lower bound on δ in Lemma 16 converges to 0 as $\beta \to 0$ we obtain Proposition 8 by comparison with weaker and weaker soft-blocking systems. $\square$

3.6.1 Asymptotics of a semilinear lattice heat equation

Here we consider the long-time behaviour of the solution of the following inhomogeneous heat equation on $\mathbb{Z}$

$$\begin{aligned} \partial_t \rho_x(t) &= \frac{1}{2}\Delta\rho_x(t) + \alpha\delta_0(x)\exp(-\gamma\rho_0(t)),\ t>0,\ x\in\mathbb{Z} \\ \rho_x(0) &\equiv 0, \end{aligned} \tag{3.38}$$

where $\Delta f_x = f_{x+1} + f_{x-1} - 2f_x$ is the 1-dimensional lattice Laplacian. Observe that $\frac{1}{2}\Delta$ is the generator of a random walk with variance t at time t. $\alpha, \gamma > 0$ are parameters.

Remark 14 Let ρ be the solution of (3.38). Then $\vartheta_x(t) := \gamma\rho_x(t)$ solves $\partial_t\vartheta_x(t) = \frac{1}{2}\Delta\vartheta_x(t) + \alpha'\delta_0(x)\exp(-\vartheta_0(t))$ with $\alpha' := \alpha\gamma$, hence it suffices to consider the case $\gamma = 1$.

Lemma 17 *Let ρ be the solution of (3.38). As t tends to infinity we have*

$$\rho_0(t) = \frac{1}{\gamma}\left\{\frac{1}{2}\log t - \log\log t + \log\sqrt{2\pi} + \log(\alpha\gamma)\right\} + o(1), \tag{3.39}$$

$$\sum_x \rho_x(t) = \alpha\int_0^t e^{-\gamma\rho_0(s)}ds \sim \frac{1}{\gamma}\left(\frac{2}{\pi}\right)^{1/2}\sqrt{t}\log t. \tag{3.40}$$

Proof of Lemma 17. We assume w.l.o.g. $\gamma = 1$, cf. Remark 14. Let $p_x(t) = \mathbb{P}_0(X_t = x)$ be the probability that a symmetric simple random walker on $\mathbb{Z}$ started at 0 is at $x \in \mathbb{Z}$ at time t. Writing (3.38) in integral form we get

$$\rho_x(t) = \int_0^t \alpha p_x(t-s)\exp(-\rho_0(s))\,ds,$$

in particular $\rho_0(t)$ is the solution of

$$f(t) = \int_0^t \alpha p_0(t-s)\exp(-f(s))\,ds, \quad t \geq 0. \tag{3.41}$$

Let us call a function $\bar{\varphi} : \mathbb{Z}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ with $\bar{\varphi}_\cdot(0) \equiv 0$ a *strict supersolution to (3.38)* if it solves

$$\begin{aligned} &\partial_t \bar{\varphi}_x(t) = \frac{1}{2}\Delta\bar{\varphi}_x(t) + \alpha r^{\bar{\varphi}}(t)\delta_0(x), \quad t > 0, x \in \mathbb{Z} \\ &\text{with an } r^{\bar{\varphi}}(t) > \exp(-\bar{\varphi}_0(t)). \end{aligned} \tag{3.42}$$

Then we see that $\bar{\varphi}_0(t) \geq \rho_0(t)$ for all $t \geq 0$:
$\psi_x(t) := \bar{\varphi}_x(t) - \rho_x(t)$ solves

$$\partial_t \psi_x(t) = \frac{1}{2}\Delta\psi_x(t) + \alpha\underbrace{\left(r^{\bar{\varphi}}(t) - e^{-\rho_0(t)}\right)}_{=:r^{\psi}(t)}\delta_0(x)$$

and $\psi_0(t) > 0$ for small t. Assume that $t_0 := \inf\{t : \psi_0(t) < 0\} < \infty$. Then we would have $\psi_0(t_0) = 0$ by continuity, but also $\psi_x(t_0) \geq 0$ for all x. To see this observe that $\psi_x(t)$, $x \neq 0$ has a representation (ψ solves the heat equation away from 0, consider $\psi_0(t)$ as exogenous input)

$$\psi_x(t) = \int_0^t \psi_0(t-s)\mathbb{P}_x(T_0 \in ds) + \mathbb{E}_x[\psi_{X_t}(0); T_0 > t] \quad \left(= \mathbb{E}_x \psi_{X_{t\wedge T_0}}(t - (t \wedge T_0))\right)$$

where $T_0 := \inf\{s : X_s = 0\}$ (see Lemma 19). Hence $\psi_x(t_0) \geq 0$ for all x because $\psi(0) \equiv 0$ and $\psi_0(s) \geq 0$ for $0 \leq s \leq t_0$ by definition. Consequently $\Delta\psi_0(t_0) \geq 0$ and we conclude that $\alpha^{-1}\partial_t\psi_0(t_0) \geq r^{\psi}(t_0) > \exp(-\bar{\varphi}_0(t_0)) - \exp(-\rho_0(t_0)) = 0$ in contradiction to the definition of t_0.

Observe that we can construct a supersolution to (3.38) from a strict subsolution to (3.41): Assume $\underline{f} : [0,\infty) \to [0,\infty)$, satisfies $\underline{f}(0) = 0$ and

$$\underline{f}(t) < \int_0^t \alpha p_0(t-s)\exp(-\underline{f}(s))\,ds \quad \text{for } t > 0. \tag{3.43}$$

Then $\bar{\varphi}_x(t) := \int_0^t \alpha p_x(t-s)\exp(-\underline{f}(s))\,ds$ solves

$$\partial_t \bar{\varphi}_x(t) = \frac{1}{2}\Delta\bar{\varphi}_x(t) + \alpha\exp(-\underline{f}(t))\delta_0(x)$$

and in particular $\bar{\varphi}_0(t) > \underline{f}(t)$, hence $\exp(-\underline{f}(t)) > \exp(-\bar{\varphi}_0(t))$.

Similarly, if $\underline{\varphi}$ is a strict subsolution we have $\underline{\varphi}_0(t) \leq \rho_0(t)$ for all $t \geq 0$ and such a $\underline{\varphi}$ can be constructed analogously from a supersolution $\bar{f}$ to (3.41).

Observe that the solution ρ of (3.38) has the property that $\rho_0(t)$ is an increasing function: Obviously $\partial_t\rho_0(t) > 0$ for t small. Assume that $t_0 := \inf\{t : \partial_t\rho_0(t) < 0\} < \infty$. Then by continuity $\partial_t\rho_0(t_0) = 0$. Let x be a neighbour of 0. We have by the representation given in Lemma 19

$$\begin{aligned}
\partial_t \rho_x(t_0) &= \lim_h \frac{1}{h}\left[\int_0^{t_0} \rho_0(t_0-s)\mathbb{P}_x(T_0\in ds) - \int_0^{t_0-h} \rho_0(t_0-h-s)\mathbb{P}_x(T_0\in ds)\right] \\
&= \lim_h \int_0^{t_0-h} \frac{1}{h}(\rho_0(t_0-s)-\rho_0(t_0-h-s))\mathbb{P}_x(T_0\in ds) \\
&\quad + \lim_h \frac{1}{h}\int_{t_0-h}^{t_0} \rho_0(t_0-s)\mathbb{P}_x(T_0\in ds) \\
&= \int_0^{t_0} \partial_t\rho_0(t-s)\mathbb{P}_x(T_0\in ds) + \underbrace{\rho_0(0)}_{=0}\frac{\mathbb{P}_x(T_0\in dt)}{dt}|_{t=t_0} > 0
\end{aligned}$$

because $\partial_t\rho_0(t)>0$ in $[0,t_0)$ and $\mathrm{supp}(\mathcal{L}_x(T_0))=\mathbb{R}_+$. So we see that

$$\partial_t^2\rho_0(t_0) = \frac{1}{2d}\sum_{x=\pm 1}(\partial_t\rho_0(t_0)-\partial_t\rho_x(t_0)) - \partial_t\rho_0(t_0)\alpha\exp(-\rho_0(t_0)) > 0,$$

contradicting the definition of t_0.

Lemma 18 *Assume* $\gamma = 1$.
i) For $C < \log\left(\alpha\sqrt{2\pi}\right)$ *there exists a* $K>0$ *such that*

$$\underline{f}(t) := \begin{cases} \frac{1}{2}\log t - \log\log t + C & \text{if } t\geq K, \\ 0 & \text{if } 0\leq t<K \end{cases}$$

is a strict subsolution for (3.41).
ii) For $C > \log\left(\alpha\sqrt{2\pi}\right)$ *there exist* $K, K'>0$ *such that*

$$\bar{f}(t) := \begin{cases} \frac{1}{2}\log t - \log\log t + C & \text{if } t\geq K, \\ K' & \text{if } 0\leq t<K \end{cases}$$

is a strict supersolution for (3.41).

Proof of Lemma 18. This is a straightforward computation using the asymptotics $p_0(t)\sim(2\pi t)^{-1/2}$. Here are some details:
i). Let $e^{-C}=(1+\varepsilon)/(\alpha\sqrt{2\pi})$. Observe that $p_0(t)=e^{-t}\mathrm{I}_0(t)\geq 1/\sqrt{2\pi t}$ for $t\geq 0.5$ (I_0 is the modified Bessel function of the first kind with index 0). For $\underline{f}$ as in *i)* and any $t>K$ we can estimate

$$\begin{aligned}
&\textstyle\int_0^t p_0(t-s)\alpha e^{-\underline{f}(s)}\,ds \geq \frac{\alpha}{e^C\sqrt{2\pi}}\int_K^{t-0.5}\frac{\log s}{\sqrt{s(t-s)}}ds = \frac{1+\varepsilon}{2\pi}\int_{K/t}^{1-\frac{0.5}{t}}\frac{\log t+\log u}{\sqrt{u(1-u)}}du \\
&\textstyle\quad\geq \frac{1+\varepsilon}{2\pi}\left\{\log t\left[\int_0^1\frac{1}{\sqrt{u(1-u)}}du - 2\sqrt{K/t}-\sqrt{2/t}\right]+\int_0^1\frac{\log u}{\sqrt{u(1-u)}}du\right\}.
\end{aligned}$$

Observing that $\int_0^1(u(1-u))^{-1/2}du=\pi$ we see that there exists $n\,(=n(\varepsilon))\geq 1$ such that for all $K\geq 1$

$$\int_0^t \alpha p_0(t-s)e^{-\underline{f}(s)}\,ds \geq \frac{1+\varepsilon/2}{2}\log t > \underline{f}(t) \quad \text{whenever } t\geq nK.$$

On the other hand for $K<t<nK$ we have

$$\begin{aligned}
\int_0^t \alpha p_0(s)\exp(-\underline{f}(t-s))\,ds &\geq \frac{\alpha}{\sqrt{2\pi}}\int_{t-K}^t\frac{e^{-0}}{\sqrt{u}}du \geq \frac{1}{2\pi}\int_{(n-1)K}^{nK}u^{-1/2}du \\
&= \alpha(2/\pi)^{1/2}\left(\sqrt{nK}-\sqrt{(n-1)K}\right)\geq\alpha\sqrt{\frac{K}{2\pi n}}
\end{aligned}$$

and $\underline{f}(t) \le \log(nK)$. So we just have to chose $K \ge 1$ so big that $\sqrt{\frac{K}{2\pi n}} > \log(nK)$.
ii) can be treated similarly.

$\square$

Proof of Lemma 17, continued. We would like to conclude directly from this lemma that $\rho_0(t) = \frac{1}{2}\log t - \log\log t + \log \alpha\sqrt{2\pi} + o(1)$, but there seems to be no formal argument. So we resort to the following:

Constructing $\bar{\varphi}$ and $\underline{\varphi}$ satisfying $\underline{\varphi} \le \rho \le \bar{\varphi}$ from the functions given in Lemma 18 we see that

$$\rho_0(t) \sim \frac{1}{2}\log t \quad \text{as } t \to \infty, \tag{3.44}$$

but we need a finer result. Denoting $\xi(t) := \alpha\exp(-\rho_0(t))$ we can write (3.41) as $\rho_0 = p_0 * \xi$, after taking Laplace transforms this reads

$$\widehat{\rho_0}(\lambda) = \widehat{p_0}(\lambda)\widehat{\xi}(\lambda), \quad \lambda > 0. \tag{3.45}$$

It is well known that $\widehat{p_0}(\lambda) \sim (2\lambda)^{-1/2}$ as $\lambda \searrow 0$. From (3.44) and a Tauberian theorem (see e.g. [10], chap. XIII.5, Thm. 4, p. 423) we conclude that $\widehat{\rho_0}(\lambda) \sim \frac{1}{2\lambda}\log(1/\lambda)$, hence $\widehat{\xi}(\lambda) \sim (2\lambda)^{-1/2}\log(1/\lambda)$ for $\lambda \searrow 0$. Invoking the Tauberian theorem in the other direction we get

$$\alpha\exp(-\rho_0(t)) = \left((2\pi)^{-1/2}t^{-1/2}\log t\right)(1 + o(1)).$$

(3.39) follows by taking logarithms. Observe that the use of (this direction of) the Tauberian theorem is justified because ρ_0 and hence g is a monotone function. Finally observe that $\int(\log s)/\sqrt{s}\,ds = 2\sqrt{s}\log s - 4\sqrt{s}$ to obtain (3.40). $\square$

Lemma 19 *Let $\psi_\cdot(0) : \mathbb{Z}^d \to \mathbb{R}$ and $\psi_0(\cdot) : \mathbb{R}_+ \to \mathbb{R}$ be given real-valued continuous functions and define ψ on $\mathbb{Z}^d \times \mathbb{R}_+$ as the solution of the heat equation away from 0 with given boundary behavior, i.e. ψ solves*

$$\partial_t\psi_x(t) = \frac{1}{2}\Delta\psi_x(t), \quad x \in \mathbb{Z}^d \setminus \{0\}, t > 0.$$

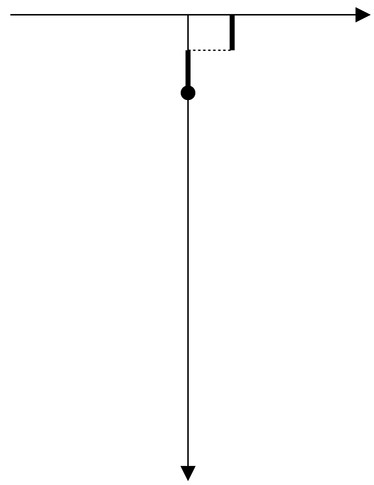
Then ψ has the stochastic representation

$$\psi_x(t) = \mathbb{E}_x\psi_{X(t\wedge T_0)}(t - (t \wedge T_0))$$

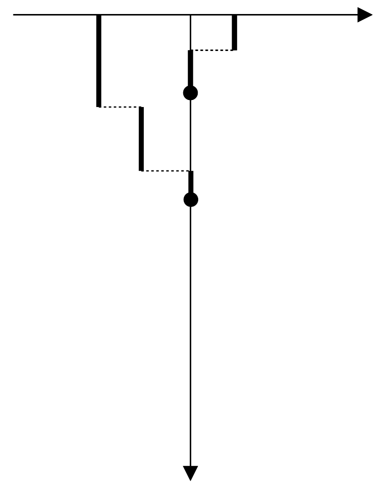
where $(X(s))$ is a continuous-time simple random walk on $\mathbb{Z}^d$, and $T_0 := \inf\{s > 0 : X(s) = 0\}$ the hitting time of the origin.

3.7 The self-blocking backward construction cartoon

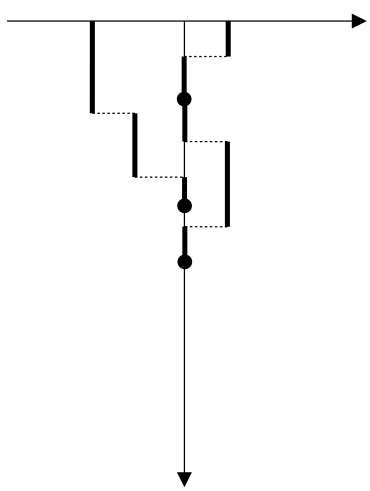
This cartoon illustrates the backward construction of section 3.2 for the historical process $(\Psi_t)_{t\leq 0}$ of a self-blocking system η: The horizontal axis denotes space (in this illustration a part of $\mathbb{R}$), the vertical axis is time in backward direction, their intersection is the space-time point $(x_0, 0)$.
We observe how at jump times $0 > \tau_1 > \tau_2 > \cdots$ of a Poisson process new paths try to enter the system but before they are allowed to do so they have to prove themselves in the space-time medium of the paths already there: The fat black dots on the time axis are "lethal" for paths that hit them from below.



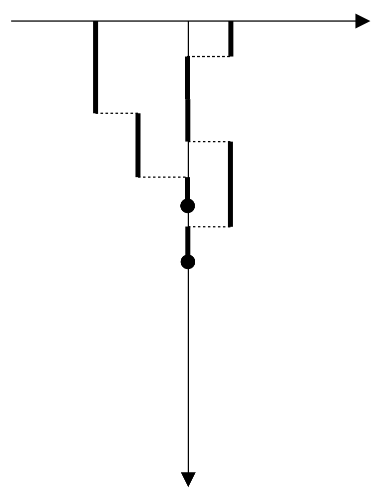
At (backward) time τ_1 the first path enters. A path is always marked by its birth-time (the fat dot). Ψ_{τ_1} consists of exactly this path.



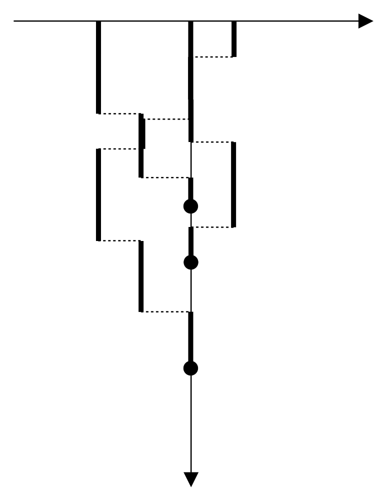
At time τ_2 a new path is added. As this path avoids the space-time point (x_0, τ_1) we have now two paths in Ψ_{τ_2}.



At time τ_3 we try to add a new particle but we see that this path hits the point(x_0, τ_1) which is the birth point of a path currently in Ψ so that we stop the new path there and glue it the old path...



... erasing the "old birthpoint": Ψ_{τ_3} still consists of two paths. The philosophy is that we realise that τ_1 can not have been a birth point because at that time there was already a particle at x_0 that was born earlier but we "recycle" the path that belonged to τ_1.



The path born at τ_4 does not hit any fat dots so far in the system so that we simply add it. Ψ_{τ_4} has three paths. Observe that the new path can pass the (former) birth point τ_1 undisturbedly (and away from birth points paths can interpenetrate anyway).

Chapter 4

Branching random walks in space-time i.i.d. random environment

4.1 Model

Individuals live on a countable state space S – we will consider mostly $\mathbb{Z}^d$ in the following – in discrete, non-overlapping generations. An individual at position x in generation n has k offspring with probability $Q_k(x,n)$, each child moves then independently to $y \in S$ with probability $p(x,y)$. The offspring of all individuals of the n-th generation form the $(n+1)$-st generation.

The random offspring distributions $Q(x,n), x \in S, n \in \mathbb{N}_0$ model the influence of a randomly fluctuating environment on the population. We assume that they are independent and identically distributed. Given the Q's, all individuals branch and move independently. For a probability measure $q = (q_k)$ on $\mathbb{N}_0$ we denote

$$m_1(q) := \sum_k k q_k, \quad m_2(q) := \sum_k k^2 q_k.$$

We assume furthermore that the random offspring distribution has at least two moments in the following sense:

$$\mathbb{E}[m_1(Q)],\ \mathbb{E}[m_2(Q)] < \infty. \tag{4.1}$$

Remark 15 1) Formally, we have a pair of processes, namely (population, environment). Our considerations center on the population (averaged over the environment), not conditional on a fixed realisation of the environments. This is sometimes called the "annealed case" in random media-parlance.

2) The following example is inspired by Greven's "coupled branching process" ([13]). Consider a $[0,1]$-valued i.i.d. space-time field U. We interpret $U(x,n)$, $x \in S, n \in \mathbb{N}_0$ as the survival probability of a child whose mother lived at position x in generation n: Each individual has a random number K of *potential offspring*, with a fixed distribution and independently of everything else, but a potential child survives only with probability $U(x,n)$. The surviving children form the next generation.

This amounts to considering a (possibly supercritical) classical spatial branching process with random thinning depending on its spatial embedding.

4.1.1 The Markov chain point of view

Let us denote the number of individuals at position x in generation n by $\eta_n(x)$. Then η_{n+1} arises, given η_n, in the following way:

$$\eta_{n+1}(x) = \sum_{y\in\mathbb{Z}^d} \sum_{i=1}^{\eta_n(y)} \sum_{j=1}^{K_n(y,i)} \mathbf{1}(Y_n(y,i,j) = x), \tag{4.2}$$

where

$Q(x,n), x \in \mathbb{Z}^d, n \in \mathbb{N}_0$ .... independent copies of Q, the random offspring distribution

$K_n(y,i)$ .... independent given $Q(\cdot,\cdot)$, with $\mathbb{P}(K_n(y,i) = k|Q(\cdot,\cdot)) = Q_k(y,n)$

$Y_n(y,i,j)$ .... independent, $\mathbb{P}(Y_\cdot(y,\cdot,\cdot) = x) = p(y,x) = p(0,y-x)$.

$Q(\cdot,\cdot)$ is the random environment, $K_n(y,i)$ is the number of descendants of the i-th particle at position y in generation n, and $Y_n(y,i,j)$ is the site that the j-th of these descendants jumps to.

Observe that, compared to the previous chapters which worked in continuous time, we have slightly changed the notation of objects which depend on space and time: It appeared more natural to us to put the discrete generation index in the subscript.

As in chapter 2, we follow Liggett and Spitzer [24] and consider the following state space:

$$E_1 := \{\eta \in (\mathbb{N}_0)^{\mathbb{Z}^d} : \textstyle\sum_x \gamma(x)\eta(x) < \infty\},$$

with $\gamma : \mathbb{Z}^d \to (0,\infty)$ summable, satisfying

$$\sum_y p(x,y)\gamma(y) \le M\gamma(x) \quad \text{for all } x \in S$$

for some suitable $M > 0$. Then our assumption $\mathbb{E}\, m_1(Q) < \infty$ implies

$$\eta_n \in E_1 \Longrightarrow \eta_{n+1} \in E_1 \text{ almost surely,}$$

because

$$\mathbb{E}[\eta_{n+1}(x)|\eta_n] = \sum_y \eta_n(y)\mathbb{E}[m_1(Q)]p(y,x),$$

and hence

$$\mathbb{E}\left[\textstyle\sum_x \eta_{n+1}(x)\gamma(x)|\eta_n\right] = \mathbb{E}[m_1(Q)] \sum_{x,y} \eta_n(y)p(y,x)\gamma(x) \le \mathbb{E}[m_1(Q)]M \sum_y \eta_n(y)\gamma(y).$$

Further on we assume that

$$\mathbb{E}[m_1(Q)] = 1, \tag{4.3}$$

and consequently a spatially homogeneous intensity measure is constant in time.

