Extinction versus unbounded growth

Habilitation Thesis of the University Erlangen-Nürnberg

Jan M. Swart

ÚTIA
Pod vodárenskou věží 4
18208 Praha 8
Czech Republic
e-mail: swart@utia.cas.cz

January 31, 2007

Abstract Certain Markov processes, or deterministic evolution equations, have the property that they are dual to a stochastic process that exhibits *extinction versus unbounded growth*, i.e., the total mass in such a process either becomes zero, or grows without bounds as time tends to infinity. If this is the case, then this phenomenon can often be used to determine the invariant measures, or fixed points, of the process originally under consideration, and to study convergence to equilibrium. This principle, which has been known since early work on multitype branching processes, is here demonstrated on three new examples with applications in the theory of interacting particle systems.

Contents

Chapter 1

Introduction

1.1 Interacting particle systems

This habilitation thesis treats three subjects from *probability theory*, and more precisely, from the field of *interacting particle systems*. The binding element is a common technique used to study these subjects, which gives the title to this thesis, which finds its origin in multitype *branching theory*, and which is applied here both to branching processes and to processes which do not have the branching property, but still are in some ways similar to branching processes, although in other aspects of their behavior they are completely different. In this introductory section, we zoom out a bit more than is usual in a research paper, and take a look at the whole area of probability theory, and the fields of interacting particle systems and branching theory in particular, to see how they arose historically and how they are related.

Probability theory established itself as a mathematical discipline relatively late in history. Its origins are often traced back to an exchange of letters between Pascal and Fermat in the mid-17th century [Apo69], although some mention Cardano, one century earlier. The theory was not put on a firm axiomatic basis until the monograph by Kolmogorov in 1933 [Kol33], who based it on abstract measure theory, which had been developed in the preceding decades following the work of Lebesgue at the turn of the century. Because of these foundations, some authors claim that probability theory is a subfield of measure theory. Although there are measures all over the place, this is probably as justified as saying that algebra is a subfield of linear algebra.

When one tries to look for reasons why probability theory rose so late (why, for example, did the Greeks show no interest?), one is reminded of Einstein's remark 'Gott würfelt nicht' (God doesn't gamble). Even today, many people, including some mathematicians, associate mathematics primarely with beautiful structures that are entirely fixed, like a Penrose tiling, while an infinite random structure of the type that occurs in percolation theory evokes a certain disdain: 'Why, that can be anything!'. Actually, it can't.

The reason is that once random structures get large, many events tend to get extremely improbable, until in the limit, for infinite systems, their probability is actually zero. The example that everybody knows are the laws of large numbers, which pertain to sums of independent identically distributed random variables. Closely related to this is the central limit theorem, which describes exactly how much randomness is left in the limit, and what the limit distribution is. Once a colleague asked what I was just working on. After hearing my explanation, his reaction was: so you are trying to prove a sort of central limit theorem? The answer is both yes and no.

Indeed, most of probability theory seems to be occupied with proving that certain things are certain in the limit that the system size, or time, or both tend to infinity, and that other things have a limit law.[1] Yet, the methods needed to prove these limit statements are in general completely different from those used in the case of independent random variables. The independent case being well-understood, probabilists nowadays investigate systems of highly dependent components. And while there is just one way in which things can be independent, there are many ways in which things can depend on each other.

Seen from this point of view, the "theory of interacting particle systems" sounds like the natural culmination point of all of probability theory. That is not quite true. In fact, the classical book by Liggett called 'Interacting Particle Systems' [Lig85] was translated into Russian as 'Markovskije Processy s Lokalnym Vzaimodejstvijem' (Markov Processes with Local Interaction), which captures the subject more precisely. Interacting particle systems are always situated in space, which is often $\mathbb{Z}^d$, sometimes $\mathbb{R}^d$, and sometimes another discrete or continuous structure that is in some way translation invariant. At each point in this space, there is some local Markov process going on, that is inherently random, and interacts with the Markov processes surrounding it. Although this interaction is only local, in the long run information can spread arbitrarily far, and therefore it is the long-time behavior of the process that is usually of interest.

This description of interacting particle systems excludes many other dependent systems, such as random walks in random environment, self-enforced and self-avoiding random walks, cellular automata and other deterministic evolutions, random matrices, and percolation theory, although many of these topics have close links with interacting particle systems. It also excludes, unrighteously, interacting particle systems in quantum probability. And, finally, it excludes other active areas of probabilistic research, such as abstract theory of Markov processes and semigroups, stochastic evolution equations, stochastic analysis, and more.

The origin of the field of interacting particle systems lies in 19-th century physics, when scientists like Bolzmann, Van der Waals, and others started to look for the molecular basis of thermodynamics. Thus, the original motivation was to study particles moving around in $\mathbb{R}^3$ according to the deterministic rules of classical Hamiltonian dynamics, or, later, its quantummechanical counterpart, which in a sense is both deterministic and inherently random. The mathematical problems arising from continuous space and deterministic motion being too difficult, people turned to models on lattices, that moreover have a local source of randomness. This class of models is still extremely rich, and apart from their original physical motivation, it was found that models of this type can be used to model many other interesting phenomena in a variety of applications in, for example, biology, sociology, and random network theory. Of the four classical models from [Lig85], namely the Ising model, voter model, contact process,

[1] I have to add a caveat here for statisticians, who are sometimes treated as probabilists, and sometimes as a species of their own, who from a practical point of view also have a lively interest in small samples, and, generally speaking, seem to be more interested in doing things and managing things, while the probabilist sensu strictu just sits down and tries to understand.

and exclusion process, only the first and last have a clear physical motivation.

As a mathematical discipline, the field of interacting particle systems started around 1970. Again, compared to other branches of mathematics, this is very recent. This time, the reasons lie probably not only in a lack of interest (after all, the physical problems had been around for a century by that time) but also in the inherent difficulty of the subject. Certain special results date back further, to the mid 40ies; this includes work on multitype branching processes, percolation, and the famous Onsager solution of the 2-dimensional equilibrium Ising model. Gradually, people had to get used to the fact that interacting particle systems rarely allow for explicit solutions, and that very little can be said about them in general. Rather, even the simplest-looking among them required the development of new tools suited exactly for them, and many naive questions remained open for many years.

The systems of interest (interacting particle systems) and the main questions (limit laws for large system sizes and large times) being defined now, we can focus on some more specific topics. The first topic we would like to mention, which motivates much of the work done in the field, is that of *phase transitions*. Originally referring to the phenomenon that certain substances (as a general rule with exceptions: pure chemical substances) can either be in a gaseous, fluid, or solid phase, and change abruptly between these phases as the temperature or pressure pass a certain point, the concept has subsequently been generalized to include more phases (e.g. graphite versus diamond) and then to describe the general phenomenon that many-particle systems may drastically change their behavior when certain parameters pass certain tresholds, called *critical points*.

Phase transitions are a central topic for a number of reasons. First of all, since finite systems running for a finite time generally depend continuously on their parameters, mathematically ideal phase transitions occur only in the limit that the system size, and time, are sent to infinity, and therefore are the typical sort of phenomenon that justifies the study of large or infinite systems. Second, detailed information about them is often hard to get, since they are out of reach of most expansion techniques that tell us something about very high or low values of our parameters. In other words, phase transitions are difficult, and therefore prestigious. The third and most important reason is probably the belief, supported by nonrigorous theory developed by theoretical physicists, that phase transitions are highly universal. Thus, different interacting particle systems may have the 'same' phase transition. Although the exact parameter values where this phase transition takes place may differ from one model to the other, zooming in on these phase transitions, and at the same time zooming out in space (and time, if we are not in equilibrium) should always yield roughly the same picture. This can for example be seen from the *critical exponents* of these phase transitions, which describe how certain quantities behave according to a certain power law as the critical point is approached. The classical paper in physics on this topic is [WK74].

Trying to prove results about *critical phenonema* that take place at, or in the immediate vicinity of the critical points, in particular, the calculation of critical exponents, has been a big aim behind much work done on interacting particle systems. Progress has been slow. In a number of cases, expansion techniques, such as the lace expansion, have been used to show that certain systems have 'trivial' exponents, that are the same as those for other, noninteracting systems. Recently, important progress has been made on critical exponents for two-dimensional systems having conformally invariant scaling limits. The key object in this work is the Stochastic Loewner Equation [Law05]. Apart from these two cases (the 'trivial' critical exponents and those from conformal field theory) there is still little process.

Where, in all of this, is the present habilitation thesis situated? No critical exponents will be calculated in what follows, but we will see critical phenomena, and even some universality. In any case, there will be phase transitions around, and we will prove limit laws as time and system size are sent to infinity. A repeating theme in the proofs will be the exploitation of the simple observation that in certain particle systems, the number of particles either becomes zero, or tends to infinity. As far as I am aware off, this idea was first used in multitype branching theory.

The theory of branching processes started with a paper by Galton and Watson in 1874 [WG74], who studied the problem of the extinction of noble names. The problem drew new interest with the rise of probability theory in the 30-ies and with the study of nuclear chain reactions, which led to the study of multitype processes. It was only in the mid-70-ies, when people started to consider $\mathbb{Z}^d$ as the space of types, that the first branching processes were studied that might truly be called interacting particle systems. Even as such, they hardly deserve the name, since they consist of particles independently hopping around on a lattice, that moreover independently of each other split into more particles or die. The only way in which dependencies arise, which make the model interesting, is through the fact that certain 'families' of particles all descend from one and the same 'ancestor'. Basic questions about their ergodic behavior were solved by Kallenberg [Kal77] using his famous 'backward tree technique'. We will use this technique in Section 2.9.2. It is moreover closely linked to the work in Chapter 4 of this thesis. The main technique that unites all chapters, however, is the use of 'extinction versus unbounded growth', as will be explained in the next section.

1.2 Extinction versus unbounded growth

Certain Markov processes, or deterministic evolution equations, have the property that they are dual to a stochastic process that exhibits *extinction versus unbounded growth*, i.e., the total mass in such a process either becomes zero, or grows without bounds as time tends to infinity. If this is the case, then this phenomenon can often be used to determine the invariant measures, or fixed points, of the process originally under consideration, and to study convergence to equilibrium. In this section, we demonstrate this principle, in the historicaly correct order, first on multitype branching processes, and then on the contact process.

1.2.1 Extinction versus unbounded growth in branching theory

Consider a collection of particles of n different types. Assume that each particle of type $i \in \{1,\ldots,n\}$ gives with *birth rate* b_{ij} birth to a particle of type $j \in \{1,\ldots,n\}$, and dies with *death rate* d_i. We will assume that $b_{ij} > 0$ and $d_i > 0$ for all i,j. Let $Y_t(i)$ denote the number of particles of type i at time $t \geq 0$. Then $Y = (Y_t)_{t\geq 0}$ is a Markov process in $\mathbb{N}^n$, which in the usual terminology is called a *continuous-time multitype binary branching process*. We write P^y for the law of Y started in $Y_0 = y$ and denote expectation with respect to P^y by E^y. It is well-known that

$$E^y\Big[\prod_{i=1}^n(1-u_0(i))^{Y_t(i)}\Big]=\prod_{i=1}^n(1-u_t(i))^{y(i)}\qquad(t\geq 0),\tag{1.2.1}$$

whenever $u_t=(u_t(1),\ldots,u_t(n))$ is a $[0,1]^n$-valued solution to the system of differential equations

$$\tfrac{\partial}{\partial t}u_t(i)=\sum_{j=1}^n b_{ij}u_t(j)(1-u_t(i))-d_iu_t(i)\qquad(t\geq 0,\ i\in\{1,\ldots,n\}).\tag{1.2.2}$$

The map that gives $(1-u_t)$ as a function of $(1-u_0)$ and t is what is classically known as the *generating function* of the branching process Y (at time t). We prefer to work with u_t (and not $1-u_t$) since this will simplify formulas later on.

Formula (1.2.1) has a useful interpretation in terms of thinning. By definition, a *thinning* of a particle configuration $y\in\mathbb{N}^n$ with a vector $v\in[0,1]^n$ is the random particle configuration obtained from y in the following manner. Independently for each particle, we decide with probability $v(i)$ (depending on the type i of the particle) whether we will keep it; with the remaining probability $1-v(i)$ we throw this particle away. If we denote the thinned collection of particles resulting from this procedure by $\mathrm{Thin}_v(y)$, then the left-hand side of (1.2.1) is just the probability that the configuration $\mathrm{Thin}_{u_t}(Y_t)$ contains no particles. Since the right-hand side of (1.2.1) has a similar interpretation, we may rewrite (1.2.1) as

$$P^y[\mathrm{Thin}_{u_0}(Y_t)=0]=P[\mathrm{Thin}_{u_t}(y)=0]\qquad(t\geq 0).\tag{1.2.3}$$

The relation (1.2.1), or its rewrite (1.2.3), are an example of a *duality relation*, where the dual of the Markov process Y is in this case the deterministic process u.

Using this duality relation, we can deduce information about Y from u, and vice versa. To demonstrate this, we will show how the fact that the process Y exhibits extinction versus unbounded growth gives information about the fixed points of the n-dimensional differential equation (1.2.2).

It is not hard to see that

$$\tfrac{\partial}{\partial t}E[Y_t(i)]=\sum_{j=1}^n M_{ji}E[Y_t(j)]\qquad(t\geq 0),\tag{1.2.4}$$

where $M_{ji}=b_{ji}-\delta_{ij}d_i$ $(i,j=1,\ldots,n)$. Since by adding a constant multiple of the identity, we can make M into a matrix with strictly positive entries, it follows from the Perron-Frobenius theorem that M has a maximal eigenvalue, say λ, that corresponds to a positive right and left eigenvector, which are the only nonnegative eigenvectors. If $\lambda<0$, we say that the branching process Y is *subcritical*, if $\lambda=0$ we say that it is *critical*, and if $\lambda>0$ we say that it is *supercritical*. In the subcritical and critical cases, Y *dies out*, i.e.,

$$P^y\big[\exists t\geq 0\text{ s.t. }Y_s=0\ \forall s\geq t\big]=1\qquad(y\in\mathbb{N}^n).\tag{1.2.5}$$

(Note that since there is no spontaneous creation of particles, the zero configuration is a trap for the Markov process Y.) On the other hand, in the supercritical case, on which we focus from now on, Y *survives* with positive probability, i.e.,

$$P^y\big[Y_t \neq 0\ \forall t \geq 0\big] > 0 \qquad (y \in \mathbb{N}^n,\ y \neq 0). \tag{1.2.6}$$

Indeed, the probability in (1.2.6) is given by $1 - \prod_{i=1}^n (1 - p(i))^{y(i)}$, where

$$p(i) := P^{\delta_i}\big[Y_t \neq 0\ \forall t \geq 0\big] > 0 \qquad (i = 1, \ldots, n), \tag{1.2.7}$$

and δ_i denotes the particle configuration with just one particle of type i.

We claim that p is the only nonzero fixed point of the differential equation (1.2.2), and the limit point started from any nonzero initial condition. To prove this, we observe that Y exhibits *extinction versus unbounded growth*, in the following sense:

$$P^y\big[\exists t \geq 0 \text{ s.t. } Y_s = 0\ \forall s \geq t \quad \text{or} \quad \lim_{t\to\infty} |Y_t| = \infty\big] = 1 \qquad (y \in \mathbb{N}^n), \tag{1.2.8}$$

where $|y| := \sum_{i=1}^n y(i)$ denotes the total number of particles in a particle configuration $y \in \mathbb{N}^n$. Why does (1.2.8) hold? We will not give a formal proof here, but just explain the main idea. (For a more formal approach, see Lemma 2.80 below.) Since we are assuming that the death rates d_i are all positive, it is not hard to show that

$$\inf_{|y| \leq K} P^y\big[\exists t \geq 0 \text{ s.t. } Y_s = 0\ \forall s \geq t\big] > 0 \qquad (K \geq 0). \tag{1.2.9}$$

Indeed, if the process Y is started with no more than K particles, then there is a positive chance that all these particles die before they have a chance to branch, and therefore the probability that the process dies out can be estimated from below uniformly in all particle configurations with no more than K particles. Now imagine that the number of particles $|Y_t|$ is less than K at a (random) sequence of times tending to infinity. Then the process would infinitely often have a (uniformly) positive chance to die out in the next time interval of a certain length, and therefore it would eventually have to die out. Since this is true for any K, the only way for the process to escape extinction is to let the number of particles tend to infinity.