Let $(S_n)_n$ denote the transition semigroup generated by the dynamics of (η_n), i.e. for $\eta \in E_1$, $f : E_1 \to \mathbb{R}$ bounded

$$(S_n f)(\eta) := \mathbb{E}[f(\eta_n)|\eta_0 = \eta]. \tag{4.4}$$

Remark 16 Observe that S_n has the following continuity property (with respect to the product topology on E_1): $\mu_m \Rightarrow \mu$ and $\sup_{x,m} \int \eta(x)\mu_m(d\eta) < \infty$ imply $\mu_m S_1 \Rightarrow \mu S_1$ (and of course also $\mu_m S_n \Rightarrow \mu S_n$ by induction). To see this it suffices to reason that under the above assumptions for any finite set $A \subset \mathbb{Z}^d$ and $\epsilon > 0$ there exists a finite subset $B \subset \mathbb{Z}^d$ with the property $\mu_m(\{\text{some offspring of a particle in } B^c \text{ reaches } A \text{ in the next step}\}) \le \epsilon$ uniformly in m.

4.1.2 The genealogical point of view

Obviously our branching dynamics gives rise to a (random) genealogy, simply by recording each individual's mother — this information had been obliterated in the Markov chain-viewpoint.

The descendants of an individual in generation 0 form a rooted, ordered, spatially embedded tree. We formalize such a tree t as set of nodes together with a (cycle free) successor relation. We write $b \in \text{offspring}(a)$ if $a, b \in t$ and b is a child of a. Let $\text{root}(t)$ be the root of t, $|a|$ the height of node $a \in t$, i.e. its genealogical distance from the root. We consider ordered (or planar) trees, i.e. the successors $b_1, \dots, b_{|\text{offspring}(a)|}$ of a node a have an order, so that it makes sense to speak of the i-th child of a given individual. Let us denote the tree we obtain by truncating t at height N by $t|_N$. It arises from t by discarding all nodes b of height $|b| > N$. Finally, any node $a \in t$ (including the root) has a spatial position $\text{pos}(a) \in \mathbb{Z}^d$.

We denote the tree formed by the descendants of the i-th individual at position x in generation 0 ($1 \le i \le \eta_0(x)$) by $\tau_{x,i}$. We can specify the joint distribution of η_0, $(Q(x,n))_{x\in\mathbb{Z}^d, n\in\mathbb{N}_0}$ and $(\tau_{y,i})_{y\in\mathbb{Z}^d, i\in\mathbb{N}}$ in the following way (μ denotes the law of Q): For $N \in \mathbb{N}$, a finite $A \subset \mathbb{Z}^d$, $n_x \in \mathbb{N}_0$ for $x \in A$, $t_{x,i}$, $x \in A, i \le n_x$ (finite, rooted, ordered, spatially embedded) trees of height at most N, $\mathbb{Z}^d \supset B \supset A$ with $|B| < \infty$ and $\text{pos}(\sigma) \in B$ for all $\sigma \in t_{x,i}$, $x \in A, i \le n_x$ and (measurable) $C_{y,j} \subset \mathcal{M}_1(\mathbb{N}_0)$ for $y \in B$, $j < N$, we have

$$
\begin{aligned}
&\mathbb{P}\Big(\begin{array}{l} \{\eta_0(x) = n_x, x \in A\} \cap \{Q(y,j) \in C_{y,j} \text{ for } y \in B \text{ and } 0 \le j < N\} \\ \quad \cap \{\tau_{z,i}|_N = t_{z,i} \text{ for } z \in A \text{ and } 1 \le i \le n_z\} \end{array} \Big) \\
&\quad = \mathbb{P}(\eta_0(x) = n_x, x \in A) \times \\
&\qquad \int_{\mathcal{M}_1(\mathbb{N}_0)} \cdots \int_{\mathcal{M}_1(\mathbb{N}_0)} \prod_{y\in B, j<N} \mu(dq(y,j)) \mathbf{1}_{C_{y,j}}(q(y,j)) \Bigg\{ \\
&\qquad \prod_{x\in A} \prod_{1\le i\le n_x} \prod_{\sigma\in t_{x,i}} q_{|\text{offspring}(\sigma)|}(\text{pos}(\sigma), |\sigma|) \prod_{\sigma'\in\text{offspring}(\sigma)} p(\text{pos}(\sigma), \text{pos}(\sigma')) \Bigg\}.
\end{aligned}
$$

The probabilities of these events characterise the joint distribution.

4.2 Collision time of two independent random walks

Here we consider certain characteristics of the underlying individual motion of particles playing a part in determining the long-time behaviour of the system. These quantities are studied more closely in chapter 5.

Let X, X' be two independent $\bar{p}$-random walks, where $\bar{p}(x,y) := p(y,x)$ is the transposed transition matrix.

$$
\alpha_2 := \sup\left\{ \alpha > 0 : \mathbb{E}_{(0,0)}[\alpha^{\#\{i\ge 1: X_i = X'_i\}}] < \infty \right\}. \tag{4.5}
$$

Let us denote the difference random walk by $Z = X - X'$. Starting from $Z_0 = 0$ the number of visits of Z to 0 is geometrically distributed with parameter $\mathbb{P}_0(Z_i \ne 0, i \ge 1)$ by the Markov property, hence

$$
\alpha_2 = \frac{1}{1 - \mathbb{P}_0(Z_i \ne 0 \text{ for all } i \ge 1)}.
$$

In particular we see that $\alpha_2 > 1$ if and only if Z transient. Using the explicit form of the geometric weights we find that $\mathbb{E}_{(0,0)}\big[\alpha_2^{\#\{i\geq 1: X_i = X_i'\}}\big] = \infty$. Furthermore let

$$\alpha_* := \sup\Big\{\alpha > 0 : \mathbb{E}_{(0,0)}[\alpha^{\#\{i\geq 1: X_i = X_i'\}}\,|X] < \infty \text{ a.s.}\Big\}. \tag{4.6}$$

Obviously we have $\alpha_* \geq \alpha_2$. We have $\alpha_* > \alpha_2 > 1$ whenever the difference random walk $Z = X - X'$ is transient and p satisfies (the mild) assumption (5.1), see Corollary 4 in chapter 5. Obviously recurrence of Z implies $\alpha_2 = \alpha_* = 1$.

Remark 17 Since for a p-random walk X the Markov chain $(-X_n)_n$ is a $\bar{p}$-random walk and obviously $X_i = X_i' \iff -X_i = -X_i'$, we have $\alpha_2 = \alpha_2(p) = \alpha_2(\bar{p})$ and $\alpha_* = \alpha_*(p) = \alpha_*(\bar{p})$. Thus for the quantities defined in this section it is irrelevant if we consider the random walk dynamics p generating the individual motion of η-particles or its time reversal $\bar{p}$.

4.3 Two moments

Using a discrete "path integral representation" we find a condition on the distribution of the random offspring distribution which (in a shift-invariant scenario with a transient random walk as individual motion and under assumption (4.3) of "criticality") is necessary and sufficient for the global boundedness of the second moments of $\eta_n(x)$, $n \in \mathbb{N}$.

Lemma 20 *We have*
1)

$$\mathbb{E}\,\eta_n(x) = \sum_y \bar{p}^n(x,y)(\mathbb{E}\,m_1(Q))^n\mathbb{E}\,\eta_0(y) = (\mathbb{E}\,m_1(Q))^n\mathbb{E}_x\eta_0(\bar{X}_n).$$

2) Assume furthermore $\sup_x \mathbb{E}[\eta_0(x)^2] < \infty$.
Then with $\varphi_{x,y}(\eta) := \eta(x)\eta(y) - \delta_{x,y}\eta(x)$

$$\begin{aligned}\mathbb{E}\,\varphi_{x,x'}(\eta_n) \;&=\; \mathbb{E}_{(x,x')}\Bigg[\varphi_{\bar{X}_n,\bar{X}_n'}(\eta_0)\alpha^{\#\{1\leq i\leq n\,:\bar{X}_i=\bar{X}_i'\}}\beta^{\#\{1\leq i\leq n\,:\bar{X}_i\neq\bar{X}_i'\}}\\ &\qquad + \mathbb{E}[m_2(Q)-m_1(Q)]\sum_{i=1}^n \mathbf{1}(\bar{X}_i=\bar{X}_i')\alpha^{\#\{1\leq j\leq i-1\,:\bar{X}_j=\bar{X}_j'\}}\\ &\qquad\qquad \times\beta^{\#\{1\leq j\leq i-1\,:\bar{X}_j\neq\bar{X}_j'\}}\mathbb{E}\eta_{n-i}(\bar{X}_i)\Bigg],\end{aligned}$$

where $\alpha := \mathbb{E}[(m_1(Q))^2]$, $\beta := (\mathbb{E}\,m_1(Q))^2$. *Here* $\bar{X}$ *and* $\bar{X}'$ *are two independent random walks with transition matrix* $\bar{p}(x) = p(-x)$, *independent of* (η_n).

Proof. 1) Condition on η_{n-1} to see

$$\begin{aligned}\mathbb{E}[\eta_n(x)] \;&=\; \mathbb{E}[\mathbb{E}[\eta_n(x)|\eta_{n-1}]] = \mathbb{E}\left[\sum_y \eta_{n-1}(y)\mathbb{E}[m_1(Q)]p(y,x)\right]\\ &=\; \mathbb{E}[m_1(Q)]\sum_y \bar{p}(x,y)\mathbb{E}\,\eta_{n-1}(y).\end{aligned}$$

2) Observe that $\mathbb{E}\varphi_{x,x'}(\eta_n)$ is the expected number of ordered pairs of particles we can form by choosing the first particle from those living at position x in generation n and the second particle from those living at x' in the same generation. We decompose this number according to the possible provenances y, y' of the ancestors of these particles in generation $n-1$.

For $y \neq y'$, the total number of offspring born at y in generation $n-1$ that move to x in the next step is independent of the corresponding number of descendants of y'-parents that move to x', given the population η_{n-1}. Thus the conditional expectation of the number of such pairs is

$$\begin{aligned}(\eta_{n-1}(y)\mathbb{E}[m_1(Q)]p(y,x)) &\times (\eta_{n-1}(y')\mathbb{E}[m_1(Q)]p(y',x')) \\ &= \bar{p}(x,y)\bar{p}(x',y')(\mathbb{E}\, m_1(Q))^2\varphi_{y,y'}(\eta_{n-1}).\end{aligned}$$

For $y = y'$ we first consider the case $x = x'$. We write the number of offspring that move to x as $A_1 + \cdots + A_{\eta_{n-1}(y)}$ (where A_i are descendants of the i-th particle at y in generation $n-1$). Conditional on $\eta_{n-1}(y)$ and $Q(y, n-1)$, the A_i are independent and identically distributed with $\mathbb{E}[A_1|\eta_{n-1}(y), Q(y, n-1)] = m_1(Q(y, n-1))p(y,x)$ and $\mathbb{E}[(A_1)^2|\eta_{n-1}(y), Q(y, n-1)] = (m_2(Q(y, n-1)) - m_1(Q(y, n-1)))p(y,x)^2 + m_1(Q(y, n-1)p(y,x)$. Consequently we have

$$\begin{aligned}
&\mathbb{E}[(A_1 + \cdots + A_{\eta_{n-1}(y)})(A_1 + \cdots + A_{\eta_{n-1}(y)} - 1)|\eta_{n-1}(y), Q(y, n-1)] \\
&\quad = \eta_{n-1}(y)(\eta_{n-1}(y)-1)(m_1(Q(y,n-1)))^2 p(y,x)^2 \\
&\qquad + \eta_{n-1}(y)\big((m_2(Q(y,n-1)) - m_1(Q(y,n-1)))p(y,x)^2 + m_1(Q(y,n-1))p(y,x)\big) \\
&\qquad - \eta_{n-1}(y)m_1(Q(y,n-1)p(y,x) \\
&\quad = p(y,x)^2\Big\{\eta_{n-1}(y)(\eta_{n-1}(y)-1)(m_1(Q(y,n-1)))^2 \\
&\qquad\qquad + \eta_{n-1}(y)(m_2(Q(y,n-1)) - m_1(Q(y,n-1))\Big\} \\
&\quad = p(y,x)^2\Big\{\varphi_{y,y}(\eta_{n-1})(m_1(Q(y,n-1)))^2 \\
&\qquad\qquad + \eta_{n-1}(y)(m_2(Q(y,n-1)) - m_1(Q(y,n-1))\Big\}.
\end{aligned}$$

Finally for $y = y'$ and $x \neq x'$ we decompose the total number of offspring of particles at y that move to x in the next step as $A_1 + \cdots + A_{\eta_{n-1}(y)}$ and the corresponding number of children that move to x' as $B_1 + \cdots + B_{\eta_{n-1}(y)}$ (again A_i and B_i count descendants of the i-th particle at position y in generation $n-1$). We have $\mathbb{E}[A_iB_j|\eta_{n-1}(y), Q(y, n-1)] = m_1(Q(ym, n-1))^2 p(y,x)p(y,x')$ for $i \neq j$ and $\mathbb{E}[A_iB_i|\eta_{n-1}(y), Q(y, n-1)] = (m_2(Q(y, n-1)) - m_1(Q(y, n-1))p(y,x)p(y,x')$. Hence we can compute

$$\begin{aligned}
&\mathbb{E}[(A_1 + \cdots + A_{\eta_{n-1}(y)})(B_1 + \cdots + B_{\eta_{n-1}(y)})|\eta_{n-1}(y), Q(y, n-1)] \\
&\quad = \eta_{n-1}(y)(\eta_{n-1}(y)-1)m_1(Q(ym,n-1))^2 p(y,x)p(y,x') \\
&\qquad + \eta_{n-1}(y)(m_2(Q(y,n-1)) - m_1(Q(y,n-1))p(y,x)p(y,x') \\
&\quad = p(y,x)p(y,x')\Big\{\varphi_{y,y}(\eta_{n-1})m_1(Q(y,n-1))^2 \\
&\qquad\qquad + \eta_{n-1}(y)(m_2(Q(y,n-1)) - m_1(Q(y,n-1))\Big\}.
\end{aligned}$$

By taking the expectation with respect to η_{n-1} and $Q(\cdot, n-1)$ in the above identities and summing over all possible pairs (y, y') we find

$$\begin{aligned}\mathbb{E}\,\varphi_{x,x'}(\eta_n) \quad = \quad & \sum_{y,y'} \bar{p}(x,y)\bar{p}(x',y') \Big\{ \alpha^{\mathbf{1}(y=y')}\beta^{\mathbf{1}(y\neq y')}\mathbb{E}\,\varphi_{y,y'}(\eta_{n-1}) \\ & \qquad + \mathbf{1}(y=y')\mathbb{E}[m_2(Q(y,n-1)) - m_1(Q(y,n-1))]\,\mathbb{E}\,\eta_{n-1}(y) \Big\}.\end{aligned}$$

Our claim follows by iteration using a "discrete Feynman-Kac formula", see e.g. Lemma 22 at the end of this section. $\square$

Let us denote by $\mathcal{R}_\theta$ the set of all shift-invariant initial conditions $\mu \in \mathcal{M}_1(E_1)$ satisfying $\int \eta(x)\,d\mu \equiv \theta$, $\int \eta(0)^2\,d\mu < \infty$ and

$$\lim_{N\to\infty} \sum_{x,y\in\mathbb{Z}^d} \bar{p}_N(0,x)\bar{p}_N(0,y) \int \eta(x)\eta(y)\,\mu(d\eta) = \theta^2. \tag{4.7}$$

Observe that (4.7) is in this situation equivalent to $\sum_x \bar{p}_N(0,x)\eta(x) \to \theta$ in $L^2(\mu)$. Any shift-ergodic μ with $\int \eta(0)\,d\mu = \theta$ and $\int \eta(0)^2 d\mu < \infty$ lies in $\mathcal{R}_\theta$, see [24], Lemma 5.2 (note that [24] treats transition kernels for continuous-time random walks, but the adaptation to discrete-time random walks is obvious).

Lemma 21 *Assume (4.3). Let $\mathcal{L}(\eta_0)$ be shift invariant with $\mathbb{E}\eta_0(0) = \theta$, $\mathbb{E}\eta_0(0)^2 < \infty$, satisfying*

$$\mathrm{Cov}(\eta_0(x), \eta_0(y)) \to 0 \quad \textit{as } |x-y| \to \infty.$$

1) For any fixed n, $\mathcal{L}(\eta_n)$ also has these properties.

2) If $\alpha\,(= \mathbb{E}[m_1(Q)^2]) < \alpha_2$ we have furthermore

$$\sup_n \mathbb{E}\,\eta_n(0)^2 < \infty, \qquad \lim_{|x-y|\to\infty} \sup_n |\mathrm{Cov}(\eta_n(x), \eta_n(y))| = 0.$$

3) If $\alpha < \alpha_2$, $\mathcal{L}(\eta_0) \in \mathcal{R}_\theta$ and $\eta_{n_k} \to \eta_\infty$ in finite dimensional distributions along some subsequence $n_k \to \infty$, then also $\mathcal{L}(\eta_\infty) \in \mathcal{R}_\theta$.

4) If $\alpha \geq \alpha_2$ we have $\lim_n \mathbb{E}\,\eta_n(0)^2 = \infty$*, for $\alpha > \alpha_2$ even* $\lim_n \frac{1}{n} \log \mathbb{E}\,\eta_n(0)^2 > 0$.

Proof. Obviously $\mathcal{L}(\eta_n)$ is shift invariant whenever this is true for $\mathcal{L}(\eta_0)$. Lemma 20, part 1) shows that in this situation we have $\mathbb{E}\,\eta_n(x) \equiv \theta$ for all n. Furthermore we have for all $x, y \in \mathbb{Z}^d$

$$\mathbb{E}[\varphi_{x,y}(\eta_0)] \leq \mathbb{E}\,\eta_0(x)\eta_0(y) \leq \left(\mathbb{E}\,\eta_0(x)^2\mathbb{E}\,\eta_0(y)^2\right)^{1/2} = \mathbb{E}\,\eta_0(0)^2 =: \theta_2 < \infty.$$

Consequently Lemma 20, 2) allows to estimate (putting $\gamma := \mathbb{E}[m_2(Q) - m_1(Q)]$ for brevity)

$$\begin{aligned}\mathbb{E}\,\eta_n(0)^2 \quad \leq \quad & \theta + \theta_2\mathbb{E}_{(0,0)}\left[\alpha^{\#\{1\leq i\leq n\,:\bar{X}_i=\bar{X}'_i\}}\right] \\ & + \gamma\theta\mathbb{E}_{(0,0)}\left[\sum_{i=1}^n \mathbf{1}(\bar{X}_i = \bar{X}'_i)\alpha^{\#\{1\leq j\leq i-1\,:\bar{X}_j=\bar{X}'_j\}}\right].\end{aligned}$$

This is apparently finite for any n. Defining $B_0 := \{i \geq 1 : \bar{X}_i = \bar{X}'_i\}$ we can estimate uniformly in n:

$$\mathbb{E}\,\eta_n(0)^2 \leq \theta + \theta_2 \mathbb{E}_{(0,0)} \alpha^{B_0} + \gamma\theta \mathbb{E}_{(0,0)} B_0 \alpha^{B_0},$$

the rhs is finite for $\alpha < \alpha_2$. In order to treat the covariances we define $T := \min\{k : \bar{X}_k = \bar{X}'_k\}$ and again use Lemma 20 (for $x \neq y$) to compute

$$\begin{aligned}
&\mathrm{Cov}(\eta_n(x), \eta_n(y)) = \mathbb{E}\varphi_{x,y}(\eta_n) - \theta^2 \\
&\quad = \mathbb{E}_{(x,y)}\Big[\mathbf{1}(T > n)\mathbb{E}^{\eta}[\varphi_{\bar{X}_n, \bar{X}'_n}(\eta_0)]\Big] - \theta^2 \\
&\qquad + \sum_{k=1}^{n} \sum_{z \in \mathbb{Z}^d} \mathbb{E}_{(x,y)}\Big[\mathbf{1}(T = k, \bar{X}_k = z)\Big\{\mathbb{E}^{\eta}[\varphi_{\bar{X}_n, \bar{X}'_n}(\eta_0)]\alpha^{\#\{1 \leq i \leq n : \bar{X}_i = \bar{X}'_i\}} \\
&\qquad\qquad + \gamma\theta \sum\nolimits_{i=1}^{n} \mathbf{1}(\bar{X}_i = \bar{X}'_i)\alpha^{\#\{1 \leq j \leq i-1 : \bar{X}_j = \bar{X}'_j\}}\Big\}\Big] \\
&\quad = \sum_{x',y'} \mathbb{P}_{(x,y)}(\bar{X}_n = x', \bar{X}'_n = y', T > n)\big(\mathbb{E}^{\eta}[\varphi_{x',y'}(\eta_0)] - \theta^2\big) - \mathbb{P}_{(x,y)}(T \leq n)\theta^2 \\
&\qquad + \sum_{k=1}^{n} \sum_{z \in \mathbb{Z}^d} \mathbb{P}_{(x,y)}(T = k, \bar{X}_k = \bar{X}'_k = z)\mathbb{E}^{\eta}[\varphi_{z,z}(\eta_{n-k})] \\
&\quad = \sum_{x',y'} \mathbb{P}_{(x,y)}(\bar{X}_n = x', \bar{X}'_n = y', T > n)\mathrm{Cov}(\eta_0(x'), \eta_0(y')) - \mathbb{P}_{(x,y)}(T \leq n)\theta^2 \\
&\qquad + \sum_{k=1}^{n} \mathbb{P}_{(x,y)}(T = k)\mathbb{E}^{\eta}[\varphi_{0,0}(\eta_{n-k})].
\end{aligned}$$

As for fixed n and any bound M it holds that $\mathbb{P}_{(x,y)}(|\bar{X}_n - \bar{X}'_n| \geq M, T > n) \to 1$ as $|x - y| \to \infty$ we see that asymptotic uncorrelatedness is retained over any finite time interval. If $\alpha < \alpha_2$ (which in particular implies that the difference random walk $\bar{X} - \bar{X}'$ is transient) we can furthermore estimate

$$\begin{aligned}
|\mathrm{Cov}(\eta_n(x), \eta_n(y))| \quad \leq \quad & \sum_{x',y'} \mathbb{P}_{(x,y)}(\bar{X}_n = x', \bar{X}'_n = y', T > n)|\mathrm{Cov}(\eta_0(x'), \eta_0(y'))| \\
& + \mathbb{P}_{(x,y)}(T < \infty)\big(\theta^2 + \sup\nolimits_m \mathbb{E}^{\eta}[\varphi_{0,0}(\eta_m)]\big)
\end{aligned}$$

uniformly in n. As $\mathbb{P}_{(x,y)}(T < \infty) \to 0$ and $\mathbb{P}_{(x,y)}(\inf_n |\bar{X}_n - \bar{X}'_n| \geq M) \to 1$ for $|x - y| \to \infty$ by the transience of the difference random walk, the second claim in 2) is proved.

3) η_∞ is shift-invariant as a limit of shift-invariant η_{n_k}. 2) implies uniform integrability of $(\eta_{n_k}(0))_k$, hence $\mathbb{E}\,\eta_\infty(0) = \theta$. It remains to verify that $\mathcal{L}(\eta_\infty)$ satisfies (4.7). Fix $N \in \mathbb{N}$ for the moment. Using Lemma 20 we compute

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\sum_x \bar{p}_N(0,x)\eta_{n_k}(x) - \theta\Big)^2\Big] = \sum_{x,y} \bar{p}_N(0,x)\bar{p}_N(0,y)\mathbb{E}\left[\varphi_{x,y}(\eta_{n_k}) + \theta\delta_{x,y} - \theta^2\right] \\
&\quad = \theta \sum_x (\bar{p}_N(0,x))^2 + \sum_{x,y} \bar{p}_N(0,x)\bar{p}_N(0,y)\mathbb{E}_{(x,y)}\Big[\varphi_{\bar{X}_{n_k}, \bar{X}'_{n_k}}(\eta_0)\alpha^{\#\{1 \leq i \leq n_k : \bar{X}_i = \bar{X}'_i\}} - \theta^2 \\
&\qquad\qquad + \theta\gamma \sum_{i=1}^{n_k} \mathbf{1}(\bar{X}_i = \bar{X}'_i)\alpha^{\#\{1 \leq j \leq i-1 : \bar{X}_j = \bar{X}'_j\}}\Big],
\end{aligned}$$

showing that

$$\begin{aligned}\mathbb{E}\Big(\sum_x \bar p_N(0,x)\eta_{n_k}(x)-\theta\Big)^2 = {} & \theta\tilde p_{2N}(0,0)+\mathbb{E}_{(0,0)}\Big[\varphi_{\bar X_{N+n_k},\bar X'_{N+n_k}}(\eta_0)-\theta^2\Big] \\ & + \mathbb{E}_{(0,0)}\Big[\varphi_{\bar X_{N+n_k},\bar X'_{N+n_k}}(\eta_0)\big(\alpha^{\#\{N+1\le i\le N+n_k:\bar X_i=\bar X'_i\}}-1\big)\Big] \\ & + \mathbb{E}_{(0,0)}\Bigg[\theta\gamma\sum_{i=N+1}^{N+n_k}\mathbf{1}(\bar X_i=\bar X'_i)\alpha^{\#\{N+1\le j\le i-1:\bar X_j=\bar X'_j\}}\Bigg].\end{aligned}$$

Observe that the second term on the righthand side vanishes as $k\to\infty$ because $\mathcal{L}(\eta_0)\in\mathcal{R}_\theta$, observe also that $\sum_x \bar p_N(0,x)\eta_{n_k}(x)\to\sum_x\bar p_N(0,x)\eta_\infty(x)$ in probability. Denoting $B_N:=\#\{i>N:\bar X_i=\bar X'_i\}$ we can thus estimate by Fatou's Lemma

$$\begin{aligned}&\mathbb{E}\Big[\Big(\sum_x \bar p_N(0,x)\eta_\infty(x)-\theta\Big)^2\Big] \\ &\quad\le \liminf_k \mathbb{E}\Big[\Big(\sum_x \bar p_N(0,x)\eta_{n_k}(x)-\theta\Big)^2\Big] \\ &\quad\le \theta\tilde p_{2N}(0,0)+\Big(\sup_{z,z'}\mathbb{E}\,\varphi_{z,z'}(\eta_0)\Big)\mathbb{E}_{(0,0)}\big[\alpha^{B_N}-1\big]+\theta\gamma\mathbb{E}_{(0,0)}\big[B_N\alpha^{B_N}\big].\end{aligned}$$

Using $B_N\le B_0$, the fact that $\alpha<\alpha_2$ implies that $\mathbb{E}_{(0,0)}\alpha^{B_0}$, $E_{(0,0)}B_0\alpha^{B_0}<\infty$ and that $B_N\to 0$ in probability as $N\to\infty$ as well as $\tilde p_{2N}(0,0)\to 0$ by transience of the difference random walk we can conclude by dominated convergence that

$$0\le\limsup_N \mathbb{E}\Big[\Big(\sum_x \bar p_N(0,x)\eta_\infty(x)-\theta\Big)^2\Big]\le 0,$$

which proves our claim.
4) For the case $\alpha\ge\alpha_2$ we observe that

$$\mathbb{E}[\varphi_{x,y}(\eta_0)]\ge C\mathbf{1}(|x-y|\ge M)$$

for suitable $C,M>0$ by the assumed shift invariance and asymptotic uncorrelatedness of η_0. Thus we see from Lemma 20

$$\mathbb{E}\,\eta_n(0)^2\ge C\,\mathbb{E}_{(0,0)}\Big[\mathbf{1}(|\bar X_n-\bar X'_n|\ge M)\alpha^{\#\{1\le i\le n:\bar X_i=\bar X'_i\}}\Big].$$

Transience of the difference random walk gives $\mathbf{1}(|\bar X_n-\bar X'_n|\ge M)\alpha^{\#\{1\le i\le n:\bar X_i=\bar X'_i\}}\to\alpha^{\#\{1\le i:\bar X_i=\bar X'_i\}}$ almost surely as $n\to\infty$ so that we can conclude with Fatou's lemma that

$$\liminf_{n\to\infty}\mathbb{E}\,\eta_n(0)^2\ge C\,\mathbb{E}_{(0,0)}\alpha^{\#\{1\le i:\bar X_i=\bar X'_i\}}=\infty.$$

Finally the results from chapter 5, section 5.4 show that $\mathbb{E}_{(0,0)}\alpha^{\#\{1\le i\le n:\bar X_i=\bar X'_i\}}$ grows exponentially in n if $\alpha>\alpha_2$. □

The following lemma is the analogue of the Feynman-Kac formula in this discrete-time setting. The proof is elementary, e.g. by induction, and we omit it here. Note that in general, one would have to work with the dual of p with respect to an invariant measure. In our situation, where p is doubly stochastic and thus counting measure is invariant, the transpose $\bar p$ coincides with the dual.

Lemma 22 ("discrete Feynman-Kac formula") *Let S be a countable set and p a doubly stochastic transition matrix on S, let $a_n : S \to \mathbb{R}$, $b_n : S \to \mathbb{R}$ for $n = 0, 1, \ldots,$ sequences of bounded functions on S and a bounded $\varphi_0 : S \to \mathbb{R}$ be given. Define (φ_n) recursively by*

$$\varphi_n(x) = \sum_{y \in S} p(y, x)\big(a_{n-1}(y)\varphi_{n-1}(y) + b_{n-1}(y)\big).$$

Then φ_n has a probabilistic representation as the expectation of a functional of a path of the Markov chain $(\bar{X}_n)_n$ with transition matrix $\bar{p}$, where $\bar{p}(x, y) = p(y, x)$:

$$\varphi_n(x) = \mathbb{E}_x \left[\varphi_0(\bar{X}_n) \prod_{i=1}^{n} a_{n-i}(\bar{X}_i) + \sum_{j=1}^{n} b_{n-j}(\bar{X}_j) \prod_{k=1}^{j-1} a_{n-k}(\bar{X}_k) \right].$$

4.4 Local size-biasing

Let the distribution of η under $\hat{\mathbb{P}}$ be size-biased with respect to $\eta_N(x_0)$, the number of particles at position x_0 in generation N, i.e. $\hat{\mathbb{P}}(\eta \in B) = (\mathbb{E}\,\eta_N(x_0))^{-1}\mathbb{E}[\eta_N(x_0)\mathbf{1}_B(\eta)]$. We derive in this section a stochastic representation of η under $\hat{\mathbb{P}}$.

To do this we take up again the genealogical point of view from section 4.1.2 and define a probability measure $\tilde{\mathbb{P}}$ on {initial configurations} $\times$ {space-time fields of offspring distributions} $\times$ {genealogies} $\times$ {tagged ancestral lines}. Here a tagged ancestral line is a ray (of length N) that ends at spatial position x_0 and is part of the genealogy. Alternatively we could speak of a tagged individual at x_0 in generation N. The gist is to construct $\tilde{\mathbb{P}}$ in such a way that the weight does not depend on which particle we tag, i.e. such that for any initial configuration $\underline{n} = (n(x))_{x \in \mathbb{Z}^d}$, space-time configuration of offspring distributions $(q(n, x))$, genealogy $(t_{x,i})$ and tagged ancestral line a (that is part of some t_{y,i^*} and ends at x_0 in generation N) we have with a fixed constant c

$$\text{“}\tilde{\mathbb{P}}(\eta_0 = \underline{n}, Q = q, \tau = t, \alpha = a) = c\,\mathbb{P}(\eta_0 = \underline{n}, Q = q, \tau = t)\text{”}$$

(of course a slight technical difficulty lies in the fact that the above equation actually reads "$0 = 0$" because there are uncountably many possibly configurations). By summing out all possible tagged ancestral lines we see that $\tilde{\mathbb{P}} = \hat{\mathbb{P}}$ as desired (formally, one has to "project away" the ancestral line component).

Here is our list of ingredients: Let Y_N be a $\mathbb{Z}^d$-valued random variable with $\tilde{\mathbb{P}}(Y_N = y) = \mathbb{E}[\eta_0(y)m^N p^N(y, x_0)]/\mathbb{E}[\eta_N(x_0)]$, where $m := \mathbb{E}\,m_1(Q)$. Given $Y_N = y$ let the path $(Y_0, Y_1, \ldots, Y_N)$ be a $\bar{p}$-bridge from x_0 to y.

Given $Y_N = y$ let the initial configuration η_0 be an independent copy of $\hat{\eta}^{(y)}$, where $\hat{\eta}^{(y)}$ has the locally (with the number of particles at y) size-biased distribution of η_0, i.e. $\mathbb{P}(\hat{\eta}^{(y)} \in B) = (\mathbb{E}\eta_0(y))^{-1}\mathbb{E}[\eta_0(y)\mathbf{1}_B(\eta_0)]$.

Let $\hat{Q}(1), \ldots, \hat{Q}(N)$ be independent, distributed according to the law of Q reweighted by its first moment $m_1(Q)$, i.e. $\tilde{\mathbb{P}}(\hat{Q} \in B) = m^{-1}\mathbb{E}[m_1(Q)\mathbf{1}_B(Q)]$. Given the $\hat{Q}(\cdot)$s let $\hat{K}_1, \ldots, \hat{K}_N$ be independent, $\hat{K}_i$ distributed according to the size-biasing of $\hat{Q}(i)$, i.e. $\tilde{\mathbb{P}}(\hat{K}_i = k|\hat{Q}(i)) = k\hat{Q}(i)_k/m_1(\hat{Q}(i))$. Given the $\hat{K}_i$ choose $I_1, \ldots, I_N$ independent, I_i uniformly distributed on $\{1, \ldots, \hat{K}_i\}$. Let I_{N+1} be uniformly distributed on $\{1, 2, \ldots, \eta_0(Y_N)\}$.

For $j < N, z \in \mathbb{Z}^d, z \neq Y_{N-j}$ let $Q(z, j)$ be independent copies of Q, define $Q(Y_j, N - j) := \hat{Q}(j)$, $j = 1, \ldots, N$.

Our system arises from these ingredients as follows: Individual number I_{N+1} at location Y_N in the initial population becomes the founding mother of the tagged ancestral line. This individual has $\hat{K}_N$ offspring, the I_N-th of them continues the tagged line. In general the individual in generation $N-j$ on the tagged ancestral line has $\hat{K}_j$ offspring, and the I_j-th of them continues the special line. The spatial embedding of the tagged ancestral line is $(Y_N,\dots,Y_0)$. Given the space-time field of offspring distributions $Q(\cdot,\cdot)$ let all the siblings branching off the tagged ancestral line take an independent p-step from the position of their respective mother. Then they found independent, spatially embedded branching trees in the already generated space-time medium. All the other individuals in the initial population, which are not related to the tagged line also found, given $Q(\cdot,\cdot)$, independent branching trees. See also the figure on page 72.

Formally we define $\tilde{\mathbb{P}}$ as follows: For a (finite) $A\subset\mathbb{Z}^d$, $n(z)\in\mathbb{N}_0$ for $z\in A$, $t_{z,i}$, $z\in A, i\le n(z)$ (finite, rooted, ordered, spatially embedded) trees of height at most N, an ancestral line a, which is a part of t_{y,i^*}, where $i^*\in\{1,\dots,\eta_0(y)\}$, with spatial embedding $(y=y_N,y_{N-1},\dots,y_0=x_0)$, $\mathbb{Z}^d\supset B\supset A$ with $|B|<\infty$ and $\mathrm{pos}(\sigma)\in B$ for all $\sigma\in t_{z,i}$, $z\in A, i\le n(z)$ and (measurable) $C_{w,j}\subset\mathcal{M}_1(\mathbb{N}_0)$, $w\in B$, $j<N$, we set

$$\begin{aligned}
&\tilde{\mathbb{P}}\Big(\{\eta_0(z)=n(z),z\in A\}\cap\{Q(w,j)\in C_{w,j},w\in B,0\le j<N\}\\
&\qquad\qquad\qquad\cap\{\tau_{z,i}|_N=t_{z,i},i\le n(x),z\in A\}\cap\{\alpha=a\}\Big)\\
&\quad=\frac{\mathbb{E}[\eta_0(y)]m^Np^N(y,x)}{\mathbb{E}[\eta_N(x)]}\times\left(\frac{1}{\bar{p}^N(x,y)}\prod_{j=1}^N\bar{p}(y_{j-1},y_j)\right)\\
&\quad\times\frac{1}{\mathbb{E}[\eta_0(y)]}n(y)\mathbb{P}(\eta_0(z)=n(z),z\in A)\\
&\quad\times\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\cdots\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\prod_{k=1}^N\mu(dq(y_k,N-k))\frac{m_1(q(y_k,N-k))}{m}\mathbf{1}_{C_{y_k,N-k}}(q(y_k,N-k))\\
&\qquad\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\cdots\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\prod_{\substack{z\in B,j<N\\ z\ne y_{N-j}}}\mu(dq(z,j))\mathbf{1}_{C_{z,j}}(q(z,j))\Bigg\{\\
&\qquad\prod_{\ell=1}^N\frac{1}{m_1(q(y_\ell,N-\ell))}|\mathrm{offspring}(a_{N-\ell})|\times q_{|\mathrm{offspring}(a_{N-\ell})|}(y_\ell,N-\ell)\\
&\qquad\times\prod_{m=1}^N\frac{1}{|\mathrm{offspring}(a_{N-m})|}\times\frac{1}{n(y)}\times\\
&\qquad\times\prod_{\sigma\in t_{y,i^*},\sigma\notin a}q_{|\mathrm{offspring}(\sigma)|}(\mathrm{pos}(\sigma),|\sigma|)\Bigg(\prod_{\sigma'\in\mathrm{offspring}(\sigma)}p(\mathrm{pos}(\sigma),\mathrm{pos}(\sigma'))\Bigg)\\
&\qquad\times\prod_{\substack{w\in A,n\le n(w)\\(w,n)\ne(y,i*)}}\prod_{\rho\in t_{w,n}}q_{|\mathrm{offspring}(\rho)|}(\mathrm{pos}(\rho),|\rho|)\Bigg(\prod_{\rho'\in\mathrm{offspring}(\rho)}p(\mathrm{pos}(\rho),\mathrm{pos}(\rho'))\Bigg)\Bigg\}\\
&\quad=\frac{1}{\mathbb{E}[\eta_N(x)]}\mathbb{P}(\eta_0(z)=n(z),z\in A)\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\cdots\int\limits_{\mathcal{M}_1(\mathbb{N}_0)}\prod_{z\in B,j<N}\mu(dq(z,j))\mathbf{1}_{C_{z,j}}(q(z,j))\Bigg\{\\
&\qquad\qquad\prod_{w\in A,i<n(w)}\prod_{\rho\in t_{w,i}}q_{|\mathrm{offspring}(\rho)|}(\mathrm{pos}(\rho),|\rho|)\Bigg(\prod_{\rho'\in\mathrm{offspring}(\rho)}p(\mathrm{pos}(\rho),\mathrm{pos}(\rho'))\Bigg)\Bigg\}.
\end{aligned}$$

Our claim follows because the probabilities of these events determine $\tilde{\mathbb{P}}$.

Remark 18 1) This genealogical construction has of course its own genealogy: Among its ancestors are at least Olav Kallenberg's pioneering [19], Dawson and Fleischmann's [6], [7], Lyons, Pemantle, Peres' [25], and work of Jochen Geiger, see e.g. [11], [12].

2) "One – two – size-bias: a doubly stochastic construction": Observe that the distribution of the number of children along the tagged ancestral line is the size-biasing of the original offspring distribution,

$$\begin{aligned}\mathbb{P}(\hat{K}=k) &= \mathbb{E}\left[\frac{k}{m_1(\hat{Q})}\hat{Q}_k\right] = \frac{1}{\mathbb{E}[m_1(Q)]}\mathbb{E}\left[\frac{k}{m_1(Q)}\times m_1(Q)Q_k\right] \\ &= \frac{1}{\mathbb{E}[m_1(Q)]}\mathbb{E}\left[kQ_k\right] = \frac{k}{\mathbb{E}\,K}\mathbb{P}(K=k).\end{aligned}$$

This is analogous to the "classical" (non-interacting) constructions mentioned above.

There is a simple observation behind this: Consider a doubly stochastic scenario where one first chooses a random distribution Q and then, given Q, a positive random variable K with this distribution. Then the size-biased law $\mathcal{L}(\hat{K})$ can also be constructed in a doubly stochastic way. First choose $\hat{Q}$ according to the law of Q reweighted by its first moment $\sum_k kQ_k$, then $\hat{K}$ according to a size-biased $\hat{Q}$.

4.4.1 ... and a condition for uniform integrability

Let us consider a shift-invariant scenario with $\mathbb{E}\,\eta_0(x)\equiv\theta$ and $\mathbb{E}\,\eta_0(x)^2<\infty$ under assumption (4.3). Then we have in particular that $\mathbb{E}\,\eta_n(x)\equiv\theta$. The question whether any limit point η_∞ also satisfies $\mathbb{E}\,\eta_\infty(x)\equiv\theta$ is equivalent to the question whether the family $(\eta_n(0))_{n\in\mathbb{N}}$ is uniformly integrable. This, in turn, is by Lemma 9 equivalent to tightness of the size-biased family $(\hat{\eta}_n(0))_{n\in\mathbb{N}}$.

We can use the above construction of $\mathcal{L}(\hat{\eta}_N(0))$ to find a (sufficient) criterion by averaging out the field of random offspring distributions *conditional on the spatial embedding* $Y=(Y_0,Y_1,\ldots)$ *of the tagged line*: Obviously $(\eta_n(0))_{n\in\mathbb{N}}$ is tight under $\tilde{\mathbb{P}}$ if even $\sup_n\tilde{\mathbb{E}}[\eta_n(0)|Y]<\infty$ holds.

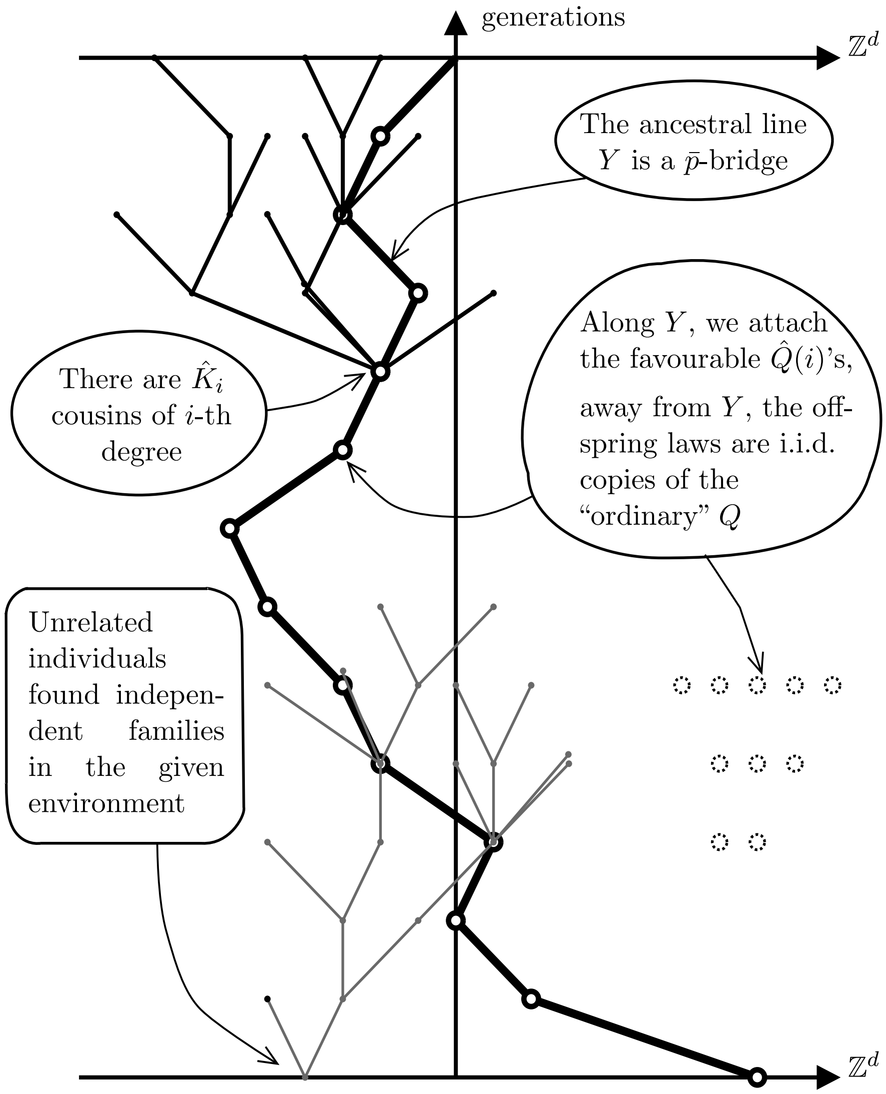
Proposition 9 *Let η_0 be shift-invariant with $\mathbb{E}\,\eta_0(x)\equiv\theta$ and $\mathbb{E}\,\eta_0(x)^2<\infty$. Assume (4.3) and $\alpha=\mathbb{E}[(m_1(Q))^2]<\alpha_*$, where α_* is defined in (4.6). Then any limit point η_∞ satisfies $\mathbb{E}\,\eta_\infty(x)=\theta$.*

Proof. Let us first consider a family $(\eta_k^{x,n})_{k\geq n}$ founded by one single ancestor at position x in generation $n\leq N$. Then conditioned on the environment $Q(\cdot,\cdot)$ we have

$$\mathbb{E}[\eta_N^{x,n}(x')|Q(\cdot,\cdot)] = \sum_{\substack{x_{n+1},\ldots,x_{N-1}\in\mathbb{Z}^d \\ x_n=x,\,x_N=x'}}\prod_{j=n}^{N-1}p(x_j,x_{j+1})m_1(Q(x_j,j)),$$

as can be proved analogously to Lemma 20, 1).