We now show how extinction versus unbounded growth (formula (1.2.8)) implies that any solution of (1.2.2) with $u_0 \neq 0$ satisfies

$$\lim_{t\to\infty} u_t = p, \tag{1.2.10}$$

where p is defined in (1.2.7). Note that $P[\mathrm{Thin}_v(\delta_i) \neq 0] = v(i)$ $(v \in [0,1]^n)$, and therefore, by (1.2.3),

$$u_t(i) = P^{\delta_i}[\mathrm{Thin}_{u_0}(Y_t) \neq 0] \qquad (t \geq 0,\ i = 1, \ldots, n). \tag{1.2.11}$$

Since we are assuming that $b_{ij} > 0$ for all i, j, it is easy to see from (1.2.11) that $u_0 \neq 0$ implies $u_t(i) > 0$ for all $i = 1, \ldots, n$ and $t > 0$, so by a restart argument we may without loss of generality assume that $u_0(i) > 0$ for all $i = 1, \ldots, n$.

Using (1.2.11) once more, and using extinction versus unbounded growth (formula (1.2.8)), we see that for large t there are up to an event with small probability only two situations to be considered. Either $Y_t = 0$, in which case $\mathrm{Thin}_{u_0}(Y_t) = 0$, or $|Y_t|$ is large, in which case, by the fact that $u_0(i) > 0$ for all i, we know that $\mathrm{Thin}_{u_0}(Y_t)$ is with large probability nonzero. Therefore, $P^{\delta_i}[\mathrm{Thin}_{u_0}(Y_t) \neq 0] \cong P[Y_t \neq 0]$ for large t, and taking the limit $t \to \infty$ in (1.2.11) we arrive at (1.2.10). This proves that p is the only nonzero fixed point of the differential equation (1.2.2), and the limit point started from any nonzero initial condition.

In a discrete time setting (but with much more general branching mechanisms), the result (1.2.10), including a proof based on extinction versus unbounded growth, can be found in Harris [Har63, Theorem II.7.2], who ascribes it to Everett and Ulam [EU48].

It is not hard to see that the positivy assumptions on the rates b_{ij} and d_i can be weakened considerably. In fact, it suffices if at least one of the d_i is nonzero, and if the b_{ij} are *irreducible*, in the sense that for each $i, j \in \{1, \ldots, n\}$, there exist $k_0, \ldots, k_m$ with $k_0 = i$, $k_m = j$, and $b_{k_{l-1},k_l} > 0$ for all $l = 1, \ldots, m$.

1.2.2 Extinction versus unbounded growth in the contact process

The standard, nearest neighbor d-dimensional *contact process* is a Markov process $\eta = (\eta_t)_{t\geq 0}$ taking values in the space of all subsets of $\mathbb{Z}^d$, with the following description. If $i \in \eta_t$, then we say that the *site* $i \in \mathbb{Z}^d$ is *infected* at time $t \geq 0$, otherwise such a site is called *healthy*. Infected sites become healthy with rate 1. Healthy sites become infected with infection rate λ times the number of neighboring infected sites. Here, we say that $i, j \in \mathbb{Z}^d$ are neighbors if $|i - j| = 1$.

It is useful to think about the contact process as a frustated branching process. Think of infected sites as being occupied by a particle. Then each particle tries with rate λ to give birth to a particle at each neighboring site. If, however, that site is already occupied by a particle, the birth fails.

Indeed, it is easy to see that $|\eta_t|$, the total number of infected sites, can be bounded from above by a binary branching process with branching rate $2d\lambda$ and death rate 1. In particular, if $\lambda \leq 1/(2d)$, this branching process is (sub)critical, and hence the contact process *dies out*. On the other hand, with considerably more effort, it is possible to show that for suffiently large λ, the contact process *survives* with positive probability, i.e.,

$$P^A[\eta_t \neq \emptyset \ \forall t \geq 0] > 0 \qquad (A \neq \emptyset). \tag{1.2.12}$$

It is easy to show that two contact processes $\eta, \tilde\eta$ with infection rates $\lambda, \tilde\lambda$ can be coupled such that $\eta_t \leq \tilde\eta_t$, so it follows that there exists a *critical infection rate* $0 < \lambda_{\rm c} < \infty$ such that the contact process dies out for $\lambda < \lambda_{\rm c}$ and survives (with positive probability) for $\lambda > \lambda_{\rm c}$. The question whether the contact process survives *at* $\lambda = \lambda_{\rm c}$ was open for almost 15 years; its solution by Bezuidenhout and Grimmett in [BG90] was a major milestone in the development of the theory of the contact process.

We will not touch this subject here, but rather show how the fact that the contact process exhibits extinction versus unbounded growth, together with self-duality, can be used to prove that if the contact process survives, then it has a unique nontrivial homogeneous invariant law. Here, we say that a probability law on the space of all subsets of $\mathbb{Z}^d$ is *nontrivial* if it gives zero probability to the empty set, and (spatially) *homogeneous* if it is invariant under translations.

It is well-known that the contact process is *self-dual*, in the following sense. Fix an infection rate λ, and for $A \subset \mathbb{Z}^d$, let η^A denote the contact process with this infection rate started in the initial state $\eta_0^A = A$. Then

$$P[\eta_t^A \cap B = \emptyset] = P[A \cap \eta_t^B = \emptyset] \qquad (t \geq 0,\ A, B \subset \mathbb{Z}^d). \tag{1.2.13}$$

Since the contact process is an attractive spin system, it follows from standard theory that it has an *upper invariant law* $\overline{\nu}$, which is the largest invariant law in the sense of stochastic ordering, and the limit law as $t \to \infty$ of the process started with all sites infected:

$$\mathcal{L}(\eta_t^{\mathbb{Z}^d}) \underset{t\to\infty}{\Longrightarrow} \overline{\nu}. \tag{1.2.14}$$

Using the self-duality (1.2.13) we can give a useful characterization of $\overline{\nu}$. Let $\eta_\infty^{\mathbb{Z}^d}$ be a random variable with law $\mathcal{L}(\eta_\infty^{\mathbb{Z}^d}) = \overline{\nu}$. Then

$$P[\eta_\infty^{\mathbb{Z}^d} \cap A = \emptyset] = \lim_{t\to\infty} P[\mathbb{Z}^d \cap \eta_t^A = \emptyset] = P[\exists t \geq 0 \text{ s.t. } \eta_t^A = \emptyset] \tag{1.2.15}$$

for all finite $A \subset \mathbb{Z}^d$. Since $\mathcal{L}(\eta_t^{\mathbb{Z}^d})$ is homogeneous for each $t \geq 0$, so is $\overline{\nu}$. Using (1.2.15) and survival, it is not hard to show that $\overline{\nu}$ is nontrivial. We claim that it is the only invariant law with this property and moreover, that

$$\mathcal{L}(\eta_t) \underset{t\to\infty}{\Longrightarrow} \overline{\nu} \tag{1.2.16}$$

when η is a contact process started in any initial law $\mathcal{L}(\eta_0) = \mu$ that nontrivial and homogeneous. To prove this, we observe that the contact process exhibits extinction versus unbounded growth in the following sense:

$$P\big[\exists t \geq 0 \text{ s.t. } \eta_t^A = \emptyset \quad \text{or} \quad \lim_{t\to\infty} |\eta_t^A| = \infty\big] = 1 \qquad (A_{\mathbb{Z}}^d), \tag{1.2.17}$$

where $|A|$ denotes the cardinality of a set A. The proof is basically the same as in the case of multitype branching (see formula (1.2.8)). Since it may happen that all infected sites become healthy before any further infection has taken place, it is easy to show that

$$\inf_{|A|\leq K} P\big[\exists t \geq 0 \text{ s.t. } \eta_t^A = \emptyset\big] > 0 \qquad (K \geq 0). \tag{1.2.18}$$

Thus, the probability that the process will die out can be estimated from below uniformly in all configurations with at most K infected sites, and therefore the only way for the process to avoid extinction is to let the number of infected sites tend to infinity.

Now let $\mathcal{L}(\eta_0) = \mu$ be nontrivial and homogeneous. Then, with a bit of trouble, it is possible to show that for each $t > 0$, the law $\mathcal{L}(\eta_t)$ has the property that

$$\lim_{K\to\infty} \sup_{|A|\leq K} P[\eta_t \cap A_n = \emptyset] = 0. \tag{1.2.19}$$

Therefore, by a restart argument, we may without loss of generality assume that $\mathcal{L}(\eta_0)$ has this property. Self-duality (formula (1.2.13)) tells us that

$$P[\eta_t \cap A = \emptyset] = P[\eta_0 \cap \eta_t^A = \emptyset] \qquad (t \geq 0), \tag{1.2.20}$$

where η_0 and η_t^A are independent. If t is large, then in evaluating the right-hand side of (1.2.20), by extinction versus unbounded growth (1.2.17), up to an event with small probability we need to consider only two cases. Either $\eta_t^A = \emptyset$, in which case $\eta_0 \cap \eta_t^A = \emptyset$, or $|\eta_t^A|$ is large, in which case $\eta_0 \cap \eta_t^A$ is with high probability not empty since $\mathcal{L}(\eta_0)$ has the property (1.2.19). It follows that $P[\eta_0 \cap \eta_t^A = \emptyset] \cong P[\eta_t^A = \emptyset]$ for large t, and taking the limit $t \to \infty$ in (1.2.20), using (1.2.15), we see that

$$\lim_{t\to\infty} P[\eta_t \cap A = \emptyset] = P[\eta_\infty^{\mathbb{Z}^d} \cap A = \emptyset], \tag{1.2.21}$$

for all finite $A \subset \mathbb{Z}^d$, which proves (1.2.16).

This argument is due to Harris [Har76, Theorem 9.2], who builds on earlier work of Vasil'ev, Vasershtein, Leontovich, and others. It can also be found in Ligget's book [Lig85, Theorem VI.4.8].

1.3 Overview of the habilitation thesis

1.3.1 Branching processes in renormalization theory

Certain problems in the study of a special type of interacting particle system, namely *linearly interacting catalytic Wright-Fisher diffusions*, lead one to study a special continuous-mass continuous- type space branching process, namely, the *super-Wright-Fisher diffusion*. This is a Markov process $\mathcal{Y} = (\mathcal{Y}_t)_{t\geq 0}$, taking values in the space of finite measures on $[0,1]$, whose transition probabilities are uniquely characterized by its Laplace functionals

$$E^\mu\big[e^{-\langle \mathcal{Y}_t, u_0\rangle}\big] = e^{-\langle \mu, u_t\rangle} \qquad (t \geq 0), \tag{1.3.1}$$

where $\langle \mu, f\rangle := \int f\,\mathrm{d}\mu$ and u is a mild solution of the semilinear Cauchy equation

$$\tfrac{\partial}{\partial t} u_t(x) = \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2} u_t(x) + \alpha u_t(x)(1 - u_t(x)) \qquad (t \geq 0), \tag{1.3.2}$$

with u_0 any nonnegative continuous function on $[0,1]$. One should think of (1.3.1) and (1.3.2) as continuous analogues of (1.2.1) and (1.2.2), respectively, where the finite type space $\{1,\ldots,n\}$ has been replaced by $[0,1]$ and the space $\mathbb{N}^n$ of all n-type particle configurations has been replaced by the space $\mathcal{M}[0,1]$ of all finite measures on $[0,1]$. We can think of $\mathcal{Y}_t$ as describing a population, consisting of many particles each of which has a very small mass, such that each particle performs a Wright-Fisher diffusion on $[0,1]$, that is, the Markov process in $[0,1]$ whose generator is (the closure of) the operator $\frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}$, and in addition, particles branch in such a way that the offspring of a bit of mass $\mathrm{d}m$ at position x during a time interval of length $\mathrm{d}t$ produces offspring with mean $(1+\alpha\mathrm{d}t)\mathrm{d}m$ and variance $\alpha\mathrm{d}t$.

The way how the super Wright-Fisher diffusion $\mathcal{Y}$ arises in a renormalization analysis of systems of linearly interacting catalytic Wright-Fisher diffusions will be explained in Chapter 2

For the moment, we take the process in (1.3.1) for granted, and ask about fixed point(s) and long-time convergence of solutions u to the Cauchy equation (1.3.2). We would like to play the same game as in Section 1.2.1 and use extinction versus unbounded growth of $\mathcal{Y}$ to prove convergence of u. Apart from the technical complications arising from continuous type space and continuous mass, we meet a more fundamental problem: our underlying motion, the Wright-Fisher diffusion, is not irreducible, i.e., it is not possible to get with positive probability from any point to any other point in the type space.

Indeed, the Wright-Fisher diffusion Y has two traps: 0 and 1, and the process started in any initial state satisfies

$$P\big[\exists\tau < \infty,\ r \in \{0,1\} \text{ s.t. } Y_t = r\ \forall t \geq \tau\big] = 1, \tag{1.3.3}$$

i.e., the process gets trapped in finite time. For the measure-valued process $\mathcal{Y}$, this means that with positive probability, in the long run most of the mass gets concentrated in 0, or 1, or both. Whether there is also a positive probability that there remains some mass in $(0,1)$ turns out to depend on the parameter α. For $\alpha > 1$, the answer is yes; otherwise it is no. As a result, we have to prove extinction versus unbounded growth on each of the part of the type space $\{0\}, \{1\}$, and $(0,1)$, and we find three or four (depending on α) different nonzero fixed points of (1.3.2), each with their own domain of attraction.

This analysis carried out in Sections 2.5–2.7 of Chapter 2. There, a similar analysis is carried out also for a related branching process in discrete time, the description of which is somewhat complicated. An important tool in this analysis is the use of embedded particle systems, as explained in Section 2.2.7. The results in this chapter are joint work with Klaus Fleischmann (WIAS, Berlin). Part of this has been published in [FS03].

1.3.2 Branching-coalescing particle systems

Consider a model of binary branching random walks, i.e., a collection of particles situated on a lattice Λ, where each particle moves independently of the others according to a continuous time random walk that jumps from site $i \in \Lambda$ to site j with rate $a(i,j)$, each particle splits with a branching rate $b \geq 0$ into two new particles, created on the position of the old one, and each particle dies with a death rate $d \geq 0$. Let $X_t(i)$ denotes the number of particles at time $t \geq 0$ at the site $i \in \Lambda$ and write $X_t := (X_t(i))_{i\in\Lambda}$. Then, in analogy with (1.2.1), one has

$$E^x\Big[\prod_{i=1}^n (1-u_0(i))^{X_t(i)}\Big] = \prod_{i=1}^n (1-u_t(i))^{x(i)} \qquad (t \geq 0), \tag{1.3.4}$$

whenever $u_t = (u_t(1), \ldots, u_t(n))$ is a $[0,1]^\Lambda$-valued solution to the system of differential equations

$$\tfrac{\partial}{\partial t} u_t(i) = \sum_j a(j,i)(u_t(j) - u_t(i)) + b u_t(i)(1-u_t(i)) - d u_t(i) \tag{1.3.5}$$

$(t \geq 0,\ i \in \Lambda)$. For each $f \in [0,1]^\Lambda$, set $U_t f := u_t$ $(t \geq 0)$ where u solves (1.2.2) with initial condition $u_0 = f$; then $(U_t)_{t\geq 0}$ is the *generating semigroup* of the branching process $X = (X_t)_{t\geq 0}$.

What happens if in the branching system X we also allow for *coalescence* of particles, i.e., if we let each pair of particles, present on the same site, coalesce with rate $2c$ (with $c \geq 0$) to one particle? In this case, we lose the branching property, i.e., we obtain a truly interacting system of particles. It turns out that although there is now no longer a generating semigroup in the classical sense, if we replace the deterministic evolution in (1.2.2) by the system of stochastic differential equations (SDE's)

$$\begin{aligned} \mathrm{d}u_t(i) = &\sum_j a(j,i)(u_t(j) - u_t(i))\,\mathrm{d}t + bu_t(i)(1-u_t(i))\,\mathrm{d}t - du_t(i)\,\mathrm{d}t \\ &+ \sqrt{2cu_t(i)(1-u_t(i))}\,\mathrm{d}B_t(i) \qquad\qquad (t \geq 0,\ i \in \Lambda), \end{aligned} \tag{1.3.6}$$

then formula (1.3.4) generalizes to the case with coalescence in the sense that

$$E\Big[\prod_{i=1}^{n}(1-u_0(i))^{X_t(i)}\Big] = E\Big[\prod_{i=1}^{n}(1-u_t(i))^{X_0(i)}\Big] \qquad (t \geq 0). \tag{1.3.7}$$

The duality (1.3.7) is due to [Shi81, SU86]. It turns out that the behavior of branching-coalescing particle systems of the type we have just described is very similar to that of the contact process. In fact, the history of this type of models seems to be as least as old as that of the contact process. In particular, our model is a special case of Schlögl's first model [Sch72].

Given the similarity of X with a contact process, and the similarity of the duality (1.3.7) with the self-duality of the contact process (1.2.13), one can try to mimick the proof of (1.2.16) in the present set-up. This was done by Shiga and Uchiyama in [SU86] for solutions u to the system of SDE's (1.3.6). More precisely, they used extinction versus unbounded growth for the particle system X to prove that the law of the system of SDE's u, started in any nontrivial homogeneous initial law, converges for $t \to \infty$ to the upper invariant law of u.