By shift invariance it is sufficient to consider $\tilde{\mathbb{P}}(\eta_N(0)\in\cdot)$. The assumption $\mathbb{E}\eta_0(x)\equiv\theta$ implies that $Y=(Y_0,\ldots,Y_N)$ is a (unconditional) $\bar{p}$-Markov chain starting from $Y_0=0$ under $\tilde{\mathbb{P}}$.



A schematic representation of the construction of η under $\tilde{\mathbb{P}}$

Let us denote by $(\eta_n^{(j)})_{n=N-j,\dots,N}$ the subpopulation of individuals arising in the construction of $\tilde{\mathbb{P}}$ whose most recent common ancestor with the tagged particle lived j generations before (the present) time N. These are all side trees that branch off the tagged line in generation $N-j$. As above we find

$$\tilde{\mathbb{E}}[\eta_N^{(j)}(0)|Q(\cdot,\cdot),(\hat{K}_i)_{i=1,\dots,N},Y] = \\ (\hat{K}_j - 1) \sum_{\substack{x_{N-j+1},\dots,x_{N-1} \\ x_N=0}} p(Y_j, x_{N-j+1}) \prod_{i=N-j+1}^{N-1} p(x_i, x_{i+1}) m_1(Q(x_i, i)).$$

Averaging out $Q(\cdot,\cdot)$ and the $\hat{K}_i$s this yields

$$\begin{aligned}
&\tilde{\mathbb{E}}[\eta_N^{(j)}(0)|Y] \\
&\quad = \mathbb{E}[\hat{K}-1] \sum_{\substack{x_{N-j+1},\dots,x_{N-1} \\ x_N=0}} p(Y_j, x_{N-j+1}) \prod_{i=N-j+1}^{N-1} p(x_i, x_{i+1}) \Big\{ \mathbb{E}[m_1(Q)]\mathbf{1}(x_i \neq Y_{N-i}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \mathbb{E}[m_1(\hat{Q})]\mathbf{1}(x_i = Y_{N-i}) \Big\} \\
&\quad = \mathbb{E}[\hat{K}-1]\mathbb{E}_{Y_j}^X \left[\mathbf{1}(X_j = 0)\alpha^{\#\{1\le i<j : X_i = Y_{j-i}\}} \,\middle|\, Y \right] \\
&\quad = \mathbb{E}[\hat{K}-1]\mathbb{E}_0^{\bar{X}} \left[\mathbf{1}(\bar{X}_j = Y_j)\alpha^{\#\{1\le i<j : \bar{X}_i = Y_i\}} \,\middle|\, Y \right],
\end{aligned}$$

where X is a p- and $\bar{X}$ an independent $\bar{p}$-random walk, both independent of all the other ingredients, and we use $\mathbb{E}[m_1(\hat{Q})] = \mathbb{E}[(m_1(Q))^2]/1 = \alpha$. Observe that $\mathbb{E}\,\hat{K} = (\mathbb{E}[m_1(Q)])^{-1}\sum_k k^2 \mathbb{E}\,Q_k = \mathbb{E}[m_2(Q)]/\mathbb{E}[m_1(Q)] < \infty$. $\mathbb{E}_z^X$ refers to expectation under the measure with $X_0 = z$, and $\mathbb{E}_0^{\bar{X}}$ refers to $\bar{X}_0 = 0$.

Let $(\eta_n^{(-)})_n$ be the subpopulation founded by individuals not related to the tagged ancestral line. We have

$$\begin{aligned}
\tilde{\mathbb{E}}[\eta_N^{(-)}(0)|Y,\eta_0] &= \sum_z (\eta_0(z) - \mathbf{1}(z = Y_N))\mathbb{E}_z^X \left[\mathbf{1}(X_N = 0)\alpha^{\#\{0\le i<N : X_i = Y_{N-i}\}} \,\middle|\, Y \right] \\
&= \mathbb{E}_0^{\bar{X}} \left[(\eta_0(\bar{X}_N) - \mathbf{1}(\bar{X}_N = Y_N))\alpha^{\#\{1\le i\le N : \bar{X}_i = Y_i\}} \,\middle|\, Y, \eta_0 \right].
\end{aligned}$$

Using

$$\tilde{\mathbb{E}}[\eta_0(z)|Y] = \frac{1}{\theta}\mathbb{E}[\eta_0(z)\eta_0(Y_N)|Y] \le \frac{1}{\theta}\mathbb{E}[\eta_0(0)^2] \quad (<\infty)$$

we estimate

$$\tilde{\mathbb{E}}[\eta_N^{(-)}(0)|Y] \le \frac{1}{\theta}\mathbb{E}[\eta_0(0)^2]\mathbb{E}_0^{\bar{X}} \left[\alpha^{\#\{1\le i\le N : \bar{X}_i = Y_i\}} \,\middle|\, Y \right].$$

Putting $B := \#\{i \ge 1 : \bar{X}_i = Y_i\}$ we can thus estimate altogether

$$\begin{aligned}
\tilde{\mathbb{E}}[\eta_N(0)|Y] &\le 1 + \mathbb{E}[\hat{K}-1]\sum_{j=1}^N \mathbb{E}_0^{\bar{X}} \left[\mathbf{1}(\bar{X}_j = Y_j)\alpha^{\#\{1\le i<j : \bar{X}_i = Y_i\}} \,\middle|\, Y \right] \\
&\qquad + \mathbb{E}_0^{\bar{X}} \left[\alpha^{\#\{1\le i\le N : \bar{X}_i = Y_i\}} \,\middle|\, Y \right] \\
&\le 1 + \text{const.} \times \mathbb{E}_0^{\bar{X}}[(B+1)\alpha^B \,|\, Y]
\end{aligned}$$

uniformly in N. Observe that the righthand side is almost surely finite if $\alpha < \alpha_*$. $\square$

4.5 Limits of Poisson systems

Similar to the classical case, Poisson initial conditions are particularly suitable for relatively explicit calculations. A complication stems from the fact that the system is only infinitely divisible if we condition on the medium $Q(\cdot,\cdot)$.

Lemma 23 *Let $\eta_0(x), x \in \mathbb{Z}^d$ be independent, Poisson(θ)-distributed. η_n converges in distribution to a shift invariant, E_1-valued η_∞. $\mathcal{L}(\eta_\infty)$ is invariant under the time evolution semigroup (S_n).*

Proof. Set $\mathcal{U}(m,n) := \sigma\left(Q(x,i) : x \in \mathbb{Z}^d, m \le i < n\right)$, $\mathcal{U} := \mathcal{U}(0,\infty)$; for $q \in \mathcal{M}_1(\mathbb{N}_0)$ let

$$\varphi_q(s) := \sum_k s^k q_k, \quad s \in [0,1]$$

be the generating function of the number of children of an individual with offspring distribution q, $\bar\varphi(s) := \mathbb{E}\varphi_Q(s) = \sum_{k\ge 0} s^k \mathbb{E}\, Q_k$ the (absolute) generating function of the number of offspring of an individual in our population. Observe that $\bar\varphi$ is convex and monotone increasing and satisfies $\bar\varphi'(1) = \mathbb{E}[m_1(Q)] = 1$ by assumption (4.3). This implies that $1 - \bar\varphi(s) \le 1 - s$ for $s \in [0,1]$.

For (fixed) $f : \mathbb{Z}^d \to \mathbb{R}_+$ and $m \le n$ define

$$w_{m,n}(x) := \mathbb{E}_{\eta_m = \delta_x}\left[\exp(-\langle \eta_n, f\rangle) \,|\, \mathcal{U}\right].$$

Apparently, $w_{m,n}$ is $\mathcal{U}(m,n)$-measurable. Decompose according to the m-th generation to find the following recursion for $m < n$:

$$\begin{aligned} w_{m,n}(x) &= \sum_k Q_k(m,x) \sum_{y_1,\dots,y_k} \prod_{i=1}^k \Big\{ p(x,y_i)\mathbb{E}_{\eta_{m+1}=\delta_{y_i}}\big[\exp(-\langle \eta_n, f\rangle) \,|\, \mathcal{U}(m+1,n)\big]\Big\} \\ &= \sum_k Q_k(m,x) \left(\sum_y p(x,y)\mathbb{E}_{\eta_{m+1}=\delta_y}\big[\exp(-\langle \eta_n, f\rangle) \,|\, \mathcal{U}(m+1,n)\big]\right)^k \\ &= \varphi_{Q(m,x)}\Big(\textstyle\sum_y p(x,y) w_{m+1,n}(y)\Big) \end{aligned}$$

For $\mathcal{L}(\eta_0) = \mathrm{Poi}(\underline{\theta})$ we can use this to compute (recall that for $Z \sim \mathrm{Poi}(\theta)$ we have $\mathbb{E}\exp(-\alpha Z) = \exp(-\theta(1-e^{-\alpha})))$

$$\begin{aligned} \ell_n(f) &:= \mathbb{E}\exp(-\langle \eta_n, f\rangle) = \mathbb{E}\left[\mathbb{E}[\exp(-\langle \eta_n, f\rangle) \,|\, \mathcal{U}]\right] \\ &= \mathbb{E}\Big[\textstyle\prod_x w_{0,n}(x)^{\eta_0(x)}\Big] = \mathbb{E}\exp\left(-\theta \displaystyle\sum_x (1 - w_{0,n}(x))\right), \end{aligned}$$

consequently (by the independence of the Qs and Jensen's inequality)

$$\begin{aligned} \ell_n(f) &\ge \mathbb{E}^{\mathcal{U}(1,n)} \exp\Big(-\mathbb{E}^{\mathcal{U}(0,0)}\left[\theta \textstyle\sum_x (1 - w_{0,n}(x))\right]\Big) \\ &= \mathbb{E}^{\mathcal{U}(1,n)} \exp\Big(-\theta\mathbb{E}^{\mathcal{U}(0,0)}\Big[\textstyle\sum_x \Big\{1 - \varphi_{Q(0,x)}\Big(\textstyle\sum_y p(x,y) w_{1,n}(y)\Big)\Big\}\Big]\Big) \\ &= \mathbb{E}^{\mathcal{U}(1,n)} \exp\Big(-\theta \textstyle\sum_x \Big\{1 - \bar\varphi\Big(\textstyle\sum_y p(x,y) w_{1,n}(y)\Big)\Big\}\Big) \\ &\ge \mathbb{E}^{\mathcal{U}(1,n)} \exp\Big(-\theta \textstyle\sum_{x,y} p(x,y)(1 - w_{1,n}(y))\Big) \\ &= \mathbb{E}^{\mathcal{U}(1,n)} \exp\Big(-\theta \textstyle\sum_y (1 - w_{1,n}(y))\Big) = \ell_{n-1}(f). \end{aligned}$$

This shows that the sequence of Laplace transforms of $\mathcal{L}(\eta_n)$ converges for any (positive) test function f, which implies convergence in (finite-dimensional) distribution(s), see [20].

Obviously, as a limit of shift invariant η_n, η_∞ is shift invariant too. By Fatou's lemma we have $\mathbb{E}\eta_\infty(x) \le \theta$, thus $\mathbb{P}(\eta_\infty \in E_1) = 1$. Using Remark 16 we see that $\mathcal{L}(\eta_\infty)S_1 = (\lim_{n\to\infty}\mathrm{Poi}(\underline{\theta})S_n)S_1 = \lim_{n\to\infty}\mathrm{Poi}(\underline{\theta})S_{n+1} = \mathcal{L}(\eta_\infty)$, proving that $\mathcal{L}(\eta_\infty)$ is invariant under time evolution. □

Combining this with Proposition 9 and Lemma 21 we find

Theorem 4 *Assume (4.3) and $\alpha < \alpha_*$.*
1) Then for any $\theta \ge 0$ we have $\mathrm{Poi}(\underline{\theta})S_n \Rightarrow \nu_\theta$, where ν_θ is invariant and has full intensity measure $\int \eta(0)\nu_\theta(d\eta) = \theta$.

2) For $\alpha \in [\alpha_2, \alpha_*)$ we have $\displaystyle\int \eta(0)^2 \nu_\theta(d\eta) = \infty,$

3) for $\alpha < \alpha_2$, $\displaystyle\int \eta(0)^2 \nu_\theta(d\eta) < \infty$ *and* $\mathcal{L}(\nu_\theta) \in \mathcal{R}_\theta$.

4.6 Order and positive correlations

Recall that the probability measures $\mathcal{M}_1(M)$ on a partially ordered set $(M, \le)$ are also partially ordered via $\mu \le \nu$ ("μ is stochastically smaller than ν") iff $\int f\,d\mu \le \int f\,d\nu$ for all bounded, monotone increasing functions $f : M \to \mathbb{R}_+$. A probability measure μ on M is said to have *positive correlations* if $\int fg\,d\mu \ge \int f\,d\mu \int g\,d\mu$ for all bounded, monotone functions f and g. See [23], chapter II.2 for more on this.

Equip $E_1 \subset (\mathbb{N}_0)^{\mathbb{Z}^d}$ with its natural componentwise partial ordering. We assume in this section that

$$\begin{array}{c}\text{the random offspring distribution } Q \text{ takes values in}\\ \text{a \textit{linearly ordered} subset } \tilde{\mathcal{M}} \subset \mathcal{M}_1(\mathbb{N}_0).\end{array} \tag{4.8}$$

Observe that (4.8) is automatically satisfied in the "random thinning" case considered in Remark 15, 2).

I am confident that the following observation is well known. For lack of reference, and since the proof is short I present it here:

Lemma 24 *Let $\eta = \sum_i \delta_{x_i}$ be a Poisson point process on a Polish space M with σ-finite intensity measure μ. Then $\mathcal{L}(\eta)$ has positive correlations.*

Proof. Consider a sequence $\mathcal{P}_n$, $n \in \mathbb{N}$ of finite families of subsets of M, $\mathcal{P}_n = \{A_{1,n}, \dots, A_{m_n,n}\}$ with $A_{i,n} \cap A_{j,n} = \emptyset$ for $i \ne j$ and $\mu(A_{i,n}) < \infty$ for all i, n. We assume that $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$, that is for all $B \in \mathcal{P}_{n+1}$ there exists an $A \in \mathcal{P}_n$ such that $B \subset A$, and each $A \in \mathcal{P}_n$ is the union of some $B_k \in \mathcal{P}_{n+1}$. Set $\mathcal{F}_n := \sigma(\eta(A), A \in \mathcal{P}_n)$, then $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. We can choose the sequence in such a way that $\sigma(\cup_n \mathcal{F}_n) = \sigma(\eta)$. Observe that then $F(\eta) = \lim_n \mathbb{E}[F(\eta)|\mathcal{F}_n]$ by the martingale convergence theorem for any bounded measurable $F : \mathcal{N}(M) \to \mathbb{R}$ (where $\mathcal{N}(M)$ denotes the counting measures on M). Now consider $F, G : \mathcal{N}(M) \to \mathbb{R}$ bounded, measurable, monotone increasing. Recall that the conditional distribution of η, given $\eta(A_{i,n}) = k_{i,n}$, $i = 1, 2, \dots, m_n$ arises as follows: For each i, η has $k_{i,n}$ atoms in $A_{i,n}$ whose positions are independent with distribution $\mu(\cdot \cap A_{i,n})/\mu(A_{i,n})$, while the restriction to $\bigcap_i A_{i,n}^c$ is conditionally distributed like a Poisson point process with intensity measure $\mu(\cdot \cap \bigcap_i A_{i,n}^c)$, independent of $\eta(\cdot \cap \bigcup_i A_{i,n})$. Thus

$$\mathbb{E}\Big[\begin{Bmatrix} F \\ G \end{Bmatrix}(\eta)\,\big|\mathcal{F}_n\Big] = \begin{Bmatrix} f_n \\ g_n \end{Bmatrix}\big(\eta(A_{1,n}),\ldots,\eta(A_{m_n,n})\big),$$

for some bounded monotone functions $f_n, g_n : (\mathbb{N}_0)^{m_n} \to \mathbb{R}$. Thus we can compute using dominated convergence

$$\begin{aligned}
\mathbb{E}\left[F(\eta)G(\eta)\right] &= \lim_n \mathbb{E}\big[\mathbb{E}[F(\eta)|\mathcal{F}_n]G(\eta)\big] = \lim_n \mathbb{E}\big[\mathbb{E}[F(\eta)|\mathcal{F}_n]\mathbb{E}[G(\eta)|\mathcal{F}_n]\big] \\
&= \lim_n \mathbb{E}\left[f_n\big(\eta(A_{1,n}),\ldots,\eta(A_{m_n,n})\big)g_n\big(\eta(A_{1,n}),\ldots,\eta(A_{m_n,n})\big)\right] \\
&\geq \lim_n \mathbb{E}\left[f_n\big(\eta(A_{1,n}),\ldots,\eta(A_{m_n,n})\big)\right]\mathbb{E}\left[g_n\big(\eta(A_{1,n}),\ldots,\eta(A_{m_n,n})\big)\right] \\
&= \mathbb{E}[F(\eta)]\mathbb{E}[G(\eta)]
\end{aligned}$$

where the inequality comes from the observation that a Poisson process on a finite set (is a product measure on a product of linearly ordered spaces and thus) has positive correlations, see e.g. [23], p. 78. □

Proposition 10 *Let $\eta_0(x)$, $x \in \mathbb{Z}^d$ be independent, Poisson(θ) and assume that (4.8) holds. Then $\mathcal{L}(\eta_n)$ has positive correlations for all n, in particular ν_θ has positive correlations.*

Proof. This is more or less "general nonsense". Let $f, g : E_1 \to \mathbb{R}_+$ be bounded and monotone increasing. We first observe that for any n

$$\mathcal{L}(\eta_n|Q(\cdot,\cdot)) \quad \text{has positive correlations for (almost) all } Q(\cdot,\cdot)\ :$$

Conditional on the offspring laws $Q(\cdot,\cdot)$ the particles branch and move independently, thus the Poisson initial configuration leads to a Poisson field of "families" of individuals alive in generation n, that is a Poisson process on $\mathcal{N}_{finite}(\mathbb{Z}^d)$. This process has positive correlations by Lemma 24. Furthermore, $F : \mathcal{N}(\mathcal{N}_{finite}(\mathbb{Z}^d)) \to R$, $\sum_i \delta_{\eta^{(i)}} \mapsto f(\sum_i \eta^{(i)})$ is monotone increasing. Using this we compute

$$\begin{aligned}
\mathbb{E}[f(\eta_n)g(\eta_n)] &= \mathbb{E}\big[\mathbb{E}\left[f(\eta_n)g(\eta_n)|Q(\cdot,\cdot)\right]\big] \\
&\geq \mathbb{E}\big[\mathbb{E}\left[f(\eta_n)|Q(\cdot,\cdot)\right]\mathbb{E}\left[g(\eta_n)|Q(\cdot,\cdot)\right]\big].
\end{aligned}$$

The function $\tilde{M}^{\mathbb{Z}^d\times\mathbb{N}_0} \ni q(\cdot,\cdot) \mapsto \mathbb{E}\left[f(\eta_n)|Q(\cdot,\cdot) = q(\cdot,\cdot)\right] \in \mathbb{R}$ is monotone increasing and $\mathcal{L}(Q)$ has positive correlations as a product measure on a product of linearly ordered spaces (see [23], chapter II.2). Hence

$$\begin{aligned}
\mathbb{E}[f(\eta_n)g(\eta_n)] &\geq \mathbb{E}\big[\mathbb{E}\left[f(\eta_n)|Q(\cdot,\cdot)\right]\big]\mathbb{E}\big[\mathbb{E}\left[g(\eta_n)|Q(\cdot,\cdot)\right]\big] \\
&= \mathbb{E}[f(\eta_n)]\mathbb{E}[g(\eta_n)].
\end{aligned}$$

□

Remark 19 1) As η is a system of branching particles the positive correlations are not surprising. For an intuitive explanation consider F and G that are some growing function of the number of particles that η_n has in $A_F \subset \mathbb{Z}^d$ respectively in $A_G \subset \mathbb{Z}^d$. Then an atypically large value of F is likely to be caused by one or more atypically large families. Given this it is also likely that some members of these large families have moved to A_G, causing G to be large, too.

2) It is well possible that ν_θ will always have positive correlations even if (4.8) fails. Alas it is unclear how to prove this.

Lemma 25 *Assume $\alpha < \alpha_2$ and (4.8). There is a $C < \infty$ such that*

$$0 \le \mathrm{Cov}_{\nu_\theta}(\eta(x), \eta(y)) \le C \times \mathbb{P}_{(x,y)}(\bar{X}_i = \bar{X}'_i \text{ for some } i > 0).$$

Proof. As ν_θ has positive correlations we have $\int \eta(x)\eta(y)\,\nu_\theta(d\eta) \ge \theta^2$. On the other hand, Fatou's lemma together with Lemma 20 shows (we assume without loss of generality that $x \neq y$)

$$\begin{aligned}\int \eta(x)\eta(y)\,\nu_\theta(d\eta) &\le \liminf_n \mathbb{E}_{\mathrm{Poi}(\theta)}[\varphi_{x,y}(\eta_n)] \\ &\le \theta^2 \mathbb{P}_{(x,y)}(\bar{X}_i \neq \bar{X}'_i \text{ for all } i > 0) + C' \mathbb{P}_{(x,y)}(\bar{X}_i = \bar{X}'_i \text{ for some } i > 0)\end{aligned}$$

for some $C' < \infty$. Thus

$$\begin{aligned}\mathrm{Cov}_{\nu_\theta}(\eta(x), \eta(y)) &= \int \eta(x)\eta(y)\,\nu_\theta(d\eta) - \theta^2 \\ &\le (\theta^2 + C')\mathbb{P}_{(x,y)}(\bar{X}_i = \bar{X}'_i \text{ for some } i > 0).\end{aligned}$$

□

Corollary 2 *Assume $\alpha < \alpha_2$ and (4.8). Then ν_θ is spatially mixing for any $\theta \ge 0$.*

Proof. We proceed as in [5], p. 844. Fix a finite $A \subset \mathbb{Z}^d$ and put

$$Y_0 := \sum_{x \in A} c_x \eta(x), \quad Y_z := \sum_{x \in A} d_x \eta(x + z),$$

where for $x \in A$, c_x, d_x are positive constants. Y_0 and Y_z are positively correlated by Proposition 10, thus by Thm. 1 of [26] we have for $s, t \in \mathbb{R}$

$$\left| \mathbb{E}_{\nu_\theta}\left[e^{i(sY_0 + tY_z)}\right] - \mathbb{E}_{\nu_\theta}\left[e^{isY_0}\right]\mathbb{E}_{\nu_\theta}\left[e^{itY_z}\right] \right| \le |st| \mathrm{Cov}_{\nu_\theta}(Y_0, Y_z).$$

Now by Lemma 25 we have

$$\begin{aligned}\mathrm{Cov}_{\nu_\theta}(Y_0, Y_z) &= \sum_{x,y \in A} \mathrm{Cov}_{\nu_\theta}(\eta(x), \eta(y + z)) \\ &\le C \sum_{x,y \in A} c_x d_y \mathbb{P}_{(x,y+z)}(\bar{X}_i = \bar{X}'_i \text{ for some } i > 0),\end{aligned}$$

which converges to 0 as $|z| \to \infty$ by transience of the difference random walk. This proves that Y_0 and Y_z become asymptotically independent, or in other words that the joint law of $\langle f, \eta \rangle$ and $\langle f, \eta(\cdot + z) \rangle$ under ν_θ converges to the corresponding product law as $|z| \to \infty$ for any two test functions $f, g : \mathbb{Z}^d \to \mathbb{R}_+$ with support contained in A. Thus ν_θ is mixing. □

4.7 Coupling and ergodic theory in the regime of globally finite second moments

In order to compare two systems $\eta_0^{(1)}$ and $\eta_0^{(2)}$ of branching random walks in random environment starting from different initial conditions it is useful to couple them together in such a way that each of $(\eta_n^{(i)})_n$, $i = 1, 2$ for itself is a system of branching random walks in random environment as above, but also both systems evolve as closely together as possible. To formalize this idea we use

$Q(n,x), x \in \mathbb{Z}^d, n \in \mathbb{N}_0$ ....independent copies of Q
$K_n^{(c)}(y,i)$, $K_n^{(1)}(y,i)$, $K_n^{(2)}(y,i)$ .... given $Q(\cdot,\cdot)$ independent with $\mathbb{P}(K_n^{(\cdot)}(y,\cdot)=k|Q(\cdot,\cdot)) = Q_k(y,n)$,
$Y_n^{(c)}(y,i,j)$, $Y_n^{(1)}(y,i,j)$, $Y_n^{(2)}(y,i,j)$ .... independent, $\mathbb{P}(Y_\cdot^{(\cdot)}(y,\cdot,\cdot)=x) = p(y,x)$.

The idea is that objects with index $^{(c)}$ refer to particles that are common to both systems, while objects with index $^{(1)}$ respectively $^{(2)}$ will be used for "overshoot" particles in system $\eta^{(1)}$ resp. $\eta^{(2)}$.

Then, given the configurations $(\eta_n^{(1)}, \eta_n^{(2)})$, the $(n+1)$-th generation arises as follows: Both systems use the same field of offspring distributions Q, and $\eta_n^{(1)} \wedge \eta_n^{(2)}$ particles behave in exactly the same way in both systems. Possible overshoot particles in one of the two systems branch and move independently, given the Qs, and their offspring are then simply added to those of the common particles. Formally we put

$$\begin{aligned}
\xi_{n+1}^{(c)}(x) &= \sum_{y\in\mathbb{Z}^d} \sum_{i=1}^{\eta_n^{(1)}(y)\wedge\eta_n^{(2)}(y)} \sum_{j=1}^{K_n^{(c)}(y,i)} \mathbf{1}(Y_n^{(c)}(y,i,j)=x), \\
\xi_{n+1}^{(1)}(x) &= \sum_{y\in\mathbb{Z}^d} \sum_{i=1}^{\left(\eta_n^{(1)}(y)-\eta_n^{(2)}(y)\right)^+} \sum_{j=1}^{K_n^{(1)}(y,i)} \mathbf{1}(Y_n^{(1)}(y,i,j)=x), \\
\xi_{n+1}^{(2)}(x) &= \sum_{y\in\mathbb{Z}^d} \sum_{i=1}^{\left(\eta_n^{(2)}(y)-\eta_n^{(1)}(y)\right)^+} \sum_{j=1}^{K_n^{(2)}(y,i)} \mathbf{1}(Y_n^{(2)}(y,i,j)=x)
\end{aligned}$$

and set

$$\eta_{n+1}^{(1)}(x) := \xi_{n+1}^{(c)}(x) + \xi_{n+1}^{(1)}(x), \quad \eta_{n+1}^{(2)}(x) := \xi_{n+1}^{(c)}(x) + \xi_{n+1}^{(2)}(x).$$

Observe that our model is *attractive* in interacting-particle-systems parlance: Starting from $\eta_0^{(1)} \geq \eta_0^{(2)}$ the coupling gives $\eta_n^{(1)} \geq \eta_n^{(2)}$ for all n.

Lemma 26 *Assume (4.3). Let μ and ν be two shift invariant probability measures on E_1 with finite intensities, i.e. $\int \eta(0)\mu(d\eta), \int \eta(0)\nu(d\eta) < \infty$. Then the coupled system, starting from $\mu \otimes \nu$, satisfies*

$$\mathbb{E}|\eta_{n+1}^{(1)}(x) - \eta_{n+1}^{(2)}(x)| \leq \mathbb{E}|\eta_n^{(1)}(x) - \eta_n^{(2)}(x)| \quad \textit{for all } x \in \mathbb{Z}^d, n \geq 0, \textit{ and} \tag{4.9}$$

$$\lim_n \mathbb{P}\left(\eta_n^{(1)}(x) > \eta_n^{(2)}(x), \eta_n^{(1)}(y) < \eta_n^{(2)}(y)\right) = 0 \quad \textit{for all } x, y \in \mathbb{Z}^d. \tag{4.10}$$

Proof. Obviously we have

$$|\eta_{n+1}^{(1)}(x) - \eta_{n+1}^{(2)}(x)| = |\xi_{n+1}^{(1)}(x) - \xi_{n+1}^{(2)}(x)| \leq \xi_{n+1}^{(1)}(x) + \xi_{n+1}^{(2)}(x).$$

Furthermore for $y \in \mathbb{Z}^d$

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{\left(\eta_n^{(1)}(y)-\eta_n^{(2)}(y)\right)^+} \sum_{j=1}^{K_n^{(1)}(y,i)} \mathbf{1}(Y_n^{(1)}(y,i,j)=x)\right] \\
&\quad = \mathbb{E}[\mathbb{E}[\ldots|(\eta_n^{(1)}, \eta_n^{(2)})]] = \mathbb{E}[m_1(Q)]p(y,x)\mathbb{E}\left(\eta_n^{(1)}(y) - \eta_n^{(2)}(y)\right)^+.
\end{aligned}$$

Summing over y and using shift-invariance this gives $\mathbb{E}\,\xi^{(1)}_{n+1}(x) = \mathbb{E}\left(\eta^{(1)}_n(y) - \eta^{(2)}_n(y)\right)^+$, and analogously $\mathbb{E}\,\xi^{(2)}_{n+1}(x) = \mathbb{E}\left(\eta^{(2)}_n(y) - \eta^{(1)}_n(y)\right)^+$. This proves (4.9).

Consider events corresponding to an overshoot of (1)- resp. (2)-particles at x in generation n

$$U^{(1)}(n,x) := \{\eta^{(1)}_n(x) > \eta^{(2)}_n(x)\}, \quad U^{(2)}(n,x) := \{\eta^{(1)}_n(x) < \eta^{(2)}_n(x)\},$$

and let

$$A(n+1,z) := \{\xi^{(1)}_{n+1}(z) \wedge \xi^{(2)}_{n+1}(z) > 0\},$$

be the event that an "annihilation" of (1)- and (2)-particles occurs at position z in generation $n+1$.

Let us first consider a pair $x \neq y$ such that there exists a $z \in \mathbb{Z}^d$ with $p(x,z), p(y,z) > 0$. Put $a_n(z) := \mathbb{E}|\eta^{(1)}_n(z) - \eta^{(2)}_n(z)|\ (\geq 0)$. Observing that

$$|\xi^{(1)}_{n+1}(z) - \xi^{(2)}_{n+1}(z)| \leq \xi^{(1)}_{n+1}(z) + \xi^{(2)}_{n+1}(z) - \mathbf{1}(\xi^{(1)}_{n+1}(z) \wedge \xi^{(2)}_{n+1}(z) > 0)$$

we see from the considerations above that

$$a_{n+1}(z) \leq a_n(z) - \mathbb{P}(A(n+1,z)) \leq a_n(z) - c_{x,y}\mathbb{P}(U^{(1)}(n,x) \cap U^{(2)}(n,y))$$

for some $c_{x,y} > 0$, because

$$\begin{aligned} &U^{(1)}(n,x) \cap U^{(2)}(n,y) \cap \{\text{some offspring from } x \text{ jumps to } z\} \\ &\quad \cap \{\text{some offspring from } y \text{ jumps to } z\} \subset A(n+1,z). \end{aligned}$$

This implies that $\mathbb{P}(U^{(1)}(n,x) \cap U^{(2)}(n,y)) \to 0$ as $n \to \infty$, i.e. (4.10) holds true for this pair (x,y), because the sequence $a_n(z)$ is bounded from below. The same argument shows that $\lim_n \mathbb{P}(A(n,z)) = 0$ for any $z \in \mathbb{Z}^d$.

Now consider a pair x, y with $p(x,z)p(y,z) = 0$ for all z, but such that there exist x', y' with $p(x,x'), p(y,y') > 0$ and $\lim_n \mathbb{P}(U^{(1)}(n,x') \cap U^{(2)}(n,y')) = 0$. Then we have

$$\begin{aligned} &U^{(1)}(n,x) \cap \{\text{some offspring from } x \text{ jumps to } x'\} \cap A(n+1,x')^c \\ &\cap U^{(2)}(n,y) \cap \{\text{some offspring from } y \text{ jumps to } y'\} \cap A(n+1,y')^c \\ &\qquad \subset U^{(1)}(n+1,x') \cap U^{(2)}(n+1,y'). \end{aligned}$$

Hence

$$\begin{aligned} &\mathbb{P}\begin{pmatrix} U^{(1)}(n,x) \cap \{\text{some offspring from } x \text{ jumps to } x'\} \\ \cap\, U^{(2)}(n,y) \cap \{\text{some offspring from } y \text{ jumps to } y'\} \end{pmatrix} \\ &\quad - \mathbb{P}(A(n+1,x') \cup A(n+1,y')) \leq \mathbb{P}(U^{(1)}(n+1,x') \cap U^{(2)}(n+1,y')), \end{aligned}$$

which gives

$$\begin{aligned} &c_{x,x',y,y'}\mathbb{P}(U^{(1)}(n,x) \cap U^{(2)}(n,y)) \\ &\quad \leq \mathbb{P}(A(n+1,x')) + \mathbb{P}(A(n+1,y')) + \mathbb{P}(U^{(1)}(n+1,x') \cap U^{(2)}(n+1,y')) \end{aligned}$$

with some $c_{x,x',y,y'} > 0$. This proves (4.10) also for this pair. Observe that by irreducibility of the difference random walk, any pair x, y can be connected via finitely many steps of the difference random walk. Thus (4.10) can be proved for any $x \neq y$ by induction. □

In the case $\alpha < \alpha_2$ we can use this together with the fact that η has asymptotic density θ under ν_θ to show that all equilibria (with finite intensity) are mixtures of the ν_θ:

Proposition 11 *Assume (4.3) and $\alpha < \alpha_2$. Then*

$$\nu \in \mathcal{M}_1(E_1) \text{ shift-ergodic, invariant under } (S_n),\ \int \eta(0)\nu(d\eta) = \theta \implies \nu = \nu_\theta.$$

The same conclusion holds if we replace the assumption of shift-ergodicity by the requirement $\nu \in \mathcal{R}_\theta$. Furthermore

$$\mu S_n \Rightarrow \nu_\theta \quad \text{as } n \to \infty \text{ for all } \mu \in \mathcal{R}_\theta.$$

Proof. Let ν be invariant under (S_n) with $\int \eta(0)\, d\nu = \theta$, and assume that either ν is shift-ergodic or $\nu \in \mathcal{R}_\theta$. We start the coupled system in $\nu_\theta \otimes \nu$ and show that the coupling is eventually successful, that is that $|\eta_n^{(1)} - \eta_n^{(2)}|$ converges (vaguely) to the zero configuration in probability. Because $\sup_{x,n} \mathbb{E}_{\nu_\theta \otimes \nu}\big[\eta_n^{(1)}(x) + \eta_n^{(2)}(x)\big] = 2\theta < \infty$ the family $\mathcal{L}((\eta_n^{(1)}, \eta_n^{(2)}))_{n \in \mathbb{N}_0}$ is tight (w.r.t. the vague topology on $E_1 \times E_1$), hence we can choose a subsequence satisfying $\mathcal{L}((\eta_{n_k}^{(1)}, \eta_{n_k}^{(2)})) \Rightarrow \tilde{\nu}$ as $k \to \infty$, where $\tilde{\nu}$ is some probability measure on $E_1 \times E_1$. $\tilde{\nu}$ has marginals ν_θ and ν as both are (S_n)-invariant; in particular the marginals are shift-invariant. By Lemma 26 we have furthermore for all $x, y \in \mathbb{Z}^d$

$$\tilde{\nu}\left(\{\eta^{(1)}(x) > \eta^{(2)}(x), \eta^{(1)}(y) < \eta^{(2)}(y)\}\right) = 0.$$

This allows to estimate for all N (using (4.10) in the second equality)

$$\begin{aligned}
&\int_{E_1 \times E_1} |\eta^{(1)}(0) - \eta^{(2)}(0)|\, \tilde{\nu}(d(\eta^{(1)}, \eta^{(2)})) \\
&\quad = \int_{E_1 \times E_1} \sum_x \bar{p}_N(0,x) |\eta^{(1)}(x) - \eta^{(2)}(x)|\, \tilde{\nu}(d(\eta^{(1)}, \eta^{(2)})) \\
&\quad = \int_{E_1 \times E_1} \Big| \sum_x \bar{p}_N(0,x)(\eta^{(1)}(x) - \eta^{(2)}(x)) \Big|\, \tilde{\nu}(d(\eta^{(1)}, \eta^{(2)})) \\
&\quad \le \int_{E_1} \Big| \sum_x \bar{p}_N(0,x)(\eta^{(1)}(x) - \theta) \Big| \nu_\theta(d\eta) \\
&\qquad + \int_{E_1} \Big| \sum_x \bar{p}_N(0,x)(\eta^{(2)}(x) - \theta) \Big| \nu(d\eta).
\end{aligned}$$

The first term in the last line converges to 0 as $N \to \infty$ because $\nu_\theta \in \mathcal{R}_\theta$ (observe that $L^2(\nu_\theta)$-convergence implies the same in $L^1(\nu_\theta)$). The second term converges to 0 by the same argument if $\nu \in \mathcal{R}_\theta$, it also converges to 0 if ν is shift-ergodic by an approximation argument involving finite boxes.

In order to prove the second claim let $\nu \in \mathcal{R}_\theta$ and consider a subsequence (n_k) such that $\nu S_{n_k} \Rightarrow \nu'$ for some $\nu' \in \mathcal{M}_1(E)$. A coupling argument similar to the first part of the proof shows that $\nu' = \nu_\theta$. Hence $\nu S_n \Rightarrow \nu_\theta$ because the family $(\nu S_n)_n$ is tight and any convergent subsequence has the right limit. □

By combining Proposition 11 and Corollary 2 we obtain

Corollary 3 *Assume $\alpha < \alpha_2$ and (4.8). Then the set of all shift-ergodic, (S_n)-invariant probability measures ν on E with $\int \eta(0)\, d\nu < \infty$ is given by $\{\nu_\theta : \theta \ge 0\}$.*

Chapter 5

Some quantities pertaining to two random walk paths

Let ξ, ξ' be two independent discrete-time random walks on $\mathbb{Z}^d$ with transition matrix $p(\cdot)$, $\xi_0 = \xi_0' = 0$. Let $\widetilde{\xi} = \xi - \xi'$ be the difference random walk with transition probabilities $\tilde{p}(x) := \sum_y p(x+y)p(y)$, and $\tilde{G}(x) := \sum_{n=0}^{\infty} \tilde{p}_n(x)$ the corresponding Green function. We assume that $\widetilde{\xi}$ is transient and furthermore that

$$\sup_{n\geq 1, x\in\mathbb{Z}^d} \frac{p_n(x)}{\tilde{p}_n(0)} < \infty. \tag{5.1}$$

The local CLT shows that p satisfies (5.1) whenever $\sum_x p(x)|x|^2 < \infty$, also any symmetric p in the domain of attraction of a stable law satisfies (5.1).

5.1 Exponential moments of the collision time of two independent discrete-time random walks, conditional on one of them

Let $V := \sum_{n=1}^{\infty} \mathbf{1}(\xi_n = \xi_n')$,

$$\alpha_* := \sup\{\alpha : \mathbb{E}[\alpha^V|\xi] < \infty \text{ almost surely}\}. \tag{5.2}$$

Defining

$$\alpha_2 := \sup\{\alpha : \mathbb{E}[\alpha^V] < \infty\}$$

we obviously have $\alpha_* \geq \alpha_2$. Our aim in this chapter is to compute α_* using a variational problem and in particular to prove that $\alpha_* > \alpha_2$ under assumption (5.1) for transient $\tilde{p}$.

Remark 20 Note that the event $\{\mathbb{E}[\alpha^V|\xi] < \infty\}$ is invariant under permutation of finitely many ξ-increments, so that we have $\mathbb{P}\left(\mathbb{E}[\alpha^V|\xi] < \infty\right) \in \{0,1\}$ by the Hewitt-Savage 0-1 law. In particular $\alpha > \alpha_*$ implies that $\mathbb{P}(\mathbb{E}[\alpha^V|\xi] = \infty) = 1$.

Thus we could alternatively choose a regular conditional distribution of V, given ξ, and define α_* through (5.2) without the term "almost surely" inside the braces. Then α_* would formally be a random variable, but its value would almost surely be equal to the righthand side of (5.2).

Theorem 5 *Assume that $\tilde{\xi}$ is transient and p satisfies (5.1). Then*

$$\alpha_* = 1 + \Big(\sum_{n \geq 1} \exp\big(- H(p_n(\cdot)) \big) \Big)^{-1},$$

where $H(p_n(\cdot)) = -\sum_x p_n(x) \log p_n(x)$ is the entropy of $p_n := p^{*n}$.

Corollary 4 *Under the assumptions of Theorem 5 we have $\alpha_* > \alpha_2 > 1$.*

As $\tilde{V} = \sum_{n=0}^{\infty} \mathbf{1}(\tilde{\xi}_n = 0)$ has a geometric distribution with parameter $\tilde{p}_{\text{esc}} = \mathbb{P}_0(\tilde{\xi}_n \neq 0, n > 0)$, we have

$$\alpha_2 = (1 - \tilde{p}_{\text{esc}})^{-1} = 1 + \frac{1}{\tilde{p}_{\text{esc}}^{-1} - 1} = 1 + \big(\tilde{G}(0) - 1\big)^{-1}. \tag{5.3}$$

Now Jensen's inequality shows that

$$\begin{aligned} \exp\left(-H(p_n(\cdot)\right) &= \exp\left(\textstyle\sum_x p_n(x) \log p_n(x)\right) \\ &< \sum_x p_n(x) e^{\log p_n(x)} = \sum_x p_n(x)^2 = \tilde{p}_n(0). \end{aligned}$$

□

Observe that $\mathbb{E}[(\alpha_2)^V] = \infty$ by the form of the geometric weights.

Proof of Thm. 5. The question whether or not the conditional distribution of V given ξ possesses an exponential moment of a given order translates naturally into the question about the growth of its moments, which we consider in the following. It will be more convenient to work with factorial moments, note that for any $\alpha \geq 1$ and $v \in \mathbb{N}_0$ we have

$$\alpha^v = \big((\alpha - 1) + 1\big)^v = \sum_{k=0}^{\infty} \binom{v}{k} (\alpha - 1)^k = 1 + \sum_{k=1}^{\infty} (\alpha - 1)^k \frac{[v]_k}{k!}$$

by the Binomial theorem, where $[v]_k := v(v-1)\cdots(v-k+1)$ is the k-th falling factorial of v. Now $V = \sum_{n=1}^{\infty} \mathbf{1}(\xi_n = \xi'_n)$, thus

$$\frac{[V]_k}{k!} = \sum_{0 < j_1 < j_2 < \cdots < j_k} \mathbf{1}(\xi_{j_1} = \xi'_{j_1}, \ldots, \xi_{j_k} = \xi'_{j_k})$$

is the number of ordered pairwise distinct k-tuples of times at which the two paths meet. This allows to rewrite

$$\mathbb{E}\left[\alpha^V \,|\, \xi\right] = 1 + \sum_{k=1}^{\infty} (\alpha - 1)^k \mathbb{E}\Big[\frac{[V]_k}{k!} |\xi\Big] = 1 + \sum_{k=1}^{\infty} (\alpha - 1)^k F_k(\xi),$$

where

$$F_k(\xi) := \sum_{0 < j_1 < j_2 < \cdots < j_k} \mathbb{P}(\xi_{j_1} = \xi'_{j_1}, \ldots, \xi_{j_k} = \xi'_{j_k} | \xi)$$

is the conditional k-th ordered falling factorial moment of V given ξ. Thus if we can show that

$$\lim_k \frac{1}{k} \log F_k(\xi) =: r \tag{5.4}$$

exits and is $\mathcal{L}(\xi)$-almost surely constant, we have shown that $\alpha_* - 1 = e^{-r}$. Of course this method gives no information about the exponential moment exactly at the threshold value α_* itself.

In order to do this let us denote the (i.i.d.) sequence of ξ-increments by $X = (X_i)$, $X_i = \xi_i - \xi_{i-1}$, $i = 1, 2, \dots$ and write

$$\begin{aligned} F_k(\xi) &= \sum_{0<j_1<j_2<\cdots<j_k} p_{j_1}(\xi_{j_1}) p_{j_2-j_1}(\xi_{j_2} - \xi_{j_1}) \times \cdots \times p_{j_k - j_{k-1}}(\xi_{j_k} - \xi_{j_{k-1}}) \\ &= \sum_{\ell_1,\dots,\ell_k=1}^{\infty} \prod_{i=1}^{k} p_{\ell_i}(\xi_{\ell_1+\cdots+\ell_i} - \xi_{\ell_1+\cdots+\ell_{i-1}}) = \sum_{\ell_1,\dots,\ell_k=1}^{\infty} \prod_{i=1}^{k} p_{\ell_i}(\Delta_i(\ell, X)) \end{aligned}$$

where $\Delta_i(\ell, X) := X_{\ell_1+\cdots+\ell_{i-1}+1} + X_{\ell_1+\cdots+\ell_{i-1}+2} + \cdots + X_{\ell_1+\cdots+\ell_{i-1}+\ell_i}$.