We note that if the death rate d is positive, then the probability that the process X will get extinct can be estimated from below uniformly in all configurations with at most K particles. Therefore, extinction versus unbounded growth for X follows by the same argument as in Sections 1.2.1 and 1.2.2. If $d = 0$, the process cannot get extinct. In this case, it is not completely trivial to show that the number of particles tends to infinity, which forced the authors of [SU86] to make some additional technical assumptions.

In Chapter 3, we turn the duality (1.3.7) around, and use extinction versus unbounded growth for the system of SDE's u to prove that the law of the particle system X started in any nontrivial homogeneous initial law, converges for $t \to \infty$ to the upper invariant law of X. This also involves some technical difficulties, since we need to show that the continuous system u may hit zero in finite time, and we need to show that X has an upper invariant law, which means that we must show that X can be started with infinitely many particles at every site.

These problems can be overcome, however, and we end up with results that are stronger than those in [SU86]. Additional tools that we use are a self-duality for the system of SDE's u, as well as the fact that the particle system X can be obtained from u by Poissonization. This is joint work with Siva Athreya (Bangalore), and has been published in [AS05].

1.3.3 The contact process seen from a typical site

In the last chapter of this thesis, we return to the classical contact process, but instead of studying the process started in a nontrivial homogeneous initial law as in Section 1.2.2, we wish to study the process started in finite initial states. It is known that questions about this sort of initial states are much more difficult than those about homogeneous initial laws. Nevertheless, a lot is known for the standard, nearest neighbor process on $\mathbb{Z}^d$. A central technical tool in this work is a dynamical block technique due to [BG90], which shows that the contact process, whenever it survives, can be compared with oriented percolation with an arbitrary high parameter. This technique finds its origin in older (although published later) work on unoriented percolation [GM90, BGN91].

While this technique has been very successful for the symmetric nearest-neighbor contact process on $\mathbb{Z}^d$, and can no doubt be extended to short-range contact processes on the same lattice, it is not obvious if it can be adapted to asymmetric processes, or to other lattices than $\mathbb{Z}^d$. Nevertheless, the study of contact processes on other lattices than $\mathbb{Z}^d$ is interesting both from a theoretical and practical poiint of view. The theoretical motivation comes from analogies with unoriented percolation on general transitive graphs, which has proved to be a fruitful topic (see, e.g., [BLPS99]). For unoriented percolation, it is known that it is important whether the underlying lattice is amenable (such as $\mathbb{Z}^d$) or not (e.g. a regular tree). Work on the contact process on regular trees by [Pem92, DS95, Lig96, Sta96] makes one suspect that a similar dichotomy could hold for the contact process.

In Chapter 4, we study contact processes on general countable groups Λ. We use a technique from the theory of branching processes, namely Palm measures, to show that indeed, certain aspects of the behavior of the contact process started in finite initial states depend on a property of the underlying lattice. The property that turns out to be important is whether Λ has subexponential growth, which is in fact a bit stronger than amenability.

Somewhat surprisingly, it turns out that in this context, extinction versus unbounded growth can again be of use to us. We will see that the local law of the process as seen from a typical 'Palmed' infected site at a typical late time can approximately be described by a monotone, translation invariant, harmonic function of the contact process. It is not hard to see that if η^Λ_∞ is a random variable with law $\mathcal{L}(\eta^\Lambda_\infty) = \overline{\nu}$, the upper invariant law, then

$$f(A) := P[\eta^\Lambda_\infty \cap A \neq \emptyset] \tag{1.3.8}$$

also defines an (a priori different) monotone, translation invariant, harmonic function f. The key argument in Chapter 4 uses extinction versus unbounded growth, plus duality, to show that this is up to a multiplicative constant the only such function. This extends the classical result, outlined in Section 1.2.2, that $\overline{\nu}$ is the only nontrivial homogeneous invariant law.

Chapter 2

Renormalization of catalytic Wright-Fisher diffusions

G

2.1 Introduction

2.1.1 Linearly interacting diffusions

Let $D \subset \mathbb{R}^d$ be open and convex, let $\overline{D}$ denote its closure, and assume that $0 \in \overline{D}$. Let Λ be a countably infinite group, with group action denoted by $(\xi, \eta) \mapsto \xi\eta$ and unit element 0. Let $a : \Lambda \times \Lambda \to \mathbb{R}$ be summable and invariant with respect to left multiplication in the group, i.e.,

$$\sum_{\eta \in \Lambda} |a(\xi, \eta)| < \infty \quad \text{and} \quad a(\xi, \eta) = a(\zeta\xi, \zeta\eta) \quad (\xi, \eta, \zeta \in \Lambda), \tag{2.1.1}$$

and assume that a is irreducible in the sense that for all $\Delta \subset \Lambda$ with $\Delta \neq \emptyset, \Lambda$, there exist $\xi \in \Delta$ and $\eta \in \Lambda \backslash \Delta$ such that either $a(\xi, \eta) \neq 0$ or $a(\eta, \xi) \neq 0$. We assume moreove that

$$a(\xi, \eta) \geq 0 \qquad (\xi \neq \eta). \tag{2.1.2}$$

Consider a collection $\mathbf{x} = (\mathbf{x}_\xi)_{\xi \in \Lambda}$ of $\overline{D}$-valued processes, solving the martingale problem for the operator

$$\mathcal{A}f(x) := \sum_{\eta, \xi \in \Lambda} a(\eta, \xi) \sum_{i=1}^{d} x_{\eta,i} \tfrac{\partial}{\partial x_{\xi,i}} f(x) + \sum_{\xi \in \Lambda} \sum_{i,j=1}^{d} w_{ij}(x_\xi) \tfrac{\partial^2}{\partial x_{\xi,i} \partial x_{\xi,j}} f(x), \tag{2.1.3}$$

where we write $x = (x_\xi)_{\xi \in \Lambda}$ and $x_\xi = (x_{\xi,1}, \ldots, x_{\xi,d})$ for a point $x \in \overline{D}^\Lambda$, and the domain of $\mathcal{A}$ consists of all functions on $\overline{D}^\Lambda$ that depend only on finitely many coordinates through a $\mathcal{C}^{(2)}$ function of compact support. It is well-known that $\overline{D}^\Lambda$-valued (weak) solutions to a system of SDE's of the form

$$\mathrm{d}\mathbf{x}_\xi(t) = \sum_{\eta \in \Lambda} a(\eta, \xi) \mathbf{x}_\eta(t) \mathrm{d}t + \sqrt{2}\sigma(\mathbf{x}_\xi(t)) \mathrm{d}B_\xi(t) \qquad (t \geq 0,\ \xi \in \Lambda), \tag{2.1.4}$$

solve the martingale problem for $\mathcal{A}$, were $(B_\xi)_{\mathbf{x}\in\Lambda}$ is a system of independent d'-dimensional Brownian motions, and the $d \times d'$ matrix-valued function σ is continuous and satisfies

$$\sum_{k=1}^{d'} \sigma_{ik}(x)\sigma_{jk}(x) = w_{ij}(x). \tag{2.1.5}$$

Conversely (see [EK86, Theorem 5.3.3] for the finite dimensional case), every solution to the martingale problem for $\mathcal{A}$ can be represented as a solution to the SDE (2.1.4), where there is some freedom in the choice of the root σ of the diffusion matrix w.

Equation (2.1.4) says that $\mathbf{x}$ is a system of linearly interacting d-dimensional diffusions. As a result of assumption (2.1.2), the linear drift causes the components $(\mathbf{x}_\xi)_{\xi\in\Lambda}$ to be positively correlated.

Set

$$\lambda := a(0,0) - \sum_\xi a(0,\xi). \tag{2.1.6}$$

For reasons that will become clear in a moment (see formula (2.1.9) (i) and the remarks below it), if $\lambda > 0$, we have to assume that $\overline{D}$ is a cone in order for solutions of (2.1.4) to exist. Under suitable assumptions on the diffusion matrix w, it can then be shown that the system of SDE's (2.1.4) defines a strong Markov process in a Ligget-Spitzer space $\mathcal{E}_\gamma(\Lambda)$, defined as

$$\mathcal{E}_\gamma(\Lambda) := \big\{x \in \overline{D}^\Lambda : \sum_{\xi\in\Lambda} \gamma_\xi |x_\xi| < \infty\big\}, \tag{2.1.7}$$

where $(\gamma_\xi)_{\xi\in\Lambda}$ are strictly positive constants such that $\sum_{\xi\in\Lambda}\gamma_\xi < \infty$ and $\sum_{\eta\in\Lambda} a(\eta,\xi)\gamma_\eta \leq K\gamma_\xi$ $(\xi \in \Lambda)$, for some $K < \infty$. The Markov process $\mathbf{x}$ is uniquely defined by the lattice Λ, the *interaction kernel* a, the domain D, and the *diffusion matrix* w.

Basic information about the process $\mathbf{x}$ can be obtained by calculating its mean and covariances. Consider a random walk $R = (R_t)_{t\geq 0}$ on Λ that jumps from a point ξ to a point η with rate $a(\xi,\eta)$ $(\xi \neq \eta)$. This random walk is called the *underlying motion* of $\mathbf{x}$. Set

$$P_t(\xi,\eta) := P^\xi[R_t = \eta]. \tag{2.1.8}$$

and recall the definition of λ in (2.1.6). Write $\mathbf{x}_\xi(t) = (\mathbf{x}_{\xi,1}(t), \ldots, \mathbf{x}_{\xi,d}(t))$. Then

$$\begin{aligned}
&\text{(i)} \quad & E[\mathbf{x}_{\xi,i}(t)] &= e^{\lambda t}\sum_{\eta\in\Lambda} P_t(\eta,\xi)E[\mathbf{x}_{\eta,i}(0)],\\
&\text{(ii)} \quad & \mathrm{Cov}(\mathbf{x}_{\xi,i}(t),\mathbf{x}_{\eta,j}(t)) &= e^{2\lambda t}\sum_{\zeta,\vartheta} P_t(\zeta,\xi)P_t(\vartheta,\eta)\mathrm{Cov}(\mathbf{x}_{\zeta,i}(0),\mathbf{x}_{\vartheta,j}(0))\\
&&&\quad + \int_0^t e^{2\lambda s}\sum_\zeta P_s(\zeta,\xi)P_s(\zeta,\eta)E[w_{ij}(\mathbf{x}_\zeta(t-s))]\mathrm{d}s.
\end{aligned} \tag{2.1.9}$$

$(t \geq 0,\ \xi,\eta \in \Lambda,\ 1 \leq i,j \leq d)$. Let us start the process $\mathbf{x}$ in an initial law $\mathcal{L}(\mathbf{x}(0))$ that is *homogeneous* in the sense that it is invariant with respect to left multiplication in the group, i.e., $\mathcal{L}((\mathbf{x}_\xi(0))_{\xi\in\Lambda}) = \mathcal{L}((\mathbf{x}_{\zeta\xi}(0))_{\xi\in\Lambda})$ for each $\zeta \in \Lambda$. Then, as a function of the parameter λ, the process $\mathbf{x}$ experiences a *phase transition* at $\lambda = 0$. If $\lambda < 0$, then in many examples it can be shown that the process started in any homogeneous initial law converges, as $t \to \infty$, to a unique homogeneous invariant law ν. Letting $t \to \infty$ in (2.1.9) (i) we see that $\int \nu(\mathrm{d}x) x_{\xi,i} = 0$ for each $\xi \in \Lambda$, $i = 1, \ldots, d$. On the other hand, as one may guess from (2.1.9) (i), for $\lambda > 0$ the process becomes unstable in the sense that the process started in a nonzero homogeneous initial state does not converge to an invariant law, but grows exponentially.

In the *critical* case $\lambda = 0$, the long-time behavior of $\mathbf{x}$ is more subtle. Let us call

$$\partial_w D := \{x \in \overline{D} : w_{ij}(x) = 0 \ \forall i, j = 1, \ldots, d\} \tag{2.1.10}$$

the *effective boundary* of D (associated with w). Note that $\partial_w D$ is the set of traps of the process $\mathbf{x}$, in the sense that the process started in a constant initial state $\mathbf{x}_\xi(0) = \theta$ $(\xi \in \Lambda)$ with $\theta \in \partial_w D$ satisfies $\mathbf{x}_\xi(t) = \theta$ $(t \geq 0,\ \xi \in \Lambda)$. Let us say an initial law $\mathcal{L}(\mathbf{x}(0))$ is *nontrivial* if $P[\exists \theta \in \partial_w D \text{ s.t. } \mathbf{x}_\xi(0) = \theta \ \forall \xi \in \Lambda] = 0$.

A natural question is whether $\mathbf{x}$ has homogeneous nontrivial invariant laws. In order to guess the answer to this question, we must look at the covariance formula (2.1.9) (ii). We observe that

$$G(\xi, \eta) := \int_0^\infty \sum_\zeta P_t(\zeta, \xi) P_t(\zeta, \eta) \mathrm{d}t = E\Big[\int_0^\infty 1_{\{R_t^{\dagger,\xi} = \tilde{R}_t^{\dagger,\eta}\}} \mathrm{d}t\Big] \tag{2.1.11}$$

is the expected time spent together by two independent random walks $R^{\dagger,\xi}$ and $\tilde{R}^{\dagger,\eta}$, started in $R_0^{\dagger,\xi} = \xi$ and $\tilde{R}_0^{\dagger,\eta} = \eta$, and jumping from a point ξ to a point η with the reversed jump rates $a^\dagger(\xi, \eta) := a(\eta, \xi)$. If Λ is an abelian group, with group action denoted by $(\xi, \eta) \mapsto \xi + \eta$, then the difference $R_t^{\dagger,\xi} - \tilde{R}_t^{\dagger,\eta}$ is itself a random walk, with symmetrized jump rates $a_{\mathrm{s}}(\xi, \eta) := a(\xi, \eta) + a(\eta, \xi)$, and G is finite if and only this random walk is recurrent. In particular, this is true for finite range jump kernels on $\mathbb{Z}^n$ if and only if $n \leq 2$.

It follows from (2.1.9) (ii) that the process $\mathbf{x}$ cannot have nontrivial homogeneous invariant laws with finite second moments if $G(0,0) = \infty$. Indeed, it has been verified for a number of examples of finite range models on $\mathbb{Z}^n$, that $\mathbf{x}$ has nontrivial homogeneous invariant laws if and only if $n > 2$. More precisely, in the transient case $n > 2$, the process has a nontrivial homogeneous invariant law with mean θ for each $\theta \in \overline{D} \backslash \partial_w D$, which is the limit law of the process started in any spatially ergodic initial law with mean θ. This type of behavior is called *stable behavior*. On the other hand, in the recurrent case $n \leq 2$, the only homogeneous invariant laws of the process are the delta-measures δ_θ on constant configurations $\theta \in \partial_w D$. In this case, the law of the process started from a spatially ergodic initial law with mean $\theta \in \overline{D} \backslash \partial_w D$ converges, as time tends to infinity, to a convex combination of these delta measures. This means that there are regions in space of growing size, called *clusters*, where the process is approximately constant and equal to some $\theta \in \partial_w D$. This type of behavior is called *clustering*.

A general result on stable behavior for $d = 1$ (i.e., for one-dimensional domains D) can be found in [Shi92]. A general result on clustering for $d = 1$ can be found in [CFG96]. Some (weak) general results in dimensions $d \geq 2$ for bounded domains D can be found in [Swa00]. Below, we list some explicit examples that have been treated in the literature.

The Ornstein-Uhlenbeck process $\overline{D} = \mathbb{R}$, $w(x) = \alpha > 0$. This is a Gaussian model that has been studied in [Deu89]. This reference also contains results for the subcritical case $\lambda < 0$.

The super-random walk $\overline{D} = [0,\infty)$, $w(x) = \alpha x$, with $\alpha > 0$. This is the discrete space analogue of the well-known super-Brownian motion [Daw77, Daw93, Eth00]. Both the super-random walk and the super-Brownian motion are continuous-mass branching processes. For these models, the dichotomy between stable behavior and clustering can be proved with the help of Kallenberg's backward tree technique [Kal77, GW91].

The stepping stone model $\overline{D} = [0,1]$, $w(x) = \alpha x(1-x)$, with resampling parameter $\alpha > 0$. This model, on rather general lattices, has been treated by Shiga [Shi80a, Shi80b], who also gives results for the subcritical case $\lambda < 0$. The diffusion function $w(x) = x(1-x)$ is called the *Wright-Fisher diffusion function* and is motivated by applications in population dynamics. Generalizations to other diffusion functions $w : [0,1] \to \mathbb{R}$ that satisfy $w(0) = w(1) = 0$ and $w > 0$ on $(0,1)$ can be found in [NS80, CG94]. The multidimensional *Wright-Fisher diffusion matrix* $w_{ij}(x) := x_i(\delta_{ij} - x_j)$ on $\overline{D} := \{x \in \mathbb{R}^d : x_i \geq 0,\ \sum_{i=1}^d x_i \leq 1\}$ can be treated with the help of Donnelly and Kurtz's look-down construction [DK96, GLW05].