Let $R_1, R_2, \dots$ be i.i.d. with $\mathbb{P}(R_1 = n) = \tilde{p}_n(0)/(\tilde{G}(0) - 1)$, $n = 1, 2, \dots$ (and independent of X), set $T_n := R_1 + \cdots + R_n$ and

$$Y_i := (X_{T_i+1}, X_{T_i+2}, \dots, X_{T_{i+1}}),\ i = 0, 1, \dots.$$

We can rewrite

$$\begin{aligned} F_k(\xi) &= (\tilde{G}(0) - 1)^k \sum_{r_1,\dots,r_k=1}^{\infty} \prod_{i=1}^{k} \mathbb{P}(R_i = r_i) \frac{p_{r_i}(\Delta_i((r_j), X))}{\tilde{p}_{r_i}(0)} \\ &= (\tilde{G}(0) - 1)^k \mathbb{E}\left[\prod_{i=1}^{k} \frac{p_{R_i}(\Delta_i((R_j), X))}{\tilde{p}_{R_i}(0)} \middle| X \right] \\ &= (\tilde{G}(0) - 1)^k \mathbb{E}\left[\prod_{i=0}^{k-1} f\big((X_{T_i+1}, X_{T_i+2}, \dots, X_{T_{i+1}})\big) \middle| X \right] \\ &= (\tilde{G}(0) - 1)^k \mathbb{E}\left[\exp\left(\sum_{i=0}^{k-1} \log f(Y_k) \right) \middle| X \right] \\ &= (\tilde{G}(0) - 1)^k \mathbb{E}\left[\exp\left(k \int \log f(y)\, \mathcal{Y}_k(dy) \right) \middle| X \right], \end{aligned} \tag{5.5}$$

where $f : \bigcup_{n=1}^{\infty} \left(\mathbb{Z}^d\right)^n \to \mathbb{R}_+$ is defined by

$$f(x) = \frac{p_n(x_1 + \cdots + x_n)}{\tilde{p}_n(0)} \qquad \text{for } x = (x_1, x_2, \dots, x_n) \in \mathbb{Z}^n$$

and

$$\mathcal{Y}_k := \frac{1}{k} \sum_{i=0}^{k-1} \delta_{Y_i}$$

is the empirical distribution of $Y_0, Y_1, \dots, Y_{k-1}$.

Now according to Theorem 6 from section 5.2 the family $\mathcal{L}(\mathcal{Y}_k | X)$ almost surely satisfies a large deviation principle with rate k and ('good') rate function

$$\tilde{h}(\nu) = \begin{cases} h(\tilde{\nu}; \mathcal{L}(R_1)) & \text{if } \nu(dy) = \sum_{n=1}^{\infty} \tilde{\nu}_n \mathbf{1}_{(\mathbb{Z}^d)^n}(y) (\mathcal{L}(X_1))^{\otimes n}(dy) \\ & \text{for some } \tilde{\nu} \in \mathcal{M}_1(\mathbb{N}), \\ \infty & \text{otherwise,} \end{cases}$$

where $h(\mu; \rho) = \sum_a \mu_a \log(\mu_a/\rho_a)$ is the relative entropy of (the probability measure) μ with respect to (the probability measure) ρ.

So we can use Varadhan's lemma (see e.g. [17], Thm. III.3 or [8], Thm. 2.1.10) to compute (here our assumption (5.1), which ensures that $\log f$ is bounded from above, comes into play)

$$\begin{aligned} s &:= \lim_{k\to\infty} \frac{1}{k} \log \mathbb{E}\left[\exp\left(k \int \log f(y)\, \mathcal{Y}_k(dy) \right) \Big| X \right] \\ &= \sup_{\nu\in\mathcal{M}_1\left(\cup_n(\mathbb{Z}^d)^n\right)} \left\{ \int \log f(y)\, \nu(dy) - \tilde{h}(\nu) \right\} \qquad (5.6) \\ &= \sup_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} \left\{ \sum_{n\geq 1} \tilde{\nu}_n \int \log f(y)\, \mathcal{L}(X_1)^{\otimes n}(dy) - h(\tilde{\nu}; \mathcal{L}(R_1)) \right\} \\ &= \sup_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} \left\{ \sum_{n\geq 1} \tilde{\nu}_n \sum_{x\in\mathbb{Z}^d} p_n(x) \log \frac{p_n(x)}{\tilde{p}_n(0)} - \sum_{n\geq 1} \tilde{\nu}_n \log \frac{\tilde{\nu}_n}{\tilde{p}_n(0)/(\tilde{G}(0)-1)} \right\} \\ &= \sup_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} \left\{ \sum_{n\geq 1} \sum_{x\in\mathbb{Z}^d} \tilde{\nu}_n p_n(x) \log \frac{p_n(x)}{\tilde{\nu}_n(\tilde{G}(0)-1)} \right\}. \end{aligned}$$

Combining this with (5.5) we conclude that $\mathcal{L}(\xi)$-almost surely

$$\begin{aligned} r &= \lim \frac{1}{k} \log F_k(\xi) = \log(\tilde{G}(0)-1) + s \\ &= \sup_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} \left\{ -\sum_{n\geq 1} \tilde{\nu}_n \log \tilde{\nu}_n + \sum_{n\geq 1} \tilde{\nu}_n \sum_{x\in\mathbb{Z}^d} p_n(x) \log p_n(x) \right\} \\ &= \sup_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} \left\{ H(\tilde{\nu}) - \sum_{n\geq 1} \tilde{\nu}_n H(p_n) \right\}, \end{aligned}$$

where $H(\mu) = -\sum_a \mu_a \log \mu_a$ is the entropy of μ. For a given sequence $f_n \geq 0$ with

$$C := \sum_{n\geq 1} \exp(-f_n) < \infty, \qquad (5.7)$$

and $\nu^* = (C^{-1}\exp(-f_n))_{n\in\mathbb{N}}$, the variational problem

$$\begin{aligned} r &= \sup\left\{ H(\tilde{\nu}) - \sum_{n\geq 1} \tilde{\nu}_n f_n \;:\; \tilde{\nu}\in\mathcal{M}_1(\mathbb{N}) \right\} \\ &= \sup\left\{ -\sum_{n\geq 1} \tilde{\nu}_n \log(\tilde{\nu}_n e^{f_n}) \;:\; \tilde{\nu}\in\mathcal{M}_1(\mathbb{N}) \right\} \\ &= \log C + \sup\left\{ -\sum_{n\geq 1} \tilde{\nu}_n \log \frac{\tilde{\nu}_n}{\frac{1}{C}e^{-f_n}} \;:\; \tilde{\nu}\in\mathcal{M}_1(\mathbb{N}) \right\} \\ &= \log C - \inf_{\tilde{\nu}\in\mathcal{M}_1(\mathbb{N})} h(\tilde{\nu}; \nu^*) \end{aligned}$$

has the obvious maximiser $\tilde{\nu} = \nu^*$, leading to $r = \log C$. We apply this to $f_n := H(p_n)$ to prove the theorem. Observe that condition (5.7) is satisfied because $\exp(-H(p_n)) < \tilde{p}_n(0)$ by Jensen's inequality, and the assumed transience of $\widetilde{\xi}$ ensures that $\tilde{p}_n(0)$ is summable. $\square$

5.2 An almost sure Sanov-type theorem involving a conditional distribution

Let $X = (X_0, X_1, \ldots)$ be a sequence of i.i.d. random variables with values in some Polish space E, $R_1, R_2, \ldots$ be i.i.d. $\mathbb{N}$-valued, independent of X. Set $T_0 := 0$, $T_n := R_1 + \cdots + R_n$.

Let us observe the sequence $X_0, X_1, \ldots$ in pieces cut out by the renewal times by setting

$$Y_i := (X_{T_i}, X_{T_i+1}, \ldots, X_{T_{i+1}-1}),\ i = 0, 1, 2, \ldots$$

Obviously Y_i are $\widetilde{E} := \cup_{n=1}^{\infty} E^n$ valued random variables. $\widetilde{E}$ is a Polish space e.g. with the metric

$$\widetilde{\mathrm{d}}(x, y) := \begin{cases} \sum_{i=1}^{n} (\mathrm{d}(x_i, y_i) \wedge 1) & \text{if } x, y \in E^n \\ 1 & \text{otherwise.} \end{cases}$$

Let us consider the empirical distributions

$$\mathcal{Y}_n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{Y_i} \quad \text{with values in } \mathcal{M}_1(\widetilde{E}).$$

By assumption the sequence $Y_0, Y_1, \ldots$ is i.i.d. with $\mathcal{L}(Y_0) = \mu$, where

$$\mu(B) = \mathbb{P}((X_0, X_1, \ldots, X_{R_1 - 1}) \in B) \quad \text{for measurable } B \subset \widetilde{E}.$$

We equip $\mathcal{M}_1(\widetilde{E})$ with the weak topology. The law of large numbers implies that

$$\mathcal{Y}_n \longrightarrow \mu \qquad \text{almost surely as } n \to \infty.$$

Furthermore Sanov's theorem states that $(\mathcal{L}(\mathcal{Y}_n))_{n \in \mathbb{N}}$ satisfies the following large deviation principle with (rate n)

$$-H(\operatorname{int} \mathcal{B}; \mu) \le \liminf_n \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{B}) \le \limsup_n \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{B}) \le -H(\operatorname{cl} \mathcal{B}; \mu)$$

for any measurable $\mathcal{B} \subset \mathcal{M}_1(\widetilde{E})$. Here $H(\mathcal{B}; \mu) = \inf_{\nu \in \mathcal{B}} h(\nu; \mu)$ and

$$h(\nu; \mu) = \begin{cases} \int \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu & \text{if } \nu \ll \mu \\ \infty & \text{otherwise} \end{cases}$$

is the relative entropy of ν with respect to μ. $h(\cdot; \mu)$ is a 'good' rate function, see e.g. [8], Thm. 3.2.17.

What can we say about $(\mathcal{Y}_n)_n$ *conditional* on the sequence X?

In the following we always think of a regular version of the conditional distribution of (X, R), given X, which is guaranteed to exist by the assumed 'Polishness' of the underlying spaces (see e.g. [27], Thm. V.8.1).
First, an almost sure convergence statement with a deterministic limit automatically translates into almost surely the same behaviour, given X:

$$\mathbb{P}(\{\mathcal{Y}_n \to \mu\}) = 1 \iff \Big(\mathbb{P}(\{\mathcal{Y}_n \to \mu\} | X) = 1 \quad \mathcal{L}(X) \text{ almost surely} \Big).$$

Consider

$$\tilde{\mathcal{M}} := \{\nu \in \mathcal{M}_1(\widetilde{E}) \,|\, \forall\, n : \ \nu(\cdot | E^n) = (\mathcal{L}(X_0))^{\otimes n}\}$$

Observe that $\tilde{\mathcal{M}}$ is a closed subset with respect to the weak topology (and is isomorphic to $\mathcal{M}_1(\mathbb{N})$ via $\tilde{\mathcal{M}} \ni \nu \mapsto \tilde{\nu} = (\nu(E_n))_{n\in\mathbb{N}} \in \mathcal{M}_1(\mathbb{N})$). Define a modified rate function

$$\tilde{h}(\nu;\mu) := \begin{cases} h(\nu;\mu) & \text{if } \nu \in \tilde{\mathcal{M}} \\ \infty & \text{otherwise.} \end{cases}$$

Set $I(\mathcal{B}) := \inf_{\nu\in\mathcal{B}} \tilde{h}(\nu;\mu)$.

Theorem 6 *The family $(\mathcal{L}(\mathcal{Y}_n|X))_{n\in\mathbb{N}}$ almost surely satisfies a large deviation principle with rate function I, more precisely the event*

$$\Big\{ \liminf \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O} \,|\, X) \geq -I(\mathcal{O}) \quad \textit{for all open } \mathcal{O} \subset \mathcal{M}_1(\widetilde{E}) \Big\} \bigcap$$

$$\Big\{ \limsup \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{C} \,|\, X) \leq -I(\mathcal{C}) \quad \textit{for all closed } \mathcal{C} \subset \mathcal{M}_1(\widetilde{E}) \Big\}$$

has probability one.

Remarks 1) The statement maybe looks more scary than necessary: As $\mathcal{M}_1(\widetilde{E})$ is Polish (see e.g. [8], Lemma 3.22) and therefore second countable, it suffices to consider *countable* families of open resp. closed sets. See below.

2) The proof follows a choreography that is quite common in large deviation theory: the upper bound uses an exponential Markov inequality, the lower bound is proved by tilting.

3) The theorem is similar to a main result in [3]. Yet there appears to be no way to derive the theorem directly from Comets' Theorem III.1.

Proof. 1) Let $\mathcal{C} \subset \mathcal{M}_1(\widetilde{E})$ be closed. We first consider the case $I(\mathcal{C}) < \infty$. Choose a decreasing sequence $\tilde{\mathcal{C}}_m$ of closed neighbourhoods of $\tilde{\mathcal{M}}$ such that $\cap_m \tilde{\mathcal{C}}_m = \tilde{\mathcal{M}}$. Put $\mathcal{C}_m := \mathcal{C} \cap \tilde{\mathcal{C}}_m$. Fix $\varepsilon > 0$ and choose m so large that[1]

$$H(\mathcal{C}_m;\mu) \geq I(\mathcal{C}) - \varepsilon. \tag{5.8}$$

The (unconditional) Sanov theorem shows that

$$\mathbb{P}(\mathcal{Y}_n \in \mathcal{C}_m) \leq \exp(-n(H(\mathcal{C}_m;\mu) - \varepsilon)) \tag{5.9}$$

for n sufficiently large.

[1] We have to show that

$$\inf_{\nu\in\mathcal{C}\cap\tilde{\mathcal{C}}_m} h(\nu;\mu) \nearrow \inf_{\nu\in\mathcal{C}\cap\tilde{\mathcal{M}}} h(\nu;\mu) \quad \text{as } m \to \infty$$

(with the usual convention $\inf_{\nu\in\emptyset} h(\nu;\mu) = +\infty$). Assume that on the contrary

$$M := \sup_m \inf_{\nu\in\mathcal{C}\cap\tilde{\mathcal{C}}_m} h(\nu;\mu) < \inf_{\nu\in\mathcal{C}\cap\tilde{\mathcal{M}}} h(\nu;\mu) =: M',$$

in particular $M < \infty$. Then there would exist $\varepsilon \in (0, M' - M)$ and for each m a $\nu_m \in \mathcal{C} \cap \tilde{\mathcal{C}}_m$ with $h(\nu_m;\mu) \leq M + \varepsilon$. As $h(\cdot;\mu)$ has compact level sets we could choose a subsequence $(m_k)_k$ such that

$$\nu_{m_k} \to \nu_* \in \cap_m \left(\mathcal{C} \cap \tilde{\mathcal{C}}_m\right) = \mathcal{C} \cap \tilde{\mathcal{M}}$$

with $h(\nu_*;\mu) \leq M + \varepsilon < M'$ in contradiction to the definition of M'.

Observe that $\mu \in \tilde{\mathcal{M}} \subset \tilde{\mathcal{C}}_m$, hence $\tilde{\mathcal{C}}_m$ is a neighbourhood of μ. From the strong law we conclude the (almost sure) existence of a finite-valued random variable $N(X)$ such that

$$\mathbb{P}\left(\{\mathcal{Y}_n \in \tilde{\mathcal{C}}_m \text{ for } n \geq N(X)\}|X\right) = 1. \tag{5.10}$$

Now consider the events

$$A(n,\varepsilon) := \{\mathbb{P}(\mathcal{Y}_n \in \mathcal{C}|X) \geq \exp(-n(I(\mathcal{C}) - 3\varepsilon))\} .$$

The (exponential) Markov inequality shows that

$$\begin{aligned}
&\exp(-n(I(\mathcal{C}) - 3\varepsilon))\mathbb{P}(A(n,\varepsilon) \cap \{n \geq N(X)\}) \\
&\qquad \leq \mathbb{E}\left[\mathbf{1}(A(n,\varepsilon) \cap \{n \geq N(X)\})\mathbb{P}(\mathcal{Y}_n \in \mathcal{C}|X)\right] \\
&\qquad = \mathbb{E}\left[\mathbf{1}(A(n,\varepsilon) \cap \{n \geq N(X)\})\mathbb{P}(\mathcal{Y}_n \in \mathcal{C}_m|X)\right] \\
&\qquad \leq \mathbb{P}(\mathcal{Y}_n \in \mathcal{C}_m).
\end{aligned}$$

Combining this with (5.8) and (5.9) we see that

$$\mathbb{P}(A(n,\varepsilon) \cap \{n \geq N(X)\}) \leq \exp(-n\varepsilon)$$

for n sufficiently large. The Borel-Cantelli lemma shows that with probability one only finitely many of the events $A(n,\varepsilon) \cap \{n \geq N(X)\}$ occur, which together with $N(X) < \infty$ almost surely proves that the event

$$B(\varepsilon) := \left\{\limsup_n \frac{1}{n} \log \mathbb{P}(\{\mathcal{Y}_n \in \mathcal{C}\}|X) \leq -(I(\mathcal{C}) - 3\varepsilon)\right\}$$

has probability one. Then of course also

$$\mathbb{U}(\mathcal{C}) := \left\{\limsup_n \frac{1}{n} \log \mathbb{P}(\{\mathcal{Y}_n \in \mathcal{C}\}|X) \leq -I(\mathcal{C})\right\} = \bigcap_k B(1/k)$$

occurs almost surely.
The case $I(\mathcal{C}) = \infty$ can be treated similarly.
For completeness, here are some details: Fix $M > 1$ and choose m big enough such that $H(\mathcal{C}_m; \mu) \geq M$, where $(\mathcal{C}_m)_m$ is as above. Then again by Sanov's theorem $\mathbb{P}(\mathcal{Y}_n \in (\mathcal{C}_m)) \leq \exp(-n(M-1))$ for n big enough. Consider

$$A(n,M) := \{\mathbb{P}(\mathcal{Y}_n \in \mathcal{C}|X) \geq \exp(-n(M-2))\} .$$

Arguing as above we see that $\mathbb{P}(A(n,M)) \leq \exp(-n)$ and we can again conclude using the Borel-Cantelli lemma.

Finally we want to improve this result to hold for all closed sets simultaneously with probability one. This is standard in a Polish setting, see e.g. [3]: Let the family $(\mathcal{O}_k)_{k\in\mathbb{N}}$ of open sets be a countable basis of the topology of $\mathcal{M}_1(\widetilde{E})$, and define

$$\mathbb{U} := \bigcap_{B \subset \mathbb{N}, |B| < \infty} \bigcap_{k \in B} \mathbb{U}((\mathcal{O}_k)^c).$$

Then $\mathbb{P}(\mathbb{U}) = 1$. Now any closed $\mathcal{C} \subset \mathcal{M}_1(\widetilde{E})$ can be written as $\mathcal{C} = \cap_{k\in B'}(\mathcal{O}_k)^c$ for some subset $B' \subset \mathbb{N}$. Fix $\varepsilon > 0$ and choose a finite $B \subset B'$ with $I(\cap_{k\in B}(\mathcal{O}_k)^c) \geq I(\mathcal{C}) - \varepsilon$. Then

$$\mathbb{U} \subset \left\{ \begin{array}{rl} \limsup \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{C}|X) \leq & \limsup \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \cap_{k\in B}(\mathcal{O}_k)^c|X) \\ \leq & -I(\cap_{k\in B}(\mathcal{O}_k)^c) \leq -I(\mathcal{C}) + \varepsilon \end{array} \right\} .$$

Take $\varepsilon \to 0$ to see that on $\mathbb{U}$, the upper bound holds simultaneously for all closed sets.

2) Let $\mathcal{O} \subset \mathcal{M}_1(\widetilde{E})$ be open. We can assume without loss of generality that $I(\mathcal{O}) < \infty$. Consider $\nu \in \mathcal{O} \cap \tilde{\mathcal{M}}$ with $h(\nu; \mu) < \infty$. We can write ν as

$$\nu(dy) = \sum_{n=1}^{\infty} \tilde{\nu}_n \mathbf{1}_{E^n}(y)(\mathcal{L}(X_0))^{\otimes n}(dy)$$

with some probability measure $\tilde{\nu} = (\tilde{\nu}_n)_{n\in\mathbb{N}} \in \mathcal{M}_1(\mathbb{N})$, and of course $h(\tilde{\nu}; \rho) = h(\nu; \mu) < \infty$, where $\rho = \mathcal{L}(R_1)$. Consider a probability measure $\tilde{\mathbb{P}}$ with density

$$\prod_{i=1}^{n} \frac{\tilde{\nu}_{R_i}}{\rho_{R_i}} \quad \text{on } \sigma(R_1, \ldots, R_n, X) \text{ with respect to } \mathbb{P},$$

that is, under $\tilde{\mathbb{P}}$ the (X_i) and the (R_i) are still independent i.i.d. sequences, but each R_i has distribution $\tilde{\nu}$. As $\mathcal{Y}_n = F_n(X, R_1, R_2, \ldots, R_n)$ is a function of X and $R_1, \ldots, R_n$, we can estimate for any $\varepsilon > 0$

$$\begin{aligned}
\mathbb{P}(\mathcal{Y}_n \in \mathcal{O}|X) &= \sum_{r_1,\ldots,r_n=1}^{\infty} \left(\prod_{i=1}^{n} \rho_{r_i}\right) \mathbf{1}(F_n(X, r_1, \ldots, r_n) \in \mathcal{O}) \\
&= \sum_{r_1,\ldots,r_n=1}^{\infty} \left(\prod_{i=1}^{n} \tilde{\nu}_{r_i}\right) \exp\Big(- \sum_{j=1}^{n} \log \frac{\tilde{\nu}_{r_j}}{\rho_{r_j}}\Big) \mathbf{1}(F_n(X, r_1, \ldots, r_n) \in \mathcal{O}) \\
&= \tilde{\mathbb{E}}\left[\exp\Big(- \sum_{j=1}^{n} \log \frac{\tilde{\nu}_{R_j}}{\rho_{R_j}}\Big) \mathbf{1}(\mathcal{Y}_n \in \mathcal{O}) \,\middle|\, X \right] \\
&\geq \exp\big(- n(h(\tilde{\nu}; \rho) + \varepsilon)\big) \tilde{\mathbb{P}}\Big(\frac{1}{n} \sum_{j=1}^{n} \log \frac{\tilde{\nu}_{R_j}}{\rho_{R_j}} \leq h(\tilde{\nu}; \rho) + \varepsilon,\ \mathcal{Y}_n \in \mathcal{O} \,\Big|\, X \Big).
\end{aligned}$$

Now the law of large numbers implies that

$$\frac{1}{n} \sum_{j=1}^{n} \log \frac{\tilde{\nu}_{R_j}}{\rho_{R_j}} \longrightarrow \tilde{\mathbb{E}}\big[\log \frac{\tilde{\nu}_{R_1}}{\rho_{R_1}} \big] = h(\tilde{\nu}; \rho) \quad \text{as } n \to \infty \text{ almost surely under } \tilde{\mathbb{P}}$$

and as $\mathcal{O}$ is an open neighbourhood of ν we also have $\tilde{\mathbb{P}}$-almost surely $\mathcal{Y}_n \in \mathcal{O}$ for all sufficiently large n, showing

$$\tilde{\mathbb{P}}\Big(\frac{1}{n} \sum_{j=1}^{n} \log \frac{\tilde{\nu}_{R_j}}{\rho_{R_j}} \leq h(\tilde{\nu}; \rho) + \varepsilon,\ \mathcal{Y}_n \in \mathcal{O} \,\Big|\, X \Big) \longrightarrow 1 \qquad \text{as } n \to \infty$$

for $\tilde{\mathbb{P}}|_{\sigma(X)}$-almost all X. But $\tilde{\mathbb{P}}|_{\sigma(X)} = \mathbb{P}|_{\sigma(X)}$, so we can conclude from the above that

$$\liminf_{n} \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O}|X) \geq -h(\tilde{\nu}; \rho) - \varepsilon$$

holds $\mathbb{P}$-almost surely. Taking $\varepsilon \searrow 0$ along a countable sequence we see that

$$\liminf_{n} \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O}|X) \geq -h(\nu; \mu) \qquad \mathbb{P}\text{-almost surely}$$

for any $\nu \in \mathcal{O} \cap \tilde{\mathcal{M}}$ with $h(\nu;\mu) < \infty$. Finally choose a sequence $(\nu_k) \subset \mathcal{O}$ with $\tilde{h}(\nu_k;\mu) \searrow I(\mathcal{O})$ to obtain the desired conclusion:

$$\mathbb{L}(\mathcal{O}) := \big\{ \liminf_n \frac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O}|X) \geq -I(\mathcal{O}) \big\} \quad \text{has probability one.}$$

Again we want to strengthen this to hold with probability one simultaneously for all open sets. Consider a family $(\mathcal{O}_k)_{k\in\mathbb{N}}$ of open sets that forms a basis of the topology of $\mathcal{M}_1(\widetilde{E})$. Of course the event $\mathbb{L} := \cap_k \mathbb{L}(\mathcal{O}_k)$ has probability one. Any open $\mathcal{O}$ can be written as $\mathcal{O} = \cup_{k\in B'} \mathcal{O}_k$ for some $B' \subset \mathbb{N}$. For $\varepsilon > 0$ there exists $k \in B'$ such that $I(\mathcal{O}_k) \leq I(\mathcal{O}) + \varepsilon$, hence

$$\mathbb{L} \subset \left\{ \liminf \tfrac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O}|X) \geq \liminf \tfrac{1}{n} \log \mathbb{P}(\mathcal{Y}_n \in \mathcal{O}_k|X) \geq -I(\mathcal{O}_k) \geq -I(\mathcal{O}) - \varepsilon \right\} .$$

Take $\varepsilon \to 0$ to see that on $\mathbb{L}$, the lower bound holds simultaneously for all open sets.