Catalytic branching $\overline{D} = [0,\infty)^2$, $w(x) = \begin{pmatrix} \alpha x_1 & 0 \\ 0 & \beta x_1 x_2 \end{pmatrix}$, with $\alpha, \beta > 0$. This model has been studied in [Pen04]. A continuous space version of this model, the catalytic super-Brownian motion, has been studied in [DF97a, DF97b, EF98, FK99]. A discrete particle version of this model has been studied in [GKW99].

Mutually catalytic branching $\overline{D} = [0,\infty)^2$, $w(x) = \begin{pmatrix} \alpha x_1 x_2 & 0 \\ 0 & \beta x_1 x_2 \end{pmatrix}$, with $\alpha, \beta > 0$. This model has been studied in [DP98]. Its continuous-space analogue, the mutually catalytic super-Brownian motion, has recieved a lot of attention [DEFMPX02a, DEFMPX02b, DF02, DFMPX03].

Catalytic Wright-Fisher diffusions $\overline{D} = [0,1]^2$, $w(x) = \begin{pmatrix} \alpha x_1(1-x_1) & 0 \\ 0 & p(x_1)x_2(1-x_2) \end{pmatrix}$, where $\alpha > 0$ and the *catalyzing function* $p : [0,1] \to [0,\infty)$ Lipschitz continuous. This model, with the first component replaced by a voter model (which heuristically corresponds to taking $\alpha = \infty$) has been studied in [GKW01]. This model will also be the main subject of our present chapter.

In the clustering regime (i.e., the case $\Lambda = \mathbb{Z}^n$ with $n \leq 2$, or more generally the case where the quantity $G(0,0)$ from (2.1.11) is infinite), it is an interesting problem to determine the *clustering distribution*

$$\lim_{t\to\infty} \mathcal{L}(\mathbf{x}_0(t)) \tag{2.1.12}$$

of the process started in a constant initial state $\mathbf{x}_\xi(0) = \theta$ $(\xi \in \Lambda)$, for all $\theta \in \overline{D}$. If this limit exists, then it will be concentrated on the effective boundary $\partial_w D$. In dimension $d = 1$, when $\partial_w D$ consists of the finite endpoints of the interval D, the clustering distribution is trivial. In particular, if $D = [0,1]$, then as a result of (2.1.9) (i), it is $\theta\delta_1 + (1-\theta)\delta_0$.

More generally, for any bounded domain D in dimensions $d \geq 1$, let H_w denote the class of w-harmonic functions, i.e., functions $h \in \mathcal{C}^{(2)}(\overline{D})$ satisfying $\sum_{ij} w_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} h(x) = 0$ on D. Assume that H_w has the property that

$$T^c_{x,t} h(H_w) \subset H_w \quad (t \geq 0,\ c > 0,\ x \in \overline{D}), \tag{2.1.13}$$

where

$$T^c_{x,t}h(y) := h(x + (y - x)e^{-ct}) \quad (t \geq 0,\ c > 0,\ x \in \overline{D}) \tag{2.1.14}$$

is the semigroup with generator $\sum_{i=1}^d c(x_i - y_i)\frac{\partial}{\partial y_i}$, i.e., the generator of a deterministic process with a linear drift with strength c towards x. Under this assumption, it has been shown in [Swa00] that (2.1.9) (i), in the critical case $\lambda = 0$, can be generalized to

$$E[h(\mathbf{x}_{\xi,i}(t))] = \sum_{\eta\in\Lambda} P_t(\eta,\xi)E[h(\mathbf{x}_{\eta,i}(0))] \qquad (t \geq 0,\ h \in H_w), \tag{2.1.15}$$

and this is enough to determine the clustering distribution uniquely. Indeed, the limit in (2.1.12) must be the unique H_w-harmonic measure on $\partial_w D$ with mean x. If (2.1.13) holds then we say that w has *invariant harmonics*. Diffusion matrices on higher-dimensional domains do not in general have invariant harmonics; this applies in particular to catalytic Wright-Fisher diffusions if the catalyzing function p satisfies $p(0) = 0$ and $p(1) > 0$.

To get an idea of what the clustering distribution could be in general, we need to analyze the behavior of $\mathbf{x}$ on large space and time scales. We start with the large space-time behavior of the usual stepping stone model.

2.1.2 Large space-time behavior

The behavior of the stepping stone model on $\mathbb{Z}^n$, with resampling parameter α, on large spatial and temporal scales can be studied with the help of its moment dual, a system of rate α coalescing random walks. In fact, it is in particular the $\alpha \to \infty$ limit of these models that has been studied in detail, that is, the voter model and its dual, a system of immediately coalescing random walks. A good reference is [CG86].

In this section, we will especially be interested in the case $n = 2$, which is the critical dimension for random walk to be recurrent. Indeed, a 2-dimensional random walk $(R_t)_{t\geq 0}$ is recurrent, but it is only barely so. This is expressed, for example, in the fact that the quantity

$$E\Big[\int_0^t 1_{\{R_s = 0\}}\Big] \tag{2.1.16}$$

tends very slowly to infinity as $t \to \infty$. (For a precise definition of critical recurrence, see [Kle96, formula (1.15)].) As a result, on $\mathbb{Z}^2$ we see critical phenomenon associated with the phase transition between recurrence and transience.

Let $\mathbf{x}$ be a finite-range stepping stone model on $\mathbb{Z}^2$, started in a constant configuration $\mathbf{x}_\xi(0) = \theta$ ($\xi \in \mathbb{Z}^2$), for some $\theta \in [0, 1]$. Let

$$\Delta^t_s := [0, t^{\frac{1}{2}e^{-s}}]^2 \cap \mathbb{Z}^2 \qquad (s, t \geq 0) \tag{2.1.17}$$

be a block of volume $t^{e^{-s}}$, and let

$$\mathbf{x}^s(t) := \frac{1}{|\Delta^t_s|}\sum_{\xi\in\Delta^t_s} \mathbf{x}_\xi(t) \qquad (s, t \geq 0) \tag{2.1.18}$$

be the average of $\mathbf{x}(t)$ over Δ^t_s. By combining [CG86, Theorem 5] and [FG94, Theorem 2] as described in [GKW01, Proposition 3.1], it follows that

$$\mathcal{L}\big((\mathbf{x}^s(t))_{s\geq 0}\big) \underset{t\to\infty}{\overset{\text{f.d.d.}}{\Longrightarrow}} (\mathbf{y}_s)_{s\geq 0}, \tag{2.1.19}$$

where $(\mathbf{y}_s)_{s\geq 0}$ is a Wright-Fisher diffusion, i.e., a solution to $\mathrm{d}\mathbf{y}_s = \sqrt{\mathbf{y}_s(1-\mathbf{y}_s)}\mathrm{d}B_s$, started in $\mathbf{y}_0 = \theta$. Here f.d.d. denotes convergence in finite dimensional distributions. (The question whether the convergence in f.d.d. can be replaced by weak convergence in path space is the subject of ongoing research.) Formula (2.1.19) shows how block averages at late times t change as we zoom in in space. Very large block avarages, over blocks of volume t, still show the original intensity θ that the process $\mathbf{x}$ was starting in. As we zoom in on smaller blocks of volume $t^{e^{-s}}$, with $s \geq 0$, the block averages change in a random way, until after some random time, the Wright-Fisher diffusion $\mathbf{y}_s$ hits 0 or 1, (with probabilities $1-\theta$ or θ, respectively), and from that random scale on, the block avarages are constant.

Note that the long-time behavior of the limiting diffusion $\mathbf{y}$ in (2.1.19) gives us the clustering distribution (2.1.12). It seems likely that similar results hold for other models as well; however, the limiting diffusion in (2.1.19) will not always be the Wright-Fisher diffusion. To find out what the limit could be more generally, it is helpful to replace the lattice $\mathbb{Z}^2$ by the hierarchical group, as explained in the next section.

2.1.3 Hierarchically interacting diffusions

For any $N \geq 2$, the *hierarchical group* with freedom N is the set Ω_N of all sequences $\xi = (\xi_1, \xi_2, \ldots)$, with coordinates ξ_k in the finite set $\{0, \ldots, N-1\}$, which are different from 0 only finitely often, equipped with componentwise addition modulo N. Setting

$$\|\xi\| := \min\{n \geq 0 : \xi_k = 0 \ \forall k > n\} \qquad (\xi \in \Omega_N), \tag{2.1.20}$$

$\|\xi - \eta\|$ is said to be the *hierarchical distance* between two sites ξ and η in Ω_N.

Let $\mathbf{x}^N = (\mathbf{x}^N_\xi)_{\xi\in\Omega_N}$ be a critical system of linearly interacting diffusions on Ω_N with interaction kernel given by

$$a_N(\xi,\eta) := \sum_{k=\|\xi-\eta\|}^{\infty} \frac{c_{k-1}}{N^{2k-1}} \quad (\xi \neq \eta), \quad a_N(\xi,\xi) := -\sum_{\eta\neq\xi} a_N(\xi,\eta), \tag{2.1.21}$$

where $(c_k)_{k\geq 0}$ are positive *migration constants* such that the quantity $\sum_\xi a_N(0,\xi) = \sum_k c_k/N^k$ is finite. The random walk associated with a_N is recurrent if and only if

$$\sum_{k=0}^{\infty} \frac{1}{d_k} = \infty, \qquad \text{where} \quad d_k := \sum_{n=0}^{\infty} \frac{c_{k+n}}{N^n} \tag{2.1.22}$$

(see [DG93a, Kle96]; a similar problem is treated in [DE68]).

Let $\Delta_k(\xi) := \{\eta : \|\xi - \eta\| \leq k\}$ denote the *k-block* around ξ and let

$$\mathbf{x}^k_\xi(t) := \frac{1}{|\Delta_k(\xi)|} \sum_{\eta:\|\xi-\eta\|\leq k} \mathbf{x}_\eta(t) \qquad\qquad (k \geq 0). \tag{2.1.23}$$

denote the *k-block average* around ξ. The sequence $(\mathbf{x}_0^0(t), \mathbf{x}_0^1(t), \ldots)$ of block-averages around the origin is called the *interaction chain.* Heuristic arguments suggest that in the *local mean field limit* $N \to \infty$, the interaction chain converges to a certain well-defined Markov chain. In order to charcterize this chain, we need a few definitions.

Definition 2.1 (Renormalization class and transformation) Let $D \subset \mathbb{R}^d$ be nonempty, convex, and open, and let $\overline{D}$ be its closure. Let $\mathcal{W}$ be a collection of continuous functions w from $\overline{D}$ into the space M_+^d of symmetric non-negative definite $d \times d$ real matrices, such that $\lambda w \in \mathcal{W}$ for every $\lambda > 0$, $w \in \mathcal{W}$. We call $\mathcal{W}$ a *prerenormalization class* on $\overline{D}$ if the following three conditions are satisfied:

(i) For each constant $c > 0$, $w \in \mathcal{W}$, and $x \in \overline{D}$, the martingale problem for the operator $A_x^{c,w}$ is well-posed, where

$$A_x^{c,w} f(y) := \sum_{i=1}^{d} c\,(x_i - y_i) \frac{\partial}{\partial y_i} f(y) + \sum_{i,j=1}^{d} w_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j} f(y) \qquad (y \in \overline{D}), \tag{2.1.24}$$

and the domain of $A_x^{c,w}$ is the space of real functions on $\overline{D}$ that can be extended to a twice continuously differentiable function on $\mathbb{R}^d$ with compact support.

(ii) For each $c > 0$, $w \in \mathcal{W}$, and $x \in \overline{D}$, the martingale problem for $A_x^{c,w}$ has a unique stationary solution with invariant law denoted by $\nu_x^{c,w}$.

(iii) For each $c > 0$, $w \in \mathcal{W}$, $x \in \overline{D}$, and $i, j = 1, \ldots, d$, one has $\int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)|w_{ij}(y)| < \infty$.

If $\mathcal{W}$ is a prerenormalization class, then we define for each $c > 0$ and $w \in \mathcal{W}$ a matrix-valued function $F_c w$ on $\overline{D}$ by

$$F_c w(x) := \int_{\overline{D}} \nu_x^{c,w}(dy) w(y) \qquad (x \in \overline{D}). \tag{2.1.25}$$

We say that $\mathcal{W}$ is a *renormalization class* on $\overline{D}$ if in addition:

(iv) For each $c > 0$ and $w \in \mathcal{W}$, the function $F_c w$ is an element of $\mathcal{W}$.

If $\mathcal{W}$ is a renormalization class and $c > 0$, then the map $F_c : \mathcal{W} \to \mathcal{W}$ defined by (2.1.25) is called the *renormalization transformation* on $\mathcal{W}$ with *migration constant* c. In (2.1.24), w is called the *diffusion matrix* and x the *attraction point.* ◊

For any renormalization class $\mathcal{W}$ and any sequence of (strictly) positive migration constants $(c_k)_{k \geq 0}$, we define *iterated renormalization transformations* $F^{(n)} : \mathcal{W} \to \mathcal{W}$, as follows:

$$F^{(n+1)} w := F_{c_n}(F^{(n)} w) \quad (n \geq 0) \quad \text{with} \quad F^{(0)} w := w \qquad (w \in \mathcal{W}_{\mathrm{cat}}). \tag{2.1.26}$$

We set $s_0 := 0$ and

$$s_n := \sum_{k=0}^{n-1} \frac{1}{c_k} \qquad (1 \leq n \leq \infty). \tag{2.1.27}$$

With these definitions, we can formulate the following conjecture about the behavior of the interaction chain in the local mean field limit $N \to \infty$.

Conjecture 2.2 *Let $\mathcal{W}$ be a renormalization class. Fix $w \in \mathcal{W}$, $\theta \in D$, and positive numbers $(c_k)_{k\geq 0}$ such that for N large enough, $\sum_k c_k/N^k < \infty$. For all N large enough, let $\mathbf{x}^N$ be a solution to (2.1.4) on $\Lambda = \Omega_N$ with $a = a_N$ from (2.1.21), and assume that t_N are constants such that, for some $n \geq 1$, $\lim_{N\to\infty} N^{-n} t_N = T \in [0,\infty)$. Then*

$$\left(\mathbf{x}_0^{N,n}(t_N), \ldots, \mathbf{x}_0^{N,0}(t_N)\right) \underset{N\to\infty}{\Longrightarrow} (I^w_{-n}, \ldots, I^w_0), \tag{2.1.28}$$

where $(I^w_{-n}, \ldots, I^w_0)$ is a Markov chain with transition laws

$$P[I^w_{-k} \in \mathrm{d}y | I^w_{-k-1} = x] = \nu_x^{c_k, F^{(k)}w}(\mathrm{d}y) \qquad (x \in \overline{D},\ 0 \leq k \leq n-1) \tag{2.1.29}$$

and initial state

$$I^w_{-n} = \mathbf{y}_T, \quad \textit{where} \quad \mathrm{d}\mathbf{y}_t = c_n(\theta - \mathbf{y}_t)\mathrm{d}t + \sqrt{2}\sigma^{(n)}(\mathbf{y}_t)\mathrm{d}B_t, \quad \mathbf{y}_0 = \theta, \tag{2.1.30}$$

and $\sigma^{(n)}$ is a root of the diffusion matrix $F^{(n)}w$.

Rigorous versions of conjecture 2.2 have been proved for renormalization classes on $\overline{D} = [0,1]$ and $\overline{D} = [0,\infty)$ in [DG93a, DG93b]. See [DG96, DGV95] for similar results. Note that the Markov chain $I^w = (I^w_{-n}, \ldots, I^w_0)$ is a sort of analogue of the block averages $(\mathbf{x}^s(t))_{s\geq 0}$ defined in (2.1.18). As we will see below, for appropriate choices of the constants $(c_k)_{k\geq 0}$, the discrete chain I^w can be approximated by a diffusion, in the spirit of (2.1.19). In order to see this, we need a few facts about renormalization classes. To keep things as simple as possible, we specialize to renormalization classes on bounded domains, although much of what we will say, with some modifications here and there, can be generalized to unbounded domains.

2.1.4 Renormalization classes

In this section, we describe some elementary properties that hold generally for (pre-) renormalization classes on bounded domains. The proofs of Lemmas 2.3–2.8 can be found in Section 2.3.1 below.

Fix a prerenormalization class $\mathcal{W}$ on a set $\overline{D}$ where $D \subset \mathbb{R}^d$ is open, bounded, and convex. Then $\mathcal{W}$ is a subset of the cone $\mathcal{C}(\overline{D}, M^d_+)$ of continuous M^d_+-valued functions on $\overline{D}$. We equip $\mathcal{C}(\overline{D}, M^d_+)$ with the topology of uniform convergence. We let $\mathcal{M}_1(\overline{D})$ denote the space of probability measures on $\overline{D}$, equipped with the topology of weak convergence. Our first lemma says that the equilibrium measures $\nu_x^{c,w}$ and the renormalized diffusion matrices $F_c w(x)$ are continuous in their parameters.

Lemma 2.3 (Continuity in parameters)

(a) *The map $(x, c, w) \mapsto \nu_x^{c,w}$ from $\overline{D} \times (0,\infty) \times \mathcal{W}$ into $\mathcal{M}_1(\overline{D})$ is continuous.*

(b) *The map $(x, c, w) \mapsto F_c w(x)$ from $\overline{D} \times (0,\infty) \times \mathcal{W}$ into M^d_+ is continuous.*

In particular, $x \mapsto \nu_x^{c,w}$ is a continuous probability kernel on $\overline{D}$, and $F_c w \in \mathcal{C}(\overline{D}, M^d_+)$ for all $c > 0$ and $w \in \mathcal{W}$. Recall from Definition 2.1 that $\lambda w \in \mathcal{W}$ for all $w \in \mathcal{W}$ and $\lambda > 0$. The reason why we have included this assumption is that it is convenient to have the next scaling lemma around, which is a consequence of time scaling.

Lemma 2.4 (Scaling property of renormalization transformations) *One has*

$$\left.\begin{array}{ll}\text{(i)} & \nu_x^{\lambda c,\lambda w} = \nu_x^{c,w} \\ \text{(ii)} & F_{\lambda c}(\lambda w) = \lambda F_c w\end{array}\right\} \quad (\lambda, c > 0,\ w \in \mathcal{W},\ x \in \overline{D}). \tag{2.1.31}$$

The following simple lemma will play a crucial role in what follows.

Lemma 2.5 (Mean and covariance matrix) *For all $x \in \overline{D}$ and $i, j = 1, \ldots, d$, the mean and covariances of $\nu_x^{c,w}$ are given by*

$$\begin{array}{ll}\text{(i)} & \displaystyle\int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)(y_i - x_i) = 0, \\ \text{(ii)} & \displaystyle\int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)(y_i - x_i)(y_j - x_j) = \tfrac{1}{c} F_c w_{ij}(x).\end{array} \tag{2.1.32}$$

Recall the definition of the effective boundary associated with a diffusion matrix w in (2.1.10). The next lemma says that the effective boundary is invariant under renormalization.

Lemma 2.6 (Invariance of effective boundary) *One has $\partial_{F_c w} D = \partial_w D$ for all $w \in \mathcal{W}$, $c > 0$.*

From now on, let $\mathcal{W}$ be a renormalization class, i.e., $\mathcal{W}$ satisfies also condition (iv) from Definition 2.1. Fix a sequence of (positive) migration constants $(c_k)_{k\geq 0}$. By definition, the *iterated probability kernels* $K^{w,(n)}$ associated with a diffusion matrix $w \in \mathcal{W}$ (and the constants $(c_k)_{k\geq 0}$) are the probability kernels on $\overline{D}$ defined inductively by

$$K_x^{w,(n+1)}(\mathrm{d}z) := \int_{\overline{D}} \nu_x^{c_n, F^{(n)} w}(\mathrm{d}y) K_y^{w,(n)}(\mathrm{d}z) \quad (n \geq 0) \quad \text{with} \quad K_x^{w,(0)}(\mathrm{d}y) := \delta_x(\mathrm{d}y), \tag{2.1.33}$$

with $F^{(n)}$ as in (2.1.26). Note that $K^{w,(n)}$ is the transition probability from time $-n$ to time 0 of the interaction chain in the local mean-field limit (see Conjecture 2.2):

$$K_x^{w,(n)}(\mathrm{d}y) := P[I_0^w \in \mathrm{d}y | I_{-n}^w = x] \qquad (x \in \overline{D},\ n \geq 0). \tag{2.1.34}$$

Note moreover that

$$F^{(n)} w(x) = \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y) w(y) \qquad (x \in \overline{D},\ n \geq 0). \tag{2.1.35}$$

The next lemma follows by iteration from Lemmas 2.3 and 2.5. It their essence, this lemma and Lemma 2.8 below go back to [BCGH95].

Lemma 2.7 (Basic properties of iterated kernels) *For each $w \in \mathcal{W}$, the $K^{w,(n)}$ are continuous probability kernels on $\overline{D}$. Moreover, for all $x \in \overline{D}$, $i, j = 1, \ldots, d$, and $n \geq 0$, the mean and covariance matrix of $K_x^{w,(n)}$ are given by*

$$\begin{array}{ll}\text{(i)} & \displaystyle\int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y)(y_i - x_i) = 0, \\ \text{(ii)} & \displaystyle\int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y)(y_i - x_i)(y_j - x_j) = s_n F^{(n)} w_{ij}(x).\end{array} \tag{2.1.36}$$

We equip the space $\mathcal{C}(\overline{D}, \mathcal{M}_1(\overline{D}))$ of continuous probability kernels on $\overline{D}$ with the topology of uniform convergence (since $\mathcal{M}_1(\overline{D})$ is compact, there is a unique uniform structure on $\mathcal{M}_1(\overline{D})$ generating the topology). For 'nice' renormalization classes, it seems reasonable to conjecture that the kernels $K^{w,(n)}$ converge as $n \to \infty$ to some limit $K^{w,*}$ in $\mathcal{C}(\overline{D}, \mathcal{M}_1(\overline{D}))$. If this happens, then formula (2.1.36) (ii) tells us that the rescaled renormalized diffusion matrices $s_n F^{(n)} w$ converge uniformly on $\overline{D}$ to the covariance matrix of $K^{w,*}$.

We will mainly be interested in the case that $\lim_{n\to\infty} s_n = \infty$. Indeed, if the iterated kernels converge to a limit $K^{w,*}$, then this condition guarantees that this limit is concentrated on the effective boundary:

Lemma 2.8 (Concentration on the effective boundary) *If $s_n \underset{n\to\infty}{\longrightarrow} \infty$, then for any $f \in \mathcal{C}(\overline{D})$ such that $f = 0$ on $\partial_w D$:*

$$\lim_{n\to\infty} \sup_{x\in\overline{D}} \Big| \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y) f(y) \Big| = 0. \tag{2.1.37}$$

Note that $s_n \to \infty$ if and only if $\sum_k 1/c_k = \infty$. We can think of this condition as the $N \to \infty$ limit of the condition $\sum_k 1/d_k = \infty$ in (2.1.22). Thus, the condition $s_n \to \infty$ guarantees that the corresponding system of linearly interacting diffusions on the hierarchical group with migration constants $(c_k)_{k\geq 0}$ *clusters in the local mean field limit.*

Most of the discussion in this section carries over to renormalization classes on unbounded D, but in this case, the second moments of the iterated kernels $K^{w,(n)}$ may diverge as $n \to \infty$. As a result, because of formula (2.1.36) (ii), the s_n may no longer be the right scaling factors to find a nontrivial limit of the renormalized diffusion matrices; see, for example, [BCGH97].

2.1.5 Rescaled transformations

We return to renormalization classes on bounded domains, and focus our attention on the clustering regime $s_n \to \infty$. Since we expect $s_n F^{(n)} w$ to converge to a limit (namely, the covariance matrix of $K^{w,*}$), we will use Lemma 2.4 to convert the rescaled iterates $s_n F^{(n)}$ into (usual, not rescaled) iterates of another transformation. For this purpose, it will be convenient to modify the definition of our scaling constants s_n a little bit. Fix some $\beta > 0$ and put

$$\overline{s}_n := \beta + s_n \qquad (n \geq 0). \tag{2.1.38}$$

Define *rescaled renormalization transformations* $\overline{F}_\gamma : \mathcal{W} \to \mathcal{W}$ by

$$\overline{F}_\gamma w := (1+\gamma) F_{1/\gamma} w \qquad (\gamma > 0,\ w \in \mathcal{W}). \tag{2.1.39}$$

Using (2.1.31) (ii), one easily deduces that

$$\overline{s}_n F^{(n)} w = \overline{F}_{\gamma_{n-1}} \circ \cdots \circ \overline{F}_{\gamma_0}(\beta w) \qquad (w \in \mathcal{W},\ n \geq 1), \tag{2.1.40}$$

where

$$\gamma_n := \frac{1}{\overline{s}_n c_n} \qquad (n \geq 0). \tag{2.1.41}$$

We can reformulate the condition $s_n \to \infty$ from Lemma 2.8 in terms of the constants $(\gamma_n)_{n\geq 0}$. Indeed, it is not hard to check[1] that the following three conditions are equivalent:

$$\text{(i)}\quad s_n \underset{n\to\infty}{\longrightarrow} \infty, \qquad \text{(ii)}\quad \overline{s}_n \underset{n\to\infty}{\longrightarrow} \infty, \qquad \text{(iii)}\quad \sum_n \gamma_n = \infty. \tag{2.1.42}$$

In view of (2.1.40), it is natural to assume that the γ_n converge to a limit $\gamma^* \in [0,\infty]$. Since $\overline{s}_{n+1}/\overline{s}_n = 1+\gamma_n$, it is not hard to see that the following conditions are equivalent:

$$\text{(i)}\quad \frac{s_{n+1}}{s_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^*, \qquad \text{(ii)}\quad \frac{\overline{s}_{n+1}}{\overline{s}_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^*, \qquad \text{(iii)}\quad \gamma_n \underset{n\to\infty}{\longrightarrow} \gamma^*. \tag{2.1.43}$$

If $0<\gamma^*<\infty$, then, in the light of (2.1.40), we expect $\overline{s}_n F^{(n)}w$ to converge to a fixed point of the transformation $\overline{F}_{\gamma^*}$. If $\gamma^*=0$, the situation is more complex. In this case, we expect the orbit $\overline{s}_n F^{(n)}w \mapsto \overline{s}_{n+1}F^{(n+1)}w \mapsto \cdots$, for large n, to approximate a continuous flow, the generator of which is

$$\lim_{\gamma\to 0} \gamma^{-1}\Big(\overline{F}_\gamma w - w\Big)(x) = \tfrac{1}{2}\sum_{i,j=1}^{d} w_{ij}(x)\tfrac{\partial^2}{\partial x_i \partial x_j} w(x) + w(x) \qquad (x\in\overline{D}). \tag{2.1.44}$$

To see that the right-hand side of this equation equals the left-hand side if w is twice continuously differentiable, one needs a Taylor expansion of w together with the moment formulas (2.1.32) for $\nu_x^{1/\gamma,w}$. Under condition condition (2.1.42) (iii), we expect this continuous flow to reach equilibrium.

In the light if these considerations, we are led to at the following general conjecture.

Conjecture 2.9 (Limits of rescaled renormalized diffusion matrices) *Assume that $s_n\to\infty$ and $s_{n+1}/s_n \to 1+\gamma^*$ for some $\gamma^*\in[0,\infty]$. Then, for any $w\in\mathcal{W}$,*

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*, \tag{2.1.45}$$

where w^* satisfies

$$\begin{array}{llll}
\text{(i)} & \overline{F}_{\gamma^*} w^* = w^* & & \textit{if}\ \ 0<\gamma^*<\infty, \\
\text{(ii)} & \tfrac{1}{2}\sum_{i,j=1}^{d} w^*_{ij}(x)\tfrac{\partial^2}{\partial x_i\partial x_j} w^*(x) + w^*(x) = 0 & (x\in\overline{D}) & \textit{if}\ \ \gamma^*=0, \\
\text{(iii)} & \lim_{\gamma\to\infty} \overline{F}_\gamma w^* = w^* & & \textit{if}\ \ \gamma^*=\infty.
\end{array} \tag{2.1.46}$$

We call (2.1.46) (ii), which is in some sense the $\gamma^*\to 0$ limit of the fixed point equation (2.1.46) (i), the *asymptotic fixed point equation.* A version of formula (2.1.46) (ii) occurred in [Swa99, formula (1.3.5)] (a minus sign is missing there).

In particular, one may hope that for a given effective boundary, the equations in (2.1.46) have a unique solution. Our main result (Theorem 2.17 below) confirms this conjecture for a renormalization class of catalytic Wright-Fisher diffusions and for $\gamma^*<\infty$. In Section 2.1.7 below, we discuss numerical evidence that supports Conjecture 2.9 in the case $\gamma^*=0$ for other renormalization classes on compacta as well.

[1] To see this, let $\overline{s}_\infty \in (0,\infty]$ denote the limit of the $\overline{s}_n$ and note that on the one hand, $\sum_n 1/(\overline{s}_n c_n) \geq \sum_n \log(1+1/(\overline{s}_n c_n)) = \log(\prod_n \overline{s}_{n+1}/\overline{s}_n) = \log(\overline{s}_\infty/\overline{s}_1)$, while on the other hand $\sum_n 1/(\overline{s}_n c_n) \leq \prod_n (1+1/(\overline{s}_n c_n)) = \prod_n \overline{s}_{n+1}/\overline{s}_n = \overline{s}_\infty/\overline{s}_1$.

2.1.6 Diffusive clustering

Assuming that the rescaled renormalized diffusion matrices $s_n F^{(n)} w$ converge to a limit w^*, we can make a guess about the limit of the iterated probability kernels $K^{w,(n)}$.

Conjecture 2.10 (Limits of iterated probability kernels) *Assume that $s_n F^{(n)} w \to w^*$ as $n \to \infty$. Then, for any $w \in \mathcal{W}$,*

$$K^{w,(n)} \underset{n\to\infty}{\longrightarrow} K^*, \tag{2.1.47}$$

where K^* has the following description:

(i) *If $0 < \gamma^* < \infty$, then*
$$K^*_x = \lim_{n\to\infty} P^x[I^{\gamma^*}_n \in \cdot\,], \tag{2.1.48}$$
where $(I^{\gamma^*}_n)_{n\geq 0}$ is the Markov chain with transition law $P[I^{\gamma^*}_{n+1} \in \cdot\,| I^{\gamma^*}_n = x] = \nu^{1/\gamma^*,w^*}$.

(ii) *If $\gamma^* = 0$, then*
$$K^*_x = \lim_{t\to\infty} P^x[I^0_t \in \cdot\,], \tag{2.1.49}$$
where $(I^0_s)_{s\geq 0}$ is the diffusion process with generator $\sum_{i,j=1}^d w^*_{ij}(y)\frac{\partial^2}{\partial y_i \partial y_j}$.

(iii) *If $\gamma^* = \infty$, then*
$$K^*_x = \lim_{\gamma\to\infty} \nu^{1/\gamma,w^*}_x. \tag{2.1.50}$$

If $\gamma^* < \infty$, this conjecture is motivated by the observation that in this case, the Markov chain $(I^w_{-n}, \ldots, I^w_0)$ from Conjecture 2.2 is approximately time homogeneous for $n \to \infty$. The case $\gamma^* = 0$ is of particular interest. In this case $I^w_{-n}, I^w_{-n+1}, \ldots$ converges, in the right scaling, to the diffusion $(I^0_s)_{s\geq 0}$ with diffusion matrix w^*. This is a sort of analogon of the diffusive clustering result (2.1.19). Based on this analogy, we can make one more conjecture.

Conjecture 2.11 (Clustering distribution on $\mathbb{Z}^2$) *Let $D \subset \mathbb{R}^d$ be open, bounded, and convex, and let $\mathcal{W}$ be a renormalization class on $\overline{D}$. Assume that the asymptotic fixed point equation (2.1.46) (ii) has a unique solution w^* in $\mathcal{W}$. Let σ be a continuous root of a diffusion matrix $w \in \mathcal{W}$. Let $\mathbf{x} = (\mathbf{x}_\xi)_{\xi\in\mathbb{Z}^2}$ be a $\overline{D}^{\mathbb{Z}^2}$-valued process, solving the system of SDE's*

$$\mathrm{d}\mathbf{x}_\xi(t) = \sum_{\eta:\,|\eta-\xi|=1} \big(\mathbf{x}_\eta(t) - \mathbf{x}_\xi(t)\big)\,\mathrm{d}t + \sigma(\mathbf{x}_\xi(t))\mathrm{d}B_\xi(t), \tag{2.1.51}$$

with initial condition $\mathbf{x}_\xi(0) = \theta \in \overline{D}$ $(\xi \in \mathbb{Z}^2)$. Then

$$\mathcal{L}(\mathbf{x}_0(t)) \underset{t\to\infty}{\Longrightarrow} P[I^0_\infty \,|\, I^0_0 = \theta] \qquad (\xi \in \mathbb{Z}^2), \tag{2.1.52}$$

where $(I^0_s)_{s\geq 0}$ is the diffusion with generator $\sum_{ij} w^*_{ij}(y)\frac{\partial^2}{\partial y_i \partial y_j}$.