Remark 21 Let us consider a related situation: Let $R_1, R_2, \ldots$ be as above and $\tilde{X}^{(0)}, \tilde{X}^{(1)}, \ldots$ be independent copies of X. Define an $\widetilde{E}$ valued i.i.d. sequence

$$\tilde{Y}_i := (\tilde{X}_0^{(i)}, \tilde{X}_1^{(i)}, \ldots, \tilde{X}_{R_{i+1}-1}^{(i)}), \quad i = 0, 1, 2, \ldots$$

and the corresponding empirical distributions

$$\tilde{\mathcal{Y}}_n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\tilde{Y}_i}.$$

Then not only are $(\tilde{\mathcal{Y}}_n)_n$ and $(\mathcal{Y}_n)_n$ equal in distribution, but also the families

$$\mathcal{L}(\mathcal{Y}_n|X) \quad \text{and} \quad \mathcal{L}(\tilde{\mathcal{Y}}_n|\tilde{X}^{(j)}, j \geq 0)$$

almost surely satisfy a large deviation principle *with the same rate function*. This follows from Comets' Theorem III.1, see [3].

5.3 A numerical example

Here we consider numerical approximations to α_2 and α_* for our favourite p, which is

$$p(x) = \begin{cases} \frac{1}{7} & x \in \mathbb{Z}^3, ||x||_2 \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

Observe that p is symmetric, aperiodic, irreducible and a p-random walk (as well as the difference of two independent p-rws) is transient. In view of (5.3) and Theorem 5 we have to compute

$$A := \sum_{n\geq 1} \tilde{p}_n(0) \quad \text{and} \quad B := \sum_{n\geq 1} \exp\left(-H(p_n(\cdot))\right).$$

For an increment X with $\mathbb{P}(X = x) = p(x)$ we have $\mathbb{E}\, X = 0$ and $\Sigma := \mathrm{Cov}(X, X) = (2/7)\mathrm{Id}_3$, so the local CLT shows

$$p_n(x) \sim \frac{1}{(2\pi|\Sigma|n)^{3/2}} \exp\left(-\frac{1}{2n} x^T \Sigma^{-1} x\right),$$

in particular

$$\tilde{p}_n(0) = p_{2n}(0) \sim (8\pi n/7)^{-3/2}. \tag{5.12}$$

As p has finite range we can compute $\tilde{p}_n$ explicitly for $n = 1, \ldots, N_0$ and use (5.12) for $n > N_0$, giving the approximation

$$\hat{A}_{N_0} = \sum_{n=1}^{N_0} \tilde{p}_n(0) + (8\pi/7)^{-3/2} \left(\zeta(3/2) - \sum_{n=1}^{N_0} \frac{1}{n^{3/2}} \right).$$

The choice $N_0 = 100$ leads to the numerical value of $\hat{A}_{100} = 0.3485214 + 0.0293247 = 0.377846$, which translates into an estimate $\hat{\alpha}_2 = 3.6466$.

To approximate α_* we consider

$$\hat{B}_{N_0} = \sum_{n=1}^{N_0} \exp\left(-H(p_n(\cdot))\right) + (4\pi e/7)^{-3/2} \left(\zeta(3/2) - \sum_{n=1}^{N_0} \frac{1}{n^{3/2}} \right)$$

($p_n \approx \mathcal{N}(0, (2/7)n\mathrm{Id}_3)$, whose entropy is $(3/2)\log(4\pi n/7) + 3/2$). $\hat{B}_{240} = 0.309944$, translating into $\hat{\alpha}_* = 4.2264$.

5.4 Exponential growth in the $\alpha > \alpha_2$ regime

Let $\widetilde{\xi}$ be a $\tilde{p}$-random walk, $\tilde{V}_n := \#\{1 \le i \le n : \widetilde{\xi}_i = 0\}$. We ask about the growth (as $n \to \infty$) of

$$a_n(\alpha) := \mathbb{E}_0\, \alpha^{\tilde{V}_n}, \quad \alpha \ge 1. \tag{5.13}$$

We have $a_{n+1}(\alpha) \ge a_n(\alpha)$ for all n and $\alpha \ge 1$, furthermore

$$\mathbb{E}_x\, \alpha^{\tilde{V}_n} = \sum_{k=0}^{n} \mathbb{P}_x(\tilde{T}_0 = k) a_{n-k}(\alpha) \le a_n$$

for any $x \in \mathbb{Z}^d$, where $\tilde{T}_0 = \min\{k \ge 0 : \widetilde{\xi}_k = 0\}$. Thus by the Markov property also

$$\begin{aligned} a_{n+m} &= \sum_x \mathbb{E}_0[\alpha^{\tilde{V}_{n+m}} \mathbf{1}(\widetilde{\xi}_m = x)] = \sum_x \mathbb{E}_0[\alpha^{\tilde{V}_m} \mathbf{1}(\widetilde{\xi}_m = x)] \mathbb{E}_x\, \alpha^{\tilde{V}_n} \\ &\le a_n \mathbb{E}_0[\alpha^{\tilde{V}_m}] = a_n\, a_m, \end{aligned}$$

so that

$$\gamma(\alpha) := \lim_n \frac{1}{n} \log a_n(\alpha) = \inf_n \frac{\log a_n(\alpha)}{n} \ (\ge 0) \tag{5.14}$$

exists by the Subadditivity lemma (see e.g. [21], Lemma 10.21). Of course for $\alpha < \alpha_2$ given by (5.3) the sequence $a_n(\alpha)$ is bounded and thus $\gamma(\alpha) = 0$. Also as $\mathbb{P}(\tilde{V}_n = k) \le (1 - \tilde{p}_{\mathrm{esc}})^k$ we see that $a_n(\alpha_2)$ grows at most linearly, proving that $\gamma(\alpha_2) = 0$. Here is a "quick'n'dirty" argument why $\gamma(\alpha) > 0$ for $\alpha = (1+\varepsilon)/(1 - \tilde{p}_{\mathrm{esc}}) > \alpha_2$:
Choose N_0 so big that $\mathbb{P}_0(\tilde{T}_0 \le N_0) \ge (1 - \tilde{p}_{\mathrm{esc}})/(1 + \varepsilon/2)$. Then

$$a_n(\alpha) \ge \alpha^{[n/N_0]} \mathbb{P}(\tilde{V}_n = [\frac{n}{N_0}]) \ge \left(\alpha \mathbb{P}_0(\tilde{T}_0 \le N_0)\right)^{[n/N_0]} \tilde{p}_{\mathrm{esc}} \ge \left(\frac{1+\varepsilon}{1+\varepsilon/2}\right)^{[n/N_0]} \tilde{p}_{\mathrm{esc}},$$

proving that $\liminf n^{-1} \log a_n(\alpha) \ge (N_0)^{-1} \log\left((1+\varepsilon)/(1+\varepsilon/2)\right) > 0$.

5.5 Analogous quantities for continuous-time random walks

Let $(p_{xy}) = (p_{0,y-x})$ be a (shift invariant) stochastic matrix on $\mathbb{Z}^d$ and let X and X' be two continuous-time random walks with rate matrix $\kappa(p_{xy} - \delta_{xy})$, starting from $X(0) = X'(0) = 0$. $V := \int_0^\infty \mathbf{1}(X(s) = X'(s))\, ds$ is their collision time. Here we are interested in

$$\beta_2 \quad := \quad \sup\{\beta \ge 0 : \mathbb{E}[\exp(\beta V)] < \infty\}, \tag{5.15}$$

$$\beta_* \quad := \quad \sup\{\beta \ge 0 : \mathbb{E}[\exp(\beta V)|X] < \infty \;\; \mathcal{L}(X)\text{-almost surely}\}. \tag{5.16}$$

As above we have $\beta_* \ge \beta_2$, and β_2 can be computed in terms of the Green function of the difference random walk $\tilde{X} := X - X'$. $\tilde{X}$ jumps at rate 2κ according to the symmetrised transition matrix $q_{xy} := (p_{xy} + p_{yx})/2$. Obviously $V =_d \int_0^\infty \mathbf{1}(\tilde{X}(s) = 0)\, ds =_d \sum_{k=1}^B Z_i$, where Z_i are independent and Exp(2κ)-distributed, B is geometric with parameter q_{esc}, and finally q_{esc} is the probability that a discrete-time random walk with transition matrix q does not return to the origin. Hence $\mathbb{E}\exp(\beta V) = \sum_{k=1}^\infty q_{\text{esc}}(1 - q_{\text{esc}})^{k-1} (\mathbb{E}\exp(\beta Z))^k < \infty$ iff $\mathbb{E}\exp(\beta Z) < (1 - q_{\text{esc}})^{-1}$. As $\mathbb{E}\exp(\beta Z) = 2\kappa/(2\kappa - \beta)$ for $0 < \beta < 2\kappa$ we see that

$$\beta_2 = 2\kappa q_{\text{esc}} = \frac{1}{(2\kappa)^{-1}\sum_{n=0}^\infty q_{00}^n} = \frac{1}{\int_0^\infty \mathbb{P}_0(\tilde{X}(s) = 0)\, ds}. \tag{5.17}$$

In order to get a feeling for β_* we approximate by discrete-time random walks: Let ξ_N and ξ'_N be two discrete-time random walks with transition matrix given by $p_{xy}^{(N)} := (1 - \kappa/N)\delta_{xy} + (\kappa/N)p_{xy}$ (for $N \ge \kappa$). Then $(\xi_N([Nt])_{t\ge 0} \Rightarrow (X(t))_{t \ge 0}$ as $N \to \infty$ and we may expect that

$$\mathbb{E}\left[\exp\big(\beta \int_0^\infty \mathbf{1}(X(s) = X'(s))\, ds\big)\Big| X\right] \approx \mathbb{E}\left[\exp\big(\beta \frac{1}{N}\sum_{k=1}^\infty \mathbf{1}(\xi_N(k) = \xi'_N(k))\big)\Big|\xi_N\right].$$

By Theorem 5 the righthand side is finite for $\exp(\beta/N) < \alpha_*(p^{(N)})$, and hence the N-th threshold value is

$$\begin{aligned} \beta_*^{(N)} \quad &= \quad N \log\left(1 + \left(\sum_{n\ge 1} \exp\Big(-H\big((p^{(N)})^n\big)\Big)\right)^{-1}\right) \\ &\longrightarrow \left(\int_0^\infty \exp\big(-H(\mathbb{P}_0(X(s) \in \cdot))\big) ds\right)^{-1} \qquad \text{as } N \to \infty \end{aligned}$$

(Observe that $N\log(1 + x_N) \sim N x_N$ if $x_N \to 0$).

This approximation indeed gives the right answer:

Theorem 7 *Assume that $p_{x,y}(t)$ satisfies*

$$\sup_{t \ge 0, x \in \mathbb{Z}^d} \frac{p_x(t)}{\tilde{p}_0(t)} < \infty. \tag{5.18}$$

Then we have

$$\beta_* = \left(\int_0^\infty \exp\big(-H(\mathbb{P}_0(X(s) \in \cdot))\big) ds\right)^{-1},$$

where β_* is defined in (5.16) and $H(\mu) = -\sum_{x\in\mathbb{Z}^d} \mu_x \log \mu_x$ is the entropy of $\mu \in \mathcal{M}_1(\mathbb{Z}^d)$.

Proof. It is unclear how to make the above approximation argument itself rigorous but the proof of Theorem 5 can be adapted to continuous time: Again we convert the question of existence of conditional exponential moments into questions about the growth of conditional moments by writing

$$\begin{aligned}
&\mathbb{E}\Big[\exp\big(\beta\int_0^\infty \mathbf{1}(X(s)=X'(s))\,ds\big)|X\Big]-1\\
&\quad=\sum_{n=1}^\infty \frac{\beta^n}{n!}\mathbb{E}\Big[\int_0^\infty\cdots\int_0^\infty \mathbf{1}(X(s_1)=X'(s_1),\dots,X(s_n)=X'(s_n))\,ds_1\dots ds_n|X\Big]\\
&\quad=\sum_{n=1}^\infty \beta^n \underset{s_1\le s_2\le\cdots\le s_n}{\int\cdots\int} \mathbb{P}(X(s_1)=X'(s_1),\dots,X(s_n)=X'(s_n)|X)\,ds_1\dots ds_n\\
&\quad=\sum_{n=1}^\infty \beta^n F_n(X),
\end{aligned}$$

where $F_n(X)$ is defined to be the n-fold integral in the last but one line. Thus we have to compute (the almost sure limit of) $(1/n)\log F_n(X)$. Denote by $\pi_x(t) := \mathbb{P}_0(X(t)=x)$ the transition probability of X and by $\tilde\pi(t) := \mathbb{P}_{(0,0)}(X(t)=X'(t))$ the return probabilities for the difference walk $\tilde X$. Let $\tilde G := \int_0^\infty \tilde\pi(s)\,ds$. As (we put $s_0 := 0$)

$$\mathbb{P}(X(s_1)=X'(s_1),\dots,X(s_n)=X'(s_n)|X) = \prod_{j=1}^n \pi_{X(s_j)-X(s_{j-1})}(s_j-s_{j-1})$$

we can write

$$\begin{aligned}
F_n(X) &= \tilde G^n \int_0^\infty\cdots\int_0^\infty \Big(\prod_{j=1}^n \frac{\tilde\pi(r_j)}{\tilde G}\Big)\\
&\qquad\qquad \times\prod_{i=1}^n \frac{\pi_{X(r_1+\cdots+r_{i-1}+r_i)-X(r_1+\cdots+r_{i-1})}(r_i)}{\tilde\pi(r_i)}dr_1\dots dr_n\\
&= \tilde G^n \mathbb{E}\Big[\prod_{i=1}^n \frac{\pi_{\Delta_i}(R_i)}{\tilde\pi(R_i)}|X\Big]
\end{aligned}$$

where $R_1, R_2, \dots$ are i.i.d. with density $\mathbf{1}_{R_+}(r)\tilde\pi(r)/\tilde G$ and independent of X, $\Delta_i := X(R_1+\cdots+R_{i-1}+R_i) - X(R_1+\cdots+R_{i-1})$. We can also formulate this in terms of the empirical distribution of the increments of the path X observed between the renewal points given by the sums of the R_is: Let

$$\mathcal{Y}_n := \frac{1}{n}\sum_{i=1}^n \delta_{\big(X(R_1+\cdots+R_{i-1}+R_i)-X(R_1+\cdots+R_{i-1}),R_i\big)}$$

with values in $\mathcal{M}_1(\mathbb{Z}^d\times\mathbb{R}_+)$, and $f:\mathbb{Z}^d\times\mathbb{R}_+\to\mathbb{R}_+$ given by $f(y,t) := \pi_y(t)/\tilde\pi_0(t)$. Then we can reformulate

$$F_n(X) = \tilde G^n\,\mathbb{E}\left[\exp\Big(n\int \log f(y,t)\,d\mathcal{Y}_n(y,t)\Big)\Big|X\right]$$

and use large deviation theory as in the proof of Theorem 5. The function f is continuous and bounded from above by assumption (5.18), so we can use use Proposition 12 and

Varadhan's lemma to compute

$$
\begin{aligned}
&\lim_n \frac{1}{n} \log F_n(X) \\
&\quad = \log(\tilde{G}) + \sup_{\nu \in \mathcal{M}_1(\{\text{paths}\} \times \mathbb{R}_+)} \Big\{ \int \log f(y,t)\, d\nu(y,t) - \tilde{H}(\nu) \Big\} \\
&\quad = \log(\tilde{G}) + \sup_{\tilde{\nu} \in \mathcal{M}_1(\mathbb{R}_+)} \Big\{ \int_{\mathbb{R}_+} \mathbb{E}_0 \Big[\log f((X(s \wedge t))_{s \geq 0}, t) \Big] \tilde{\nu}(dt) - h(\tilde{\nu}; \tilde{\mu}) \Big\} \\
&\quad = \log(\tilde{G}) + \sup_{\tilde{\nu} \in \mathcal{M}_1(\mathbb{R}_+)} \left\{ \int_{\mathbb{R}_+} \sum_{x \in \mathbb{Z}^d} \pi_x(t) \log \frac{\pi_x(t)}{\tilde{\pi}(t)} \tilde{\nu}(dt) - h(\tilde{\nu}; \tilde{\mu}) \right\} \\
&\quad = \sup_{0 \leq g,\, \int g\, d\tilde{\mu} = 1} \left\{ \int_{\mathbb{R}_+} \sum_{x \in \mathbb{Z}^d} \pi_x(t) \log \frac{\pi_x(t)}{\tilde{\pi}(t)/\tilde{G}}\, g(t) \frac{\tilde{\pi}(t)}{\tilde{G}} dt - \int_{\mathbb{R}_+} g(t) \log g(t) \frac{\tilde{\pi}(t)}{\tilde{G}} dt \right\} \\
&\quad = \sup_{0 \leq g,\, \int g\, d\tilde{\mu} = 1} \Big\{ - \int_{\mathbb{R}_+} H(\pi_\cdot(t))\, g(t) \frac{\tilde{\pi}(t)}{\tilde{G}} dt - \int_{\mathbb{R}_+} \log \Big(g(t) \frac{\tilde{\pi}(t)}{\tilde{G}} \Big) g(t) \frac{\tilde{\pi}(t)}{\tilde{G}} dt \Big\} \\
&\quad = \sup_{0 \leq \varphi,\, \int \varphi(t)\, dt = 1} \Big\{ - \int_{\mathbb{R}_+} \varphi(t) \log \varphi(t) dt - \int_{\mathbb{R}_+} H(\pi_\cdot(t)) \varphi(t) dt \Big\} \\
&\quad = \sup_{0 \leq \varphi,\, \int \varphi(t)\, dt = 1} \Big\{ - \int_{\mathbb{R}_+} \varphi(t) \log \Big(\varphi(t) \exp\big(H(\pi_\cdot(t)) \big) \Big) dt \Big\} \\
&\quad = \log C_* - \inf_{0 \leq \varphi,\, \int \varphi(t)\, dt = 1} \left\{ \int_{\mathbb{R}_+} \varphi(t) \log \frac{\varphi(t)}{C_*^{-1} \exp\big(- H(\pi_\cdot(t)) \big)} \right\} \\
&\quad = \log C_* - \inf_{0 \leq \varphi,\, \int \varphi(t)\, dt = 1} h(\varphi(t) dt; \varphi_* dt) \\
&\quad = \log C_* - 0 = \log \int_0^\infty \exp\big(- H(\pi_\cdot(t)) \big) dt
\end{aligned}
$$

where $\varphi_*(t) = C_*^{-1} \exp\big(-H(\pi_\cdot(t))\big)$ is the obvious maximiser (and C_* is the normalizing constant to make it a probability density). Observe that $\beta_* = \exp\big(- \lim \frac{1}{n} \log F_n(X) \big)$ to conclude. □

Proposition 12 *Let X be a continuous-time random walk on $\mathbb{Z}^d$ starting from $X(0) = 0$, and let $R_1, R_2, \dots$ be i.i.d. non-negative random variables independent of X. Let $T_0 := 0$, $T_i := T_{i-1} + R_i$, $i = 1, 2, \dots$ be the renewal process constructed from the waiting times R_i. Define*

$$
\mathcal{Y}_n := \frac{1}{n} \sum_{i=1}^n \delta_{\big(X(T_i) - X(T_{i-1}), R_i\big)},
$$

the joint empirical distribution of the increments of X observed along the renewal times and the lengths of the corresponding renewal intervals.

The family $\big(\mathcal{L}(\mathcal{Y}_n | X)\big)_{n \in \mathbb{N}}$ almost surely satisfies a large deviation principle with rate n and rate function

$$
\tilde{H}(\nu) = \begin{cases} h(\tilde{\nu}; \tilde{\mu}) & \text{if } \nu = \int_0^\infty \big\{ \mathcal{L}\big(X(t)\big) \otimes \delta_t \big\} \tilde{\nu}(dt) \\ & \text{for some } \tilde{\nu} \in \mathcal{M}_1(\mathbb{R}_+) \text{ with } \tilde{\nu} \ll \tilde{\mu} \\ \infty & \text{otherwise,} \end{cases}
$$

where $\tilde{\mu}(dt)$ is the distribution of R_1 and $h(\tilde{\nu}; \tilde{\mu})$ is the relative entropy of $\tilde{\nu}$ with respect to $\tilde{\mu}$.