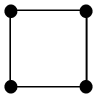
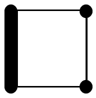
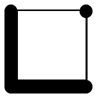
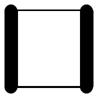
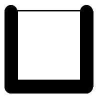
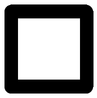
case	effective boundary	fixed points w^* of (2.1.53)
1		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & x_2(1-x_2) \end{pmatrix}$
2		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & p^*_{0,1,0}(x_1)x_2(1-x_2) \end{pmatrix}$
3		$\begin{pmatrix} q^*(x_1,x_2) & 0 \\ 0 & q^*(x_2,x_1) \end{pmatrix}$
4		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & 0 \end{pmatrix}$
5		$\begin{pmatrix} x_1(1-x_1)1_{\{x_2>0\}} & 0 \\ 0 & 0 \end{pmatrix}$
6		$g^*(x_1,x_2)\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$

Figure 2.1: Fixed points of the flow (2.1.53).

2.1.7 Numerical solutions to the asymptotic fixed point equation

Let $t \mapsto w(t,\,\cdot\,)$ be a solution to the continuous flow with the generator in (2.1.44), i.e., w is an M^d_+-valued solution to the nonlinear partial differential equation

$$\frac{\partial}{\partial t}w(t,x) = \tfrac{1}{2}\sum_{i,j=1}^{d} w_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}w(t,x) + w(t,x) \qquad (t \geq 0,\ x \in \overline{D}). \tag{2.1.53}$$

Solutions to (2.1.53) are quite easy to simulate on a computer. We have simulated solutions for all kind of diffusion matrices (including nondiagonal ones) on the unit square $[0,1]^2$, with the effective boundaries 1–6 depicted in Figure 2.1. For all initial diffusion matrices $w(0,\,\cdot\,)$ we tried, the solution converged as $t \to \infty$ to a fixed point w^*. In all cases except case 6, the fixed point was unique. The fixed points are listed in Figure 2.1. The functions $p^*_{0,1,0}$ and q^* from Figure 2.1 are plotted in Figure 2.2.

The fixed points for the effective boundaries in cases 1,2, and 4 will be described in Theorem 2.17 below. In particular, $p^*_{0,1,0}$ is the function from Theorem 2.17 (c). The simulations suggest that the domain of attraction of these fixed points (within the class of "all" diffusion matrices on $[0,1]^2$) is actually a lot larger than the classes for which we are able to prove convergence in Theorem 2.17.

The function q^* from case 3 satisfies $q^*(x_1,1) = x_1(1-x_1)$ and is zero on the other parts

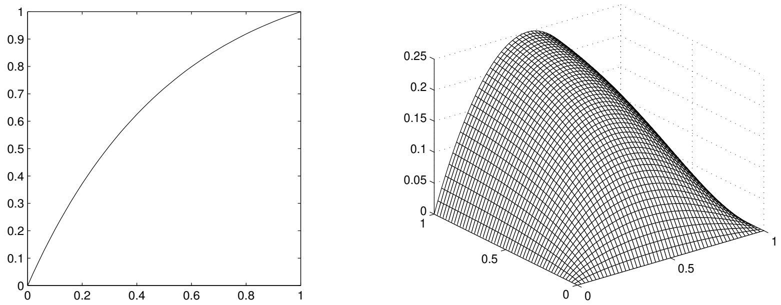


Figure 2.2: The functions $p^*_{0,1,0}$ and q^* from cases 2 and 3 of Figure 2.1.

of the boundary. In contrast to what one might perhaps guess in view of case 2, q^* is *not* of the form $q^*(x_1,x_2) = f(x_2)x_1(1-x_1)$ for some function f.

Case 5 is somewhat degenerate since in this case the fixed point is not continuous.

The only case where the fixed point is not unique is case 6. Here, m can be any positive definite matrix, while g^*, depending on m, is the unique solution on $(0,1)^2$ of the equation $1+\frac{1}{2}\sum_{i,j=1}^2 m_{ij}\frac{\partial^2}{\partial x_i \partial x_j} g^*(x) = 0$, with zero boundary conditions. Some diffusion matrices that are in the domain of attraction of these fixed points are described in Theorem 2.14 below. The simulations indicate that the true domain of attraction is much larger than what can be proved (and includes nonisotropic matrices).

2.1.8 Known results

In this section we discuss some results that have been derived previously for renormalization classes on compact sets.

Theorem 2.12 [BCGH95, DGV95] (Universality class of Wright-Fisher models) *Let $D := \{x \in \mathbb{R}^d : x_i > 0\ \forall i,\ \sum_{i=1}^d x_i < 1\}$, and let $\{e_0,\ldots,e_d\}$, with $e_0 := (0,\ldots,0)$ and $e_1 := (1,0,\ldots,0),\ldots,\ e_d := (0,\ldots,0,1)$ be the extremal points of $\overline{D}$. Let $w^*_{ij}(x) := x_i(\delta_{ij}-x_j)$ $(x \in \overline{D},\ i,j = 1,\ldots,d)$ denote the standard Wright-Fisher diffusion matrix, and assume that $\mathcal{W}$ is a renormalization class on $\overline{D}$ such that $w^* \in \mathcal{W}$ and $\partial_w \overline{D} = \{e_0,\ldots,e_d\}$ for all $w \in \mathcal{W}$. Let $(c_k)_{k\geq 0}$ be migration constants such that $s_n \to \infty$ as $n \to \infty$. Then, for all $w \in \mathcal{W}$, uniformly on $\overline{D}$,*

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*. \tag{2.1.54}$$

The convergence in (2.1.54) is a consequence of Lemmas 2.7 and 2.8: The first moment formula (2.1.36) (i) and (2.1.37) show that $K^{w,(n)}_x$ converges to the unique distribution on $\{e_0,\ldots,e_d\}$ with mean x, and by the second moment formula (2.1.36) (ii) this implies the convergence of $s_n F^{(n)} w$.

In order for the iterates in (2.1.54) to be well-defined, Theorem 2.12 *assumes* that a renormalization class $\mathcal{W}$ of diffusion matrices w on $\overline{D}$ with effective boundary $\{e_0,\ldots,e_d\}$ is given. The problem of finding a nontrivial example of such a renormalization class is open in dimensions greater than one. In the one-dimensional case, however, the following result is known.

Lemma 2.13 [DG93b] (Renormalization class on the unit interval) *The set*

$$\mathcal{W}_{\mathrm{DG}} := \{w \in \mathcal{C}[0,1] : w = 0 \text{ on } \{0,1\},\ w > 0 \text{ on } (0,1),\ w \text{ Lipschitz}\} \tag{2.1.55}$$

is a renormalization class on $[0,1]$.

About renormalization of isotropic diffusions, the following result is known. Below, $\partial D := \overline{D}\backslash D$ denotes the topological boundary of D.

Theorem 2.14 [HS98] (Universality class of isotropic models) *Let $D \subset \mathbb{R}^d$ be open, bounded, and convex and let $m \in M_+^d$ be fixed and (strictly) positive definite. Set $w^*_{ij}(x) := m_{ij}g^*(x)$, where g^* is the unique solution of $1 + \frac{1}{2}\sum_{ij} m_{ij}\frac{\partial^2}{\partial x_i \partial x_j} g^*(x) = 0$ for $x \in D$ and $g^*(x) = 0$ for $x \in \partial D$. Assume that $\mathcal{W}$ is a renormalization class on $\overline{D}$ such that $w^* \in \mathcal{W}$ and such that each $w \in \mathcal{W}$ is of the form*

$$w_{ij}(x) = m_{ij}g(x) \qquad (x \in \overline{D},\ i,j = 1,\ldots,d), \tag{2.1.56}$$

for some $g \in \mathcal{C}(\overline{D})$ satisfying $g > 0$ on D and $g = 0$ on ∂D. Let $(c_k)_{k\geq 0}$ be migration constants such that $s_n \to \infty$ as $n \to \infty$. Then, for all $w \in \mathcal{W}$, uniformly on $\overline{D}$,

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*. \tag{2.1.57}$$

The proof of Theorem 2.14 follows the same lines as the proof of Theorem 2.12, with the difference that in this case one needs to generalize the first moment formula (2.1.36) (i) in the sense that $\int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)h(y) = h(x)$ for any m-harmonic function h, i.e., $h \in \mathcal{C}^{(2)}(\overline{D})$ satisfying $\sum_{ij} m_{ij}\frac{\partial^2}{\partial x_i \partial x_j}h(x) = 0$ for $x \in D$. The kernel $K^{w,(n)}_x$ now converges to the m-harmonic measure on ∂D with mean x, and this implies (2.1.57).

Again, in dimensions $d \geq 2$, the problem of finding a 'reasonable' class $\mathcal{W}$ satisfying the assumptions of Theorem 2.14 is so far unresolved. The problem with verifying conditions (i)–(iv) from Definition 2.1 in an explicit set-up is that (i) and (ii) usually require some smoothness of w, while (iv) requires that one can prove the same smoothness for $F_c w$, which is difficult.

The proofs of Theorems 2.12 and 2.14 are both based on invariant harmonics (see (2.1.13)). Since diffusion matrices of catalytic Wright-Fisher diffusions do not in general have invariant harmonics, in order to prove our main result (Theorem 2.17 below), we will need quite different techniques.

Closely related to this is the fact that in the renormalization classes from Theorems 2.12 and 2.14, the unique attraction point w^* does not depend on the parameter γ^* from (2.1.43). As a result, it turns out that the class $\{\lambda w^* : \lambda > 0\}$ is a fixed shape. Here, for any prerenormalization class $\mathcal{W}$, a *fixed shape* is a subclass $\hat{\mathcal{W}} \subset \mathcal{W}$ of the form $\hat{\mathcal{W}} = \{\lambda w : \lambda > 0\}$ with $0 \neq w \in \mathcal{W}$, such that $F_c(\hat{\mathcal{W}}) \subset \hat{\mathcal{W}}$ for all $c > 0$. The next lemma, which will be proved in Section 2.3.1 below, describes how fixed shapes for renormalization classes on compact sets typically arise.

Lemma 2.15 (Fixed shapes) *Assume that for each $0 < \gamma^* < \infty$, there is a $0 \neq w^* = w^*_{\gamma^*} \in \mathcal{W}$ such that $s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*_{\gamma^*}$ whenever $w \in \mathcal{W}$, $s_n \to \infty$, and $s_{n+1}/s_n \to 1+\gamma^*$. Then:*

(a) *$w^*_{\gamma^*}$ is the unique solution in $\mathcal{W}$ of equation (2.1.46) (i).*

(b) *If $w^* = w^*_{\gamma^*}$ does not depend on γ^*, then*

$$F_c(\lambda w^*) = (\tfrac{1}{\lambda} + \tfrac{1}{c})^{-1} w^* \qquad (\lambda, c > 0). \tag{2.1.58}$$

Moreover, $\{\lambda w^* : \lambda > 0\}$ is the unique fixed shape in $\mathcal{W}$.

(c) *If the $w^*_{\gamma^*}$ for different values of γ^* are not constant multiples of each other, then $\mathcal{W}$ contains no fixed shapes.*

In our main result (Theorem 2.17 below), we will describe a renormalization class which we believe contains no fixed shape.

2.2 Catalytic Wright-Fisher diffusions

2.2.1 Main result

Motivated by the previous sections, we will now take the abstract definition of a renormalization class as our starting point, and study iterated renormalization transformations on one such class. Earlier work of this sort has been done in [BCGH95, BCGH97, HS98, Sch98, CDG04]. The subject of our study will be the following renormalization class on $[0,1]^2$.

Definition 2.16 (Renormalization class of catalytic Wright-Fisher diffusions) We set $\mathcal{W}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{H}\}$, where

$$w^{\alpha,p}(x) := \begin{pmatrix} \alpha x_1(1-x_1) & 0 \\ 0 & p(x_1)x_2(1-x_2) \end{pmatrix} \qquad (x = (x_1,x_2) \in [0,1]^2), \tag{2.2.1}$$

and

$$\mathcal{H} := \{p : p \text{ a real function on } [0,1],\ p \geq 0,\ p \text{ Lipschitz continuous}\}. \tag{2.2.2}$$

Moreover, we put

$$\mathcal{H}_{l,r} := \{p \in \mathcal{H} :\ 1_{\{p(0)>0\}} = l,\ 1_{\{p(1)>0\}} = r\} \qquad (l,r = 0,1), \tag{2.2.3}$$

and set $\mathcal{W}^{l,r}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{H}_{l,r}\} \quad (l,r = 0,1)$. $\Diamond$

Solutions $\mathbf{y} = (\mathbf{y}^1, \mathbf{y}^2)$ to the martingale problem for $A^{c,w^{\alpha,p}}_x$ (recall (2.1.24)) can be represented as solutions to the SDE

$$\begin{aligned} &\text{(i)} \quad \mathrm{d}\mathbf{y}^1_t = c\,(x_1 - \mathbf{y}^1_t)\mathrm{d}t + \sqrt{2\alpha \mathbf{y}^1_t(1-\mathbf{y}^1_t)}\,\mathrm{d}B^1_t, \\ &\text{(ii)} \quad \mathrm{d}\mathbf{y}^2_t = c\,(x_2 - \mathbf{y}^2_t)\mathrm{d}t + \sqrt{2p(\mathbf{y}^1_t)\mathbf{y}^2_t(1-\mathbf{y}^2_t)}\,\mathrm{d}B^2_t. \end{aligned} \tag{2.2.4}$$

We call $\mathbf{y}^1$ the Wright-Fisher *catalyst* with *resampling rate* α and $\mathbf{y}^2$ the Wright-Fisher *reactant* with *catalyzing function* p.

Here is our main result:

Theorem 2.17 (Main result)
(a) *The set* $\mathcal{W}_{\rm cat}$ *is a renormalization class on* $[0,1]^2$ *and* $F_c(\mathcal{W}^{l,r}_{\rm cat}) \subset \mathcal{W}^{l,r}_{\rm cat}$ $(c>0,\ l,r=0,1)$.

(b) *Fix (positive) migration constants* $(c_k)_{k\geq 0}$ *such that*

$$\text{(i)}\quad s_n \underset{n\to\infty}{\longrightarrow} \infty \qquad \textit{and} \qquad \text{(ii)}\quad \frac{s_{n+1}}{s_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^* \tag{2.2.5}$$

for some $\gamma^*\geq 0$. *If* $w\in\mathcal{W}^{l,r}_{\rm cat}$ $(l,r=0,1)$, *then uniformly on* $[0,1]^2$,

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*, \tag{2.2.6}$$

where the limit w^* *is the unique solution in* $\mathcal{W}^{l,r}_{\rm cat}$ *to the equation*

$$\begin{array}{llll} \text{(i)} & (1+\gamma^*)F_{1/\gamma^*}w^* = w^* & & \textit{if}\ \ \gamma^*>0,\\ \text{(ii)} & \frac{1}{2}\sum_{i,j=1}^{2} w^*_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}w^*(x)+w^*(x)=0 & (x\in[0,1]^2) & \textit{if}\ \ \gamma^*=0. \end{array} \tag{2.2.7}$$

(c) *The matrix* w^* *is of the form* $w^*=w^{1,p^*}$, *where* $p^*=p^*_{l,r,\gamma^*}\in\mathcal{H}_{l,r}$ *depends on* l,r, *and* γ^*. *One has*

$$p^*_{0,0,\gamma^*}\equiv 0 \quad \textit{and} \quad p^*_{1,1,\gamma^*}\equiv 1 \qquad \textit{for all}\ \gamma^*\geq 0. \tag{2.2.8}$$

For each $\gamma^*\geq 0$, *the function* $p^*_{0,1,\gamma^*}$ *is concave, nondecreasing, and satisfies* $p^*_{0,1,\gamma^*}(0)=0$, $p^*_{0,1,\gamma^*}(1)=1$. *By symmetry, analogous statements hold for* $p^*_{1,0,\gamma^*}$.

Conditions (2.2.5) (i) and (ii) are satisfied, for example, for $c_k=(1+\gamma^*)^{-k}$. Note that the functions $p^*_{0,0,\gamma^*}$ and $p^*_{1,1,\gamma^*}$ are independent of $\gamma^*\geq 0$. We believe that on the other hand, $p^*_{0,1,\gamma^*}$ is not constant as a function of γ^*, but we have not proved this.[2] If this is confirmed, then by Lemma 2.15, it follows that $\mathcal{W}^{0,1}_{\rm cat}$, unlike all renormalization classes studied previously, contains no fixed shapes.