Proof. The arguments are very similar to the proof of Theorem 6, so we only state the replacements necessary to transfer it to the present continuous-time setting: Instead of $\tilde{E}$ we work with $\mathbb{Z}^d \times \mathbb{R}_+$, which is again a Polish space, and hence $\mathcal{M}_1(\mathbb{Z}^d \times \mathbb{R}_+)$ with the topology of weak convergence is also a Polish space (see e.g. [8], Lemma 3.2.2). The sequence $\big((X(T_i) - X(T_{i-1}), R_i\big)_{i=1,2,\dots}$ is i.i.d, and μ, the distribution of $\big((X(R_1) - X(0), R_1\big)$, is given by $\mu(x, dt) = p_t(x)\tilde{\mu}(dt)$. Thus by Sanov's theorem the unconditional distributions of $\mathcal{Y}_n$ satisfy a large deviation principle with rate n and rate function $H(\nu) = h(\nu; \mu)$. Set

$$\tilde{\mathcal{M}} := \big\{\nu \in \mathcal{M}_1(\mathbb{Z}^d \times \mathbb{R}_+) : \text{there is a } \tilde{\nu} \in \mathcal{M}_1(\mathbb{R}_+) \text{ such that } \nu(x, dt) = p_t(x)\tilde{\nu}(dt)\big\},$$

this is the relevant subset on which the rate function of the conditional laws concentrates. $\tilde{\mathcal{M}}$ is closed with respect to the weak topology. (In order to see this consider $(\nu_n) \subset \tilde{\mathcal{M}}$ with $\nu_n \stackrel{w}{\to} \nu$ as $n \to \infty$. Then there is $(\tilde{\nu}_n) \subset \mathcal{M}_1(\mathbb{R}_+)$ such that $\nu_n(x, dt) = p_t(x)\tilde{\nu}_n(dt)$. Note that $(\tilde{\nu}_n)_{n\in\mathbb{N}}$ is tight because $\liminf \tilde{\nu}_n([0, K)) = \liminf \nu_n(\mathbb{Z}^d \times [0, K)) \geq \nu(\mathbb{Z}^d \times [0, K)) \geq 1-\varepsilon$ for K large enough. Thus there is a subsequence (n_k) and a $\tilde{\nu}_\infty \in \mathcal{M}_1(\mathbb{R}_+)$ such that $\tilde{\nu}_n \stackrel{w}{\to} \tilde{\nu}_\infty$, hence $\nu = \lim_k \nu_{n_k}$ is of the form $\nu(x, dt) = p_t(x)\tilde{\nu}_\infty(dt)$ [and $\tilde{\nu}$ is of course uniquely determined by ν].)

With these replacements the proof of is a more or less verbatim copy of the proof of Theorem 6. $\square$

Bibliography

[1] René A. Carmona, Stanislav Molchanov, Parabolic Anderson problem and intermittency, *Mem. Amer. Math. Soc.* **108** (1994), no. 518.

[2] Mu Fa Chen, *From Markov chains to nonequilibrium particle systems*, World Scientific Publishing Co., River Edge, NJ, 1992.

[3] Francis Comets, Large deviation estimates for a conditional distribution. Application to random interaction Gibbs measures, *Probab. Theory Related Fields* **80** (1989), no. 3, 407–432.

[4] J. Theodore Cox, Klaus Fleischmann, Andreas Greven, Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Related Fields* **105** (1996), no. 4, 513–528.

[5] J. Theodore Cox, Andreas Greven, Ergodic theorems for infinite systems of locally interacting diffusions, *Annals of Probability* **22** (1994), no. 2, 833-853.

[6] Donald A. Dawson, Klaus Fleischmann, On spatially homogeneous branching processes in a random environment, *Math. Nachr.* **113** (1983), 249–257.

[7] Donald A. Dawson, Klaus Fleischmann, Critical dimension for a model of branching in a random medium, *Z. Wahrsch. Verw. Gebiete* **70** (1985), no. 3, 315–334.

[8] Jean-Dominique Deuschel, Daniel W. Stroock, *Large deviations*, Academic Press, Boston, 1989.

[9] Stewart N. Ethier, Thomas G. Kurtz, *Markov processes. Characterization and convergence*, John Wiley & Sons, New York, 1986.

[10] William Feller, *An introduction to probability theory and its applications, Vol. II.* John Wiley & Sons, New York, 1966.

[11] Jochen Geiger, Size-biased and conditioned random splitting trees, *Stoch. Proc. Appl.* **65** (1996), 187–207.

[12] Jochen Geiger, Elementary new proofs of classical limit theorems for Galton-Watson processes, *J. Appl. Prob.* **36** (1999), 301–309.

[13] Andreas Greven, A phase transition for the coupled branching process. I. The ergodic theory in the range of finite second moments. *Probab. Theory Related Fields* **87** (1991), no. 4, 417–458.

[14] Andreas Greven, On phase-transitions in spatial branching systems with interaction. *Stochastic models (Ottawa, ON, 1998)*, 173–204, CMS Conf. Proc., 26, Amer. Math. Soc., Providence, RI, 2000.

[15] Andreas Greven, Frank den Hollander, Longtime behaviour of interacting diffusions in a new universality class, Manuscript, 2002.

[16] Andreas Greven, Achim Klenke, Anton Wakolbinger, The longtime behavior of branching random walk in a catalytic medium. *Electron. J. Probab.* **4** (1999), no. 12.

[17] Frank den Hollander, *Large deviations*. Fields Institute Monographs, 14. American Mathematical Society, Providence, RI, 2000.

[18] Kiyosi Itō, *Stochastic Processes*, Lecture Notes Series, No. 16, Matematisk Institut, Aarhus Universitet, Aarhus, 1969.

[19] Olav Kallenberg, Stability of critical cluster fields. *Math. Nachr.* **77** (1977) 7–43.

[20] Olav Kallenberg, *Random measures*. Third edition. Akademie-Verlag, Berlin; Academic Press, Inc., London, 1983.

[21] Olav Kallenberg, *Foundations of modern probability*, Second edition, Springer-Verlag, New York, 2002.

[22] Claude Kipnis, Claudio Landim, *Scaling limits of interacting particle systems*, Springer-Verlag, New York, 1999.

[23] Thomas M. Liggett, *Interacting Particle Systems*, Springer-Verlag, New York, 1985.

[24] Thomas M. Liggett, Frank Spitzer, Ergodic theorems for coupled random walks and other systems with locally interacting components. *Z. Wahrsch. Verw. Gebiete* **56** (1981), no. 4, 443–468.

[25] Russell Lyons, Robin Pemantle, Yuval Peres, Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** (1995), no. 3, 1125–1138.

[26] Charles M. Newman, A. Larry Wright, An invariance principle for certain dependent sequences. *Ann. Probab.* **9** (1981), no. 4, 671–675.

[27] Kalyanapuram Rangachari Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York, 1967.

[28] Amnon Pazy, *Semigroups of linear operators and applications to partial differential equations*. Applied Mathematical Sciences, 44. Springer-Verlag, New York, 1983.

[29] Tokuzo Shiga, Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems. *Osaka J. Math.* **29** (1992), no. 4, 789–807.

Deutsche Zusammenfassung

Teilchensysteme mit lokal abhängiger Verzweigung: Langzeitverhalten, Genealogie und kritische Parameter

Wir betrachten Systeme unabhängiger Irrfahrer auf $\mathbb{Z}^d$ mit Verzweigung, wobei ein Individuum (oder Teilchen) bei seinem Tod eine zufällige Anzahl Nachkommen hinterläßt, im Mittel genau einen. Das Langzeitverhalten wird durch das Zusammenwirken zweier gegensätzlicher Kräfte bestimmt: Die zufällige Verzweigung erzeugt Fluktuationen in der räumlichen Verteilung der Teilchen, während die unabhängige Bewegung dazu neigt, die Teilchen relativ homogen im Raum zu verteilen und daher glättend wirkt. Welcher der beiden Mechanismen setzt sich auf lange Sicht durch? Wenn wir mit einer räumlich homogenen Teilchenkonfiguration beginnen, konvergiert das System dann gegen ein nicht-triviales Gleichgewicht, oder sind die Fluktuationen so stark, dass die Population lokal ausstirbt? Wenn sich das System langfristig stabilisiert, wird die Gleichgewichtspopulation immer noch endliche Varianz der lokalen Teilchenzahl aufweisen, oder können die Fluktuationen dazu führen, dass die zweiten Momente verloren gehen? Das sind die zentralen Fragen, denen sich die vorliegende Arbeit stellt.

Die Antworten sind wohlbekannt im klassischen Fall unabhängiger Verzweigung: Sofern die Kinderzahlen endliche Varianz haben, hängt alles davon ab, ob die Symmetrisierung der zugrundeliegenden Irrfahrt rekurrent oder transient ist. Ist sie rekurrent, so stirbt das System lokal aus. Im transienten Fall gibt es eine (natürlicherweise mit der erwarteten Teilchendichte parametrisierte) Schar von Gleichgewichten, und endliche zweite Momente bleiben im Limes erhalten.

Wir untersuchen den Fall lokaler Abhängigkeit im Verzweigungsmechanismus, in dem die Kinderzahlen verschiedener Individuen am selben Ort korreliert sind, sei es durch echte Interaktion oder einfach dadurch, dass sie durch diesselbe Umgebung beeinflußt werden. Diese alternativen Mechanismen zur Erzeugung lokaler Abhängigkeiten verkörpern wir in zwei Klassen von Modellen:

1) Systeme mit zustandsabhängiger Verzweigungsrate: Wenn an einen Ort k Teilchen sind, kommt es mit Rate $\sigma(k)$ zu einem Verzweigungsereignis. Dann stirbt ein Teilchen an diesem Ort und hinterläßt eine zufällige Anzahl Nachkommen gemäß einer vorgegebenen Verteilung ν mit Erwartungswert 1 und endlicher Varianz. Der klassische Fall unabhängiger Verzweigung ist durch die Wahl einer linearen Funktion σ gegeben.

2) Verzweigungssysteme in zufälliger Umgebung: Die Kinderzahlverteilungen $Q(x, n)$ am Ort x in Generation n sind unabhängige Kopien einer zufälligen Nachkommensverteilung Q. Wir nehmen an, dass der Erwartungswert $m_1(Q)$ von Q Mittelwert 1 und endliche Varianz hat. Gegeben die $Q(z, j)$, $z \in \mathbb{Z}^d, j \in \mathbb{N}$, verhalten sich alle Individuen unabhängig, Teilchen am Ort x in Generation n haben eine zufällige Anzahl Nachkommen mit

Verteilung $Q(x,n)$. Die Abhängigkeit zwischen verschiedenen Teilchen entsteht hier also indirekt durch den Einfluß des zufälligen Mediums. Wir bemerken, dass die Annahme eines unabhängigen Raum-Zeit-Feldes von zufälligen Kinderzahlverteilungen natürlicherweise diskrete Zeit erzwingt, da in zeitstetigen Modellen generisch niemals zwei Teilchen exakt zur gleichen Zeit verzweigen und so die Abhängigkeit, die wir studieren möchten, verloren ginge.

Für Populationen mit zustandsabhängiger Verzweigungsrate erweisen sich die beiden folgenden Extremfälle als besonders interessant:

a) $\sigma(k) = C_\sigma k^2$

b) $\sigma(k) = 1_{\{k=1\}}$.

Die Verzweigungsrate im Fall b) ist extrem niedrig, Teilchen können nur verzweigen, solange sie alleine an einem Ort sind. Auf der anderen Seite ist das quadratische Wachstum von σ im Fall a) ein Grenzfall in dem Sinn, dass ein noch stärkeres Wachstum zur Explosion der zweiten Momente in endlicher Zeit führt.

Nehmen wir an, dass die symmetrisierte Individualbewegung transient und dass die Startbedingung in Verteilung räumlich homogen ist mit endlicher Varianz der lokalen Teilchenanzahl. Dann hängt das Langzeitverhalten von drei Parametern b, β_2 und β_* ab. b mißt die Variabilität des Verzweigungsmechanismus: Im Fall 1a) ist $b := C_\sigma \mathrm{Var}(\nu)$, im Fall 2) ist $b := \mathrm{Var}(m_1(Q))$. Die Stärke der Irrfahrtsbewegung drückt sich folgendermaßen in β_2 und β_* aus: Sei V die Kollisionszeit zweier unabhängiger Kopien S und S' der zugrundeliegenden Irrfahrt, $\mathbb{E}[e^{\beta V}]$ bzw. $\mathbb{E}[(1+\beta)^V]$ das exponentielle Moment der Ordnung β von V im Fall 1a) bzw. 2). β_2 ist das Supremum über alle Ordnungen β, für die das exponentielle Moment von V endlich ist; β_* ist das Supremum über alle β, für die das entsprechende exponentielle Moment von V, bedingt auf einen der Irrfahrtspfade, endlich ist.

Wir zeigen, dass sich β_2 jeweils mittels der Greenfunktion der Differenzirrfahrt $S-S'$ berechnen läßt, und dass im Fall 2) unter schwachen Voraussetzungen an die Irrfahrt β_*^{-1} gegeben ist als $\sum_{n\geq 1}\exp(-H_n)$, wo H_n die Entropie der n-Schritt Übergangsverteilung der Irrfahrt ist. Darüberhinaus vermuten wir, dass β_* im Fall 1a) durch einen entsprechenden Ausdruck gegeben ist, in dem $n \in \mathbb{N}$ durch $t \in \mathbb{R}_+$ und die Summe durch ein Integral ersetzt wird. Wir skizzieren ein analoges Programm zum Beweis dieser Vermutung. Insbesondere folgt aus diesen Darstellungen, dass $\beta_2 < \beta_*$.

Das Langzeitverhalten in den Fällen 1b) und 2) wird folgendermaßen durch b, β_2 und β_* bestimmt:

(i) Für $b < \beta_2$ sind die zweiten Momente gleichmäßig beschränkt, und das System konvergiert gegen ein Gleichgewicht, das (nur) von der Anfangsintensität abhängt.

(ii) Für $\beta_2 \leq b < \beta_*$ wachsen die zweiten Momente mit der Zeit über alle Schranken, sogar exponentiell wenn $\beta_2 < b$, aber das System konvergiert immer noch gegen ein Gleichgewicht, das die Anfangsintensität erhält. Im Gegensatz zu „klassischen“ unabhängigen räumlichen Verzweigungssystemen zeigt hier insbesondere jedes nicht-triviale Gleichgewicht unendliche Varianz der lokalen Teilchenanzhl.

Zum Beweis von (i) benutzen wir Kopplungsargumente und Darstellungen von (gemischten) Momenten solcher räumlicher Verzweigungssysteme mittels Funktionalen unabhängiger Irrfahrten. Der Beweis von (ii) benutzt eine stochastische Darstellung der „Verwandten“ eines zufällig aus der Population herausgegriffenen Individuums mittels seines genealogischen Baums. Die Idee, das Langzeitverhalten mittels räumlich eingebetteter, lokal größenverzerrrter genealogischer Bäume zu analysieren geht im Fall unabhängiger Verzweigung auf Kallenberg (1977) zurück. Wir konstruieren entsprechende stochastische Darstellungen für die lokal größenverzerrten Bäume (oder „Kallenberg-Bäume“) in den von uns betrachteten Modellen 1) und 2).

Eine verwandte Konstruktion für ein sehr spezielles Verzweigungssystem in zufälliger Umgebung in kontinuierlicher Zeit, der sogenannte “coupled branching process”, findet sich in einer Arbeit von Greven (2000). Er betrachtet einen klassischen superkritischen räumlichen Verzweigungsprozess, bei dem an jedem Ort ein Poissonprozess von „Katastrophen“ ab und zu die lokale Population auslöscht. Die vorliegende Arbeit ist zum Teil motiviert durch den Versuch, die Argumente in der Arbeit von Greven besser zu verstehen und eine allgemeinere Struktur hinter ihnen zu erkennen. Während letzteres geglückt zu sein scheint, war ersterem nur partieller Erfolg beschieden. Insbesondere bleibt die große offene Frage, ob $b > \beta_*$ lokales Aussterben erzwingt.

Eine weitere Quelle der Inspiration waren Untersuchungen, die Greven und den Hollander derzeit über das parabolische Anderson-Modell durchführen. Das ist in unserem Kontext ein System $(X_x(t))_{t\geq 0, x\in\mathbb{Z}^d}$ linear gekoppelter Diffusionen, Lösung von

$$dX_x(t) = \sum_y p_{y-x}(X_y(t) - X_x(t))dt + b\, dB_x(t),$$

wo B_x unabhänginge Standard-Brownbewegungen sind. Shiga (1992) konnte mit Methoden der Stochastischen Analyis zeigen, dass das parabolische Anderson-Modell lokal ausstirbt, sofern b genügend groß ist, konnte aber den genauen Schwellwert nicht bestimmen. Greven und den Hollander vermuten, dass es gerade β_* ist.

Die Parallelen zwischen Systemen mit quadratisch von der Teilchenzahl abhängiger Verzweigungsrate einerseits und Verzweigungssystemen in zufälliger Umgebung andererseits sind vielleicht weniger überraschend, wenn man sich vergegenwärtigt, dass beide die Eigenschaft haben, dass die Varianz der lokalen Populationsänderung, gegeben die derzeitige Teilchenzahl, proportional zum *Quadrat* der derzeitigen lokalen Dichte ist. Daher ist es nicht unplausibel, dass beide in derselben „Universalitätsklasse“ liegen wie das parabolische Anderson-Modell.

Für Systeme mit allgemeiner zustandsabhängiger Verzweigungsratenratenfunktion $\sigma(\cdot)$ treffen wir die Annahme, dass entweder σ nicht-fallend ist oder die Kinderzahlverteilung binär, d.h. es gibt stets entweder null oder zwei Nachkommen. Diese Annahme garantiert, dass das System “attraktiv” (in der Nomenklatur interagierender Teilchensysteme) ist. Wir zeigen, dass für transiente symmetrisierte Individualbewegung und $\limsup \sigma(k)/k^2 < \beta_2/\mathrm{Var}(\nu)$ das System sich gewissermaßen ähnlich einem klassischen unabhängigen Verzweigungssystem verhält: Es gibt eine einparametrige Schar von Gleichgewichten, alle mit endlichen zweiten Momenten, und wir beschreiben ihre Anziehungsbereiche. Diese Ergebnisse liegen parallel zu entsprechenden Resultaten von Cox und Greven (1994) über interagierende Diffusionen, die man als Skalierungslimes von Systemen vom Typ 1) erhalten kann.

Darüberhinaus zeigen wir für Systeme vom Typ 1) einen Vergleichssatz, der besagt, dass für zwei solche Systeme ξ und ξ', die sich nur in der Verzweigungsratenfunktion $\sigma(\cdot) \geq \sigma'(\cdot)$ unterscheiden, die Erwartungswerte gewisser konvexer Funktionen der Teilchenkonfiguration in derselben Weise geordnet sind. Dieses Ergebnis ist wiederum ein “Teilchen-Kollege” eines entsprechenden Resultats über interagierende Diffusionen, das Cox, Fleischmann und Greven 1996 bewiesen haben. Eine unmittelbare Anwendung des Vergleichsatzes auf die Laplace-Transformierten zeigt, dass ein System um so leichter ausstirbt, je größer seine Verzweigungsratenfunktion ist. Das ist auch intuitiv einsichtig, da größere Fluktuationen das System schneller an den absorbierenden Rand, nämlich den Leerzustand, befördern können.

Schließlich wenden wir uns auch Systemen mit zustandsabhängiger Verzweigungsrate und *rekurrenter* Individualbewegung zu, hier erhalten wir aber nur recht partielle Resultate. Angesichts des Vergleichssatzes und des Verhaltens unabhängiger räumlicher Verzweigungssysyeme ist klar, das eine solche Population stets ausstirbt, sofern σ mindestens linear wächst. Andererseits bleibt die Frage: Kann ein starkes „Herunterregulieren" der Verzeigungsrate die „Klumpigkeit" der klassischen Systeme verhindern und so langfristiges Überleben ermöglichen? Angesichts des Vergleichsresultats bietet sich hier auf natürliche Weise der Fall 1b) zum Studium an. Wenn diese sogenannten "lonely branchers" lokal aussterben, dann lautet die Antwort generisch „nein". Die Frage nach dem Langzeitverhalten von Fall 1b) für $\mathbb{Z}$ und $\mathbb{Z}^2$ wurde ursprünglich von Ted Cox (1998, private Mitteilung) gestellt, sie ist bis heute offen. Ein denkbarer Ansatz liegt im Studium der zugehörigen Kallenberg-Bäume: Das System stirbt genau dann aus, wenn der entsprechende lokal größenverzerrte genealogische Baum lokal über alle Grenzen wächst. Da wir diese Frage nicht entscheiden konnten, haben wir stattdessen ein Karikaturmodell studiert, bei dem der Stamm des Baumes sich nicht bewegt und die Seitenlinien nicht weiter verzweigen. Wir klären das Langzeitverhalten dieser Systeme mit selbstblockierender Immigration. Der Name rührt daher, dass sie aus unabhängigen Irrfahrern bestehen, zu denen sich zu den Sprungzeitpunkten eines Poissonprozesses an einen festen Raumpunkt, einer „Quelle", jeweils ein weiteres Teilchen hinzu gesellt, aber nur, wenn die Quelle derzeit nicht gerade besetzt ist. Es stellt sich heraus, dass die lokale Teilchenzahl in Wahrscheinlichkeit über alle Schranken wächst genau dann, wenn die Individualbewegung rekurrent ist. Wir vermuten daher, dass auch die wahren Kallenberg-Bäume klumpen und und die lonely branchers folglich aussterben.

Wir erhalten einige Resultate bezüglich des quantitativen Langzeitverhaltens von Systemen mit selbstblockierender Immigration mit rekurrenter Individualbewegung, d.h. wir beantworten (teilweise) die Frage, wie schnell ein solches System wächst, wenn man es aus der leeren Konfiguration startet. Für positiv rekurrente Individualbewegung zeigt sich, dass die Gesamtanzahl an Teilchen logarithmisch in der Zeit wächst. Für den interessanten Fall der (nullrekurrenten) gewöhnlichen Irrfahrt auf $\mathbb{Z}$ leiten wir mittels Ideen aus der Theorie der hydrodynamischen Limiten eine „effektive Gleichung" für die Teilchendichte her, deren Langzeitasymptotik wir analysieren. Dies führt zu der Vorhersage, dass die Gesamtanzahl in diesem Fall wie $C\sqrt{t}\log t$ wachsen sollte. Mittels der der aus der Theorie hydrodynamischen Limiten bekannten Methode der relativen Entropie können wir zeigen, dass die Vorhersage mindestens die richtige t-Potenz beschreibt, nämlich dass das System nicht schneller wachsen kann als $t^{1/2+\epsilon}$. Diese Wachstumseigenschaften zusammen mit der „Invertierung des Karikatur-Schritts" würden eine interessante Eigenschaft der eindimensionalen lonely brancher aufzeigen: Die Familie eines zur Zeit t lebenden Individuums wäre von der Größenordnung $\sqrt{t}\log t$, nicht von der Ordnung t wie im Fall unabhängigen Verzweigens.

Lebenslauf

Matthias Birkner

geboren am 10.12.1971 in Frankfurt am Main

1978-91	Grundschule und Gymnasium in Frankfurt am Main
1991	Abitur, Carl-Schurz-Gymnasium Frankfurt
1991-92	Zivildienst, Universitätsklinikum Frankfurt
1992-98	Studium der Mathematik mit Nebenfach Informatik an der Johann Wolfgang Goethe-Universität Frankfurt
April 1998	Diplom in Mathematik mit Nebenfach Informatik, Diplomarbeit „Ein Vergleich des Langzeitverhaltens von Wählermodell und frei verzweigenden Markow-Ketten" betreut von Prof. Wakolbinger, Zweitgutachter Prof. Kersting. Diplomprüfung bei Prof. Wakolbinger, Prof. Bliedtner und Prof. Schnitger (Nebenfach).
Oktober 1998-	Wissenschaftlicher Mitarbeiter im Institut für Stochastik und Mathematische Informatik am FB Mathematik der Universität Frankfurt