The function $p^*_{0,1,0}$ is the unique nonnegative solution to the equation

$$\tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p(x)+p(x)(1-p(x))=0 \qquad (x\in[0,1]) \tag{2.2.9}$$

with boundary conditions $p(0)=0$ and $p(1)>0$. This function occurred before in the work of Greven, Klenke, and Wakolbinger [GKW01, formulas (1.10)–(1.11)], who studied linearly interacting catalytic Wright-Fisher diffusions catalyzed by a voter model. They believe their results to hold for a Wright-Fisher catalyst too, i.e., for a model of the form

$$\begin{aligned} \mathrm{d}\mathbf{x}^1_\xi(t) &= \sum_{\eta:\,|\eta-\xi|=1}\big(\mathbf{x}^1_\eta(t)-\mathbf{x}^1_\xi(t)\big)\,\mathrm{d}t+\sqrt{2\alpha\mathbf{x}^1_\xi(t)(1-\mathbf{x}^1_\xi(t))}\,\mathrm{d}B^1_\xi(t),\\ \mathrm{d}\mathbf{x}^2_\xi(t) &= \sum_{\eta:\,|\eta-\xi|=1}\big(\mathbf{x}^2_\eta(t)-\mathbf{x}^2_\xi(t)\big)\,\mathrm{d}t+\sqrt{2p(\mathbf{x}^1_\xi(t))\mathbf{x}^2_\xi(t)(1-\mathbf{x}^2_\xi(t))}\,\mathrm{d}B^2_\xi(t), \end{aligned} \tag{2.2.10}$$

[2] In support of this, if $\mathcal{U}_\gamma$ $(\gamma>0)$ are transformations such that $\overline{F}^{1,p}_\gamma=w^{1,\mathcal{U}_\gamma p}$ (see (2.2.21) below), then a heuristic calculation for $p=p^*_{0,1,0}$ yields $\mathcal{U}_\gamma p(x)=p(x)+\gamma^2x(1-x)\{\frac{1}{2}p''(x)-\frac{4}{3}(p'(x))^2-\frac{4}{3}xp'''(x)\}+O(\gamma^3)$, which implies that $p^*_{0,1,0}\neq p^*_{0,1,\gamma^*}$ for γ^* small enough.

where $\alpha > 0$ is a constant, p is a nonnegative function on $[0,1]$ satisfying $p(0) = 0$ and $p(1) > 0$, but they could not prove this due to certain technical difficulties that a $[0,1]$-valued catalyst would create, compared to the simpler $\{0,1\}$-valued voter model. They determined the clustering distribution of their model on $\mathbb{Z}^2$, which turns out to coincide with the prediction made based on renormalization theory in Conjecture 2.11, with $w^* = w^{1,p^*_{0,1,0}}$ as in our Theorem 2.17.

The work in [GKW01] not only provides the main motivation for the present chapter, but also inspired some of our techniques for proving Theorem 2.17. This concerns in particular the proof of Proposition 2.18 below, which makes the connection between renormalization transformations and a branching process. We hope that conversely, our techniques may shed some light on the problems left open by [GKW01], in particular, the question whether their results stay true if the voter model catalyst is replaced by a Wright-Fisher catalyst. It seems plausible that their results may not hold for the model in (2.2.10) if the catalyzing function p grows too fast at 0. On the other hand, our proofs suggest that p with a finite slope at 0 should be OK. (In particular, while deriving formula (2.2.51) below, we use that p can be bounded from above by $r_+h_{0,1}$ for some $r_+ > 0$, which requires that p has a finite slope at 0.)

2.2.2 Open problems

The general program of studying renormalization classes in the sense of Definition 2.1 contains a wealth of open problems. In our proofs, we make heavy use of the single-way nature of the catalyzation in (2.2.4), in particular, the fact that $\mathbf{y}^1$ is an autonomous process which allows one to condition on $\mathbf{y}^1$ and consider $\mathbf{y}^2$ as a process in a random environment created by $\mathbf{y}^1$. As soon as one leaves the single-way catalytic regime one runs into several difficulties, both technically (it is hard to prove that a given class of matrices is a renormalization class in the sense of Definition 2.1) and conceptually (it is not clear when solutions to the asymptotic fixed shape equation (2.1.46) (ii) are unique). Therefore, it seems at present hard to verify the complete picture for renormalization classes on the unit square that arises from the numerical simulations described in Section 2.1.7 and Figures 2.1 and 2.2, unless one or more essential new ideas are added.

In this context, the study of the nonlinear partial differential equation (2.1.53) and its fixed points seems to be a challenging problem. This may be a hard problem from an analytic point of view, since the equation is degenerate and not in divergence form. For the renormalization class $\mathcal{W}_{\rm cat}$, the quasilinear equation (2.1.53) reduces to the semilinear equation (2.2.26), which is analytically easier to treat and moreover has a probabilistic interpretation in terms of a superprocess. We do not know whether solutions to equation (2.1.53) can in general be represented in terms of a stochastic process of some sort.

Even for the renormalization class $\mathcal{W}_{\rm cat}$, several interesting problems are left open. One of the most urgent ones is to prove that the functions $p^*_{0,1,\gamma^*}$ are not constant in γ^*, and therefore, by Lemma 2.15 (c), $\mathcal{W}^{0,1}_{\rm cat}$ contains no fixed shapes. Moreover, we have not investigated the iterated renormalization transformations in the regime $\gamma^* = \infty$. Also, we believe that the convergence in (2.2.39) (ii) does not hold if the condition that p is Lipschitz is dropped, in particular, if $p(0) = 0$ and p has an infinite slope at 0. For $p \in \mathcal{H}_{0,0}$, it seems plausible that a properly rescaled version of the iterates $\mathcal{U}^{(n)}p$, with $\mathcal{U}_\gamma$ as in (2.2.20) below, converges to a universal limit, but we have not investigated this either. Finally, we have not investigated the convergence of the iterated kernels $K^{w,(n)}$ from (2.1.33) (in particular, we have not verified Conjecture 2.10) for the renormalization class $\mathcal{W}_{\rm cat}$.

Our methods, combined with those in [BCGH95], can probably be extended to study the action of iterated renormalization transformations on diffusion matrices of the following more general form (compared to (2.2.1)):

$$w(x) = \begin{pmatrix} g(x_1) & 0 \\ 0 & p(x_1)x_2(1-x_2) \end{pmatrix} \qquad (x = \in [0,1]^2), \tag{2.2.11}$$

where $g : [0,1] \to \mathbb{R}$ is Lipschitz, $g(0) = g(1) = 0$, $g > 0$ on $(0,1)$, and $p \in \mathcal{H}$ as before. This would, however, require a lot of extra technical work and probably not generate much new insight. The numerical simulations mentioned in Section 2.1.7 suggest that many diffusion matrices of an even more general form than (2.2.11) also converge under renormalization to the limit points w^* from Theorem 2.17, but we don't know how to prove this.

In the next sections, we will show that for the renormalization class $\mathcal{W}_{\rm cat}$, the rescaled renormalization transformations $\overline{F}_\gamma$ from (2.1.39) can be expressed in terms of the log-Laplace operators of a discrete time branching process on $[0,1]$. This will allow us to use techniques from the theory of spatial branching processes to verify Conjecture 2.9 for the renormalization class $\mathcal{W}_{\rm cat}$ in the case $\gamma^* < \infty$.

2.2.3 Poisson-cluster branching processes

We first need some concepts and facts from branching theory. Finite measure-valued branching processes (on $\mathbb{R}$) in discrete time have been introduced by Jiřina [Jir64]. We need to consider only a special class.

Let E be a separable, locally compact, and metrizable space. We let $\mathcal{C}(E)$ and $B(E)$ denote the spaces of all continuous, and bounded Borel measurable, real functions on E, respectively. We put $\mathcal{C}_+(E) := \{f \in \mathcal{C}(E) : f \geq 0\}$ and define $B_+(E)$ analogously. We let $\mathcal{M}(E)$ denote the space of all finite measures on E, equipped with the topology of weak convergence. The subspace of probability measures is denoted by $\mathcal{M}_1(E)$. For $\mu \in \mathcal{M}(E)$ and $f \in B(E)$ we use the notation $\langle \mu, f\rangle := \int_E f\,\mathrm{d}\mu$ and $|\mu| := \mu(E)$.

We call a continuous map $\mathcal{Q}$ from E into $\mathcal{M}_1(\mathcal{M}(E))$ a *continuous cluster mechanism.* By definition, an $\mathcal{M}(E)$-valued random variable $\mathcal{X}$ is a *Poisson cluster measure* on E with locally finite *intensity measure* μ and continuous cluster mechanism $\mathcal{Q}$, if its log-Laplace transform satisfies

$$-\log E\big[e^{-\langle \mathcal{X}, f\rangle}\big] = \int_E \mu(\mathrm{d}x)\Big(1 - \int_{\mathcal{M}(E)} \mathcal{Q}(x,\mathrm{d}\chi)\,e^{-\langle \chi, f\rangle}\Big) \quad (f \in B_+(E)). \tag{2.2.12}$$

For given μ and $\mathcal{Q}$, such a Poisson cluster measure exists, and is unique in distribution, provided that the right-hand side of (2.2.12) is finite for $f = 1$. It may be constructed as $\mathcal{X} = \sum_i \chi_{x_i}$, where $\sum_i \delta_{x_i}$ is a (possibly infinite) Poisson point measure with intensity μ, and given $x_1, x_2, \ldots$, the $\chi_{x_1}, \chi_{x_2}, \ldots$ are independent random variables with laws $\mathcal{Q}(x_1, \cdot\,), \mathcal{Q}(x_2, \cdot\,), \ldots$, respectively.

Now fix a finite sequence of functions $q_k \in \mathcal{C}_+(E)$ and continuous cluster mechanisms $\mathcal{Q}_k$ $(k = 1, \ldots, n)$, define

$$\mathcal{U}_k f(x) := q_k(x)\Big(1 - \int_{\mathcal{M}(E)} \mathcal{Q}_k(x, \mathrm{d}\chi)\, e^{-\langle \chi, f\rangle}\Big) \qquad (x \in E,\ f \in B_+(E),\ k = 1, \ldots, n), \tag{2.2.13}$$

and assume that

$$\sup_{x\in E} \mathcal{U}_k 1(x) < \infty \qquad (k = 1, \ldots, n). \tag{2.2.14}$$

Then $\mathcal{U}_k$ maps $B_+(E)$ into $B_+(E)$ for each k, and for each $\mathcal{M}(E)$-valued initial state $\mathcal{X}_0$, there exists a (time-inhomogeneous) Markov chain $(\mathcal{X}_0, \ldots, \mathcal{X}_n)$ in $\mathcal{M}(E)$, such that $\mathcal{X}_k$, given $\mathcal{X}_{k-1}$, is a Poisson cluster measure with intensity $q_k\mathcal{X}_{k-1}$ and cluster mechanism $\mathcal{Q}_k$. It is not hard to see that the process started in μ satisfies

$$E^\mu\big[e^{-\langle \mathcal{X}_n, f\rangle}\big] = e^{-\langle \mu, \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_n f\rangle} \qquad (\mu \in \mathcal{M}(E),\ f \in B_+(E)). \tag{2.2.15}$$

We call $\mathcal{X} = (\mathcal{X}_0, \ldots, \mathcal{X}_n)$ the *Poisson-cluster branching process* on E with *weight functions* $q_1, \ldots, q_n$ and cluster mechanisms $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$. The operator $\mathcal{U}_k$ is called the *log-Laplace operator* of the transition law from $\mathcal{X}_{k-1}$ to $\mathcal{X}_k$. Note that we can write (2.2.15) in the suggestive form

$$P^\mu\big[\mathrm{Pois}(f\mathcal{X}_n) = 0\big] = P\big[\mathrm{Pois}\big((\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_n f)\mu\big) = 0\big]. \tag{2.2.16}$$

Here, if μ is an $\mathcal{M}(E)$-valued random variable, then $\mathrm{Pois}(\mu)$ denotes an $\mathcal{N}(E)$-valued random variable such that conditioned on μ, $\mathrm{Pois}(\mu)$ is a Poisson point measure with intensity μ.

2.2.4 The renormalization branching process

We will now construct a Poisson-cluster branching process on $[0, 1]$ of a special kind, and show that the rescaled renormalization transformations on $\mathcal{W}_{\mathrm{cat}}$ can be expressed in terms of the log-Laplace operators of this branching process.

By Lemma 2.30 below, for each $\gamma > 0$ and $x \in [0, 1]$, the SDE

$$\mathrm{d}\mathbf{y}(t) = \tfrac{1}{\gamma}\,(x - \mathbf{y}(t))\mathrm{d}t + \sqrt{2\mathbf{y}(t)(1 - \mathbf{y}(t))}\,\mathrm{d}B(t), \tag{2.2.17}$$

has a unique (in law) stationary solution. We denote this solution by $(\mathbf{y}^\gamma_x(t))_{t\in\mathbb{R}}$. Let τ_γ be an independent exponentially distributed random variable with mean γ, and set

$$\mathcal{Z}^\gamma_x := \int_0^{\tau_\gamma} \delta_{\mathbf{y}^\gamma_x(-t/2)}\mathrm{d}t \qquad (\gamma > 0,\ x \in [0, 1]). \tag{2.2.18}$$

Define constants q_γ and continuous (by Corollary 2.36 below) cluster mechanisms $\mathcal{Q}_\gamma$ by

$$q_\gamma := \tfrac{1}{\gamma} + 1 \qquad \text{and} \qquad \mathcal{Q}_\gamma(x, \cdot\,) := \mathcal{L}(\mathcal{Z}^\gamma_x) \qquad (\gamma > 0,\ x \in [0, 1]), \tag{2.2.19}$$

and let $\mathcal{U}_\gamma$ denote the log-Laplace operator with (constant) weight function q_γ and cluster mechanism $\mathcal{Q}_\gamma$, i.e.,

$$\mathcal{U}_\gamma f(x) := q_\gamma\Big(1 - \int_{\mathcal{M}([0,1])} \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\, e^{-\langle \chi, f\rangle}\Big) \qquad (x \in [0, 1],\ f \in B_+[0, 1],\ \gamma > 0). \tag{2.2.20}$$

We now establish the connection between renormalization transformations on $\mathcal{W}_{\rm cat}$ and log-Laplace operators.

Proposition 2.18 (Identification of the renormalization transformation) *Let $\overline{F}_\gamma$ be the rescaled renormalization transformation on $\mathcal{W}_{\rm cat}$ defined in (2.1.39). Then*

$$\overline{F}_\gamma w^{1,p} = w^{1,\mathcal{U}_\gamma p} \qquad (p \in \mathcal{H},\ \gamma > 0). \tag{2.2.21}$$

Fix a diffusion matrix $w^{\alpha,p} \in \mathcal{W}_{\rm cat}$ and migration constants $(c_k)_{k\geq 0}$. Define constants $\overline{s}_n$ and γ_n as in (2.1.38) and (2.1.41), respectively, where $\beta := 1/\alpha$. Then Proposition 2.18 and formula (2.1.40) show that

$$\overline{s}_n F^{(n)} w^{\alpha,p} = w^{1,\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}(\frac{p}{\alpha})}. \tag{2.2.22}$$

Here $\mathcal{U}_{\gamma_{n-1}}, \ldots, \mathcal{U}_{\gamma_0}$ are the log-Laplace operators of the Poisson-cluster branching process $\mathcal{X} = (\mathcal{X}_{-n}, \ldots, \mathcal{X}_0)$ with weight functions $q_{\gamma_{n-1}}, \ldots, q_{\gamma_0}$ and cluster mechanisms $\mathcal{Q}_{\gamma_{n-1}}, \ldots, \mathcal{Q}_{\gamma_0}$. We call $\mathcal{X}$ (started at some time $-n$ in an initial law $\mathcal{L}(\mathcal{X}_{-n})$) the *renormalization branching process*. By formulas (2.2.15) and (2.2.22), the study of the limiting behavior of rescaled iterated renormalization transformations on $\mathcal{W}_{\rm cat}$ reduces to the study of the renormalization branching process $\mathcal{X}$ in the limit $n \to \infty$.

2.2.5 Convergence to a time-homogeneous process

Let $\mathcal{X} = (\mathcal{X}_{-n}, \ldots, \mathcal{X}_0)$ be the renormalization branching process introduced in the last section. If the constants $(\gamma_k)_{k\geq 0}$ satisfy $\sum_n \gamma_n = \infty$ and $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$, then $\mathcal{X}$ is almost time-homogeneous for large n. More precisely, we will prove the following convergence result.

Theorem 2.19 (Convergence to a time-homogenous branching process) *Assume that $\mathcal{L}(\mathcal{X}_{-n}) \underset{n\to\infty}{\Longrightarrow} \mu$ for some probability law μ on $\mathcal{M}([0,1])$.*

(a) *If $0 < \gamma^* < \infty$, then*

$$\mathcal{L}(\mathcal{X}_{-n}, \mathcal{X}_{-n+1}, \ldots) \underset{n\to\infty}{\Longrightarrow} \mathcal{L}(\mathcal{Y}^{\gamma^*}_0, \mathcal{Y}^{\gamma^*}_1, \ldots), \tag{2.2.23}$$

where $\mathcal{Y}^{\gamma^*}$ is the time-homogenous branching process with log-Laplace operator $\mathcal{U}_{\gamma^*}$ in each step and initial law $\mathcal{L}(\mathcal{Y}^{\gamma^*}_0) = \mu$.

(b) *If $\gamma^* = 0$, then*

$$\mathcal{L}\Big(\big(\mathcal{X}_{-k_n(t)}\big)_{t\geq 0}\Big) \underset{n\to\infty}{\Longrightarrow} \mathcal{L}\Big(\big(\mathcal{Y}^0_t\big)_{t\geq 0}\Big), \tag{2.2.24}$$

where $\Rightarrow$ denotes weak convergence of laws on path space, $k_n(t) := \min\{k : 0 \leq k \leq n, \sum_{l=k}^{n-1} \gamma_l \leq t\}$, and $\mathcal{Y}^0$ is the superprocess on $[0,1]$ with underlying motion generator $\frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}$ and activity and growth parameter both identically 1, started in the initial law $\mathcal{L}(\mathcal{Y}^0_0) = \mu$.

We call the superprocess $\mathcal{Y}^0$ from part (b) the *super-Wright-Fisher diffusion.* It is the time-homogeneous Markov process in $\mathcal{M}[0,1]$ with continuous sample paths, whose Laplace functionals are given by

$$E^\mu\big[e^{-\langle \mathcal{Y}^0_t, f\rangle}\big] = e^{-\langle \mu, \mathcal{U}^0_t f\rangle} \qquad (\mu \in \mathcal{M}[0,1],\ f \in B_+[0,1],\ t \geq 0), \tag{2.2.25}$$

where $\mathcal{U}^0_t f = u_t$ is the unique mild solution of the semilinear Cauchy equation

$$\begin{cases} \frac{\partial}{\partial t} u_t(x) = \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}u_t(x) + u_t(x)(1-u_t(x)) & (t \geq 0,\ x \in [0,1]), \\ u_0 = f. \end{cases} \tag{2.2.26}$$

For a further study of the renormalization branching process $\mathcal{X}$ and its limiting processes $\mathcal{Y}^{\gamma^*}$ ($\gamma^* \geq 0$) we will use the technique of embedded particle systems, which we explain in the next section.

2.2.6 Weighted and Poissonized branching processes

In this section, we explain how from a Poisson-cluster branching process it is possible to construct other branching processes by weighting and Poissonization. We first need to introduce spatial branching particle systems in some generality.

Let E again be separable, locally compact, and metrizable. We set $\mathcal{C}_{[0,1]}(E) := \{f \in \mathcal{C}(E) : 0 \leq f \leq 1\}$ and define $B_{[0,1]}(E)$ analogously. We write $\mathcal{N}(E)$ for the space of finite counting measures, i.e., measures of the form $\nu = \sum_{i=1}^m \delta_{x_i}$ with $x_1, \ldots, x_m \in E$ ($m \geq 0$). We interpret ν as a collection of particles, situated at positions $x_1, \ldots, x_m$. For $\nu \in \mathcal{N}(E)$ and $f \in B_{[0,1]}(E)$, we adopt the notation

$$f^0 := 1 \quad \text{and} \quad f^\nu := \prod_{i=1}^m f(x_i) \quad \text{when} \quad \nu = \sum_{i=1}^m \delta_{x_i} \quad (m \geq 1). \tag{2.2.27}$$

We call a continuous map $x \mapsto Q(x, \cdot\,)$ from E into $\mathcal{M}_1(\mathcal{N}(E))$ a *continuous offspring mechanism.*

Fix continuous offspring mechanisms Q_k ($1 \leq k \leq n$), and let $(X_0, \ldots, X_n)$ be a Markov chain in $\mathcal{N}(E)$ such that, given that $X_{k-1} = \sum_{i=1}^m \delta_{x_i}$, the next step of the chain X_k is a sum of independent random variables with laws $Q_k(x_i, \cdot\,)$ ($i = 1, \ldots, m$). Then

$$E^\nu\big[(1-f)^{X_n}\big] = (1 - U_1 \circ \cdots \circ U_n f)^\nu \qquad (\nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)), \tag{2.2.28}$$

where $U_k : B_{[0,1]}(E) \to B_{[0,1]}(E)$ is defined as

$$U_k f(x) := 1 - \int_{\mathcal{N}(E)} Q^k(x, \mathrm{d}\nu)(1-f)^\nu \qquad (1 \leq k \leq n,\ x \in E,\ f \in B_{[0,1]}(E)). \tag{2.2.29}$$

We call U_k the *generating operator* of the transition law from X_{k-1} to X_k, and we call $X = (X_0, \ldots, X_n)$ the *branching particle system* on E with generating operators $U_1, \ldots, U_n$. It is often useful to write (2.2.28) in the suggestive form

$$P^\nu\big[\mathrm{Thin}_f(X_n) = 0\big] = P\big[\mathrm{Thin}_{U_1 \circ \cdots \circ U_n f}(\nu) = 0\big] \qquad (\nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)). \tag{2.2.30}$$

Here, if ν is an $\mathcal{N}(E)$-valued random variable and $f \in B_{[0,1]}(E)$, then $\mathrm{Thin}_f(\nu)$ denotes an $\mathcal{N}(E)$-valued random variable such that conditioned on ν, $\mathrm{Thin}_f(\nu)$ is obtained from ν by independently throwing away particles from ν, where a particle at x is kept with probability $f(x)$. One has the elementary relations

$$\mathrm{Thin}_f(\mathrm{Thin}_g(\nu)) \stackrel{\mathcal{D}}{=} \mathrm{Thin}_{fg}(\nu) \quad \text{and} \quad \mathrm{Thin}_f(\mathrm{Pois}(\mu)) \stackrel{\mathcal{D}}{=} \mathrm{Pois}(f\mu), \tag{2.2.31}$$

where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution.

We are now ready to describe weighted and Poissonized branching processes. Let $\mathcal{X} = (\mathcal{X}_0, \ldots, \mathcal{X}_n)$ be a Poisson-cluster branching process on E, with continuous weight functions $q_1, \ldots, q_n$, continuous cluster mechanisms $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$, and log-Laplace operators $\mathcal{U}_1, \ldots, \mathcal{U}_n$ given by (2.2.13) and satisfying (2.2.14). Let $\mathcal{Z}^k_x$ denote an $\mathcal{M}(E)$-valued random variable with law $\mathcal{Q}_k(x, \cdot\,)$. Let $h \in \mathcal{C}_+(E)$ be bounded, $h \neq 0$, and put $E^h := \{x \in E : h(x) > 0\}$. For $f \in B_+(E^h)$, define $hf \in B_+(E)$ by $hf(x) := h(x)f(x)$ if $x \in E^h$ and $hf(x) := 0$ otherwise.

Proposition 2.20 (Weighting of Poisson-cluster branching processes) *Assume that there exists a constant $K < \infty$ such that $\mathcal{U}_k h \leq Kh$ for all $k = 1, \ldots, n$. Then there exists a Poisson-cluster branching process $\mathcal{X}^h = (\mathcal{X}^h_0, \ldots, \mathcal{X}^h_n)$ on E^h with weight functions $(q^h_1, \ldots, q^h_n)$ given by $q^h_k := q_k/h$, continuous cluster mechanisms $\mathcal{Q}^h_1, \ldots, \mathcal{Q}^h_n$ given by*

$$\mathcal{Q}^h_k(x, \cdot\,) := \mathcal{L}(h\mathcal{Z}^k_x) \qquad (x \in E^h), \tag{2.2.32}$$

and log-Laplace operators $\mathcal{U}^h_1, \ldots, \mathcal{U}^h_n$ satisfying

$$h\mathcal{U}^h_k f := \mathcal{U}_k(hf) \qquad (f \in B_+(E^h)). \tag{2.2.33}$$

The processes $\mathcal{X}$ and X^h are related by

$$\mathcal{L}(\mathcal{X}^h_0) = \mathcal{L}(h\mathcal{X}_0) \quad \textit{implies} \quad \mathcal{L}(\mathcal{X}^h_k) = \mathcal{L}(h\mathcal{X}_k) \qquad (0 \leq k \leq n). \tag{2.2.34}$$

Proposition 2.21 (Poissonization of Poisson-cluster branching processes) *Assume that $\mathcal{U}_k h \leq h$ for all $k = 1, \ldots, n$. Then there exists a branching particle system $X^h = (X^h_0, \ldots, X^h_n)$ on E^h with continuous offspring mechanisms $Q^h_1, \ldots, Q^h_n$ given by*

$$Q^h_k(x, \cdot\,) := \frac{q_k(x)}{h(x)} P\big[\mathrm{Pois}(h\mathcal{Z}^k_x) \in \cdot\,\big] + \Big(1 - \frac{q_k(x)}{h(x)}\Big)\delta_0(\,\cdot\,) \qquad (x \in E^h), \tag{2.2.35}$$

and generating operators $U^h_1, \ldots, U^h_n$ satisfying

$$hU^h_k f := \mathcal{U}_k(hf) \qquad (f \in B_{[0,1]}(E^h)). \tag{2.2.36}$$

The processes $\mathcal{X}$ and X^h are related by

$$\mathcal{L}(X^h_0) = \mathcal{L}(\mathrm{Pois}(h\mathcal{X}_0)) \quad \textit{implies} \quad \mathcal{L}(X^h_k) = \mathcal{L}(\mathrm{Pois}(h\mathcal{X}_k)) \qquad (0 \leq k \leq n). \tag{2.2.37}$$

Here, the right-hand side of (2.2.35) is always a probability measure, despite that it may happen that $q_k(x)/h(x) > 1$. The (straightforward) proofs of Propositions 2.20 and 2.21 can be found in Section 2.8.1 below. If (2.2.34) holds then we say that $\mathcal{X}^h$ is obtained from $\mathcal{X}$ by *weighting* with density h. If (2.2.37) holds then we say that X^h is obtained from $\mathcal{X}$ by *Poissonization* with density h. Proposition 2.21 says that a Poisson-cluster branching process $\mathcal{X}$ contains, in a way, certain 'embedded' branching particle systems X^h. Poissonization relations for superprocesses and embedded particle systems have enjoyed considerable attention, see [FS04] and references therein.

A function $h \in B_+(E)$ such that $\mathcal{U}_k h \leq h$ is called *$\mathcal{U}_k$-superharmonic.* If the reverse inequality holds we say that h is *$\mathcal{U}_k$-subharmonic.* If $\mathcal{U}_k h = h$ then h is called *$\mathcal{U}_k$-harmonic.*

2.2.7 Extinction versus unbounded growth for embedded particle systems

In this section we explain how embedded particle systems can be used to prove Theorem 2.17. Throughout this section $(\gamma_k)_{k\geq 0}$ are positive constants such that $\sum_n \gamma_n = \infty$ and $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$, and $\mathcal{X} = (\mathcal{X}_{-n},\ldots,\mathcal{X}_0)$ is the renormalization branching process on $[0,1]$ defined in Section 2.2.4. We write

$$\mathcal{U}^{(n)} := \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}. \tag{2.2.38}$$

In view of formula (2.2.22), in order to prove Theorem 2.17, we need the following result.

Proposition 2.22 (Limits of iterated log-Laplace operators) *Uniformly on* $[0,1]$,

$$\begin{array}{llr} \text{(i)} & \lim_{n\to\infty} \mathcal{U}^{(n)} p = 1 & (p \in \mathcal{H}_{1,1}), \\ \text{(ii)} & \lim_{n\to\infty} \mathcal{U}^{(n)} p = 0 & (p \in \mathcal{H}_{0,0}), \\ \text{(iii)} & \lim_{n\to\infty} \mathcal{U}^{(n)} p = p^*_{0,1,\gamma^*} & (p \in \mathcal{H}_{0,1}), \end{array} \tag{2.2.39}$$

where $p^*_{0,1,\gamma^*} : [0,1] \to [0,1]$ *is a function depending on* γ^* *but not on* $p \in \mathcal{H}_{0,1}$.

In our proof of Proposition 2.22, we will use embedded particle systems $X^h = (X^h_{-n},\ldots,X^h_0)$ obtained from $\mathcal{X}$ by Poissonization with certain h taken from the classes $\mathcal{H}_{1,1}$, $\mathcal{H}_{0,0}$, and $\mathcal{H}_{0,1}$. Below, P^{-n,δ_x} denotes the law of the process started at time $-n$ with one particle at x.

Lemma 2.23 (Embedded particle system with $h_{1,1}$) *The constant function* $h_{1,1}(x) := 1$ *is* $\mathcal{U}_\gamma$*-harmonic for each* $\gamma > 0$. *The corresponding embedded particle system* $X^{h_{1,1}}$ *on* $[0,1]$ *satisfies*

$$P^{-n,\delta_x}\big[|X_0^{h_{1,1}}| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \delta_\infty \tag{2.2.40}$$

uniformly[3] *for all* $x \in [0,1]$.

In (2.2.40) and similar formulas below, $\Rightarrow$ denotes weak convergence of probability measures on $[0,\infty]$. Thus, (2.2.40) says that for processes started with one particle on the position x at times $-n$, the number of particles at time zero converges to infinity as $n \to \infty$.

[3]Since $\mathcal{M}_1[0,\infty]$ is compact in the topology of weak convergence, there is a unique uniform structure compatible with the topology, and therefore we can unambiguously talk about uniform convergence of $\mathcal{M}_1[0,\infty]$-valued functions (in this case, $x \mapsto P^{-n,\delta_x}\big[|X_0^{h_{1,1}}| \in \cdot\,\big]$).

Lemma 2.24 (Embedded particle system with $h_{0,0}$) *The function $h_{0,0}(x) := x(1-x)$ $(x \in [0,1])$ is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. The corresponding embedded particle system $X^{h_{0,0}}$ on $(0,1)$ is critical and satisfies*

$$P^{-n,\delta_x}\left[|X_0^{h_{0,0}}| \in \cdot\,\right] \underset{n\to\infty}{\Longrightarrow} \delta_0 \tag{2.2.41}$$

locally uniformly for all $x \in (0,1)$.

Here, we say that a branching particle system X is *critical* if each particle produces on average one offspring (in each time step and independent of its position). Formula (2.2.41) says that the embedded particle system $X^{h_{0,0}}$ gets extinct during the time interval $\{-n,\ldots,0\}$ with probability tending to one as $n \to \infty$. We can summarize Lemmas 2.23 and 2.24 by saying that the embedded particle system associated with $h_{1,1}$ grows unboundedly while the embedded particle system associated with $h_{0,0}$ becomes extinct as $n \to \infty$.

We will also consider an embedded particle system $X^{h_{0,1}}$ for a certain $h_{0,1}$ taken from $\mathcal{H}_{0,1}$. It turns out that this system either gets extinct or grows unboundedly, each with a positive probability. In order to determine these probabilities, we need to consider embedded particle systems for the time-homogeneous processes $\mathcal{Y}^{\gamma^*}$ $(\gamma^* \in [0,\infty))$ from (2.2.23) and (2.2.24). If $h \in \mathcal{H}_{0,1}$ is $\mathcal{U}_{\gamma^*}$-superharmonic for some $\gamma^* > 0$, then Poissonizing the process $\mathcal{Y}^{\gamma^*}$ with h yields a branching particle system on $(0,1]$ which we denote by $Y^{\gamma^*,h} = (Y_0^{\gamma^*,h}, Y_1^{\gamma^*,h},\ldots)$. Likewise, if $h \in \mathcal{H}_{0,1}$ is twice continuously differentiable and satisfies

$$\tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}h(x) - h(x)(1-h(x)) \le 0, \tag{2.2.42}$$

then Poissonizing the super-Wright-Fisher diffusion $\mathcal{Y}^0$ with h yields a continuous-time branching particle system on $(0,1]$, which we denote by $Y^{0,h} = (Y_t^{0,h})_{t\ge0}$. For example, for $m \ge 4$, the function $h(x) := 1-(1-x)^m$ satisfies (2.2.42).

Lemma 2.25 (Embedded particle system with $h_{0,1}$) *The function $h_{0,1}(x) := 1-(1-x)^7$ is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. The corresponding embedded particle system $X^{h_{0,1}}$ on $(0,1]$ satisfies*

$$P^{-n,\delta_x}\left[|X_0^{h_{0,1}}| \in \cdot\,\right] \underset{n\to\infty}{\Longrightarrow} \rho_{\gamma^*}(x)\delta_\infty + (1-\rho_{\gamma^*}(x))\delta_0, \tag{2.2.43}$$

locally uniformly for all $x \in (0,1]$, where

$$\rho_{\gamma^*}(x) := \begin{cases} P^{\delta_x}[Y_k^{\gamma^*,h_{0,1}} \neq 0 \ \forall k \ge 0] & (0 < \gamma^* < \infty), \\ P^{\delta_x}[Y_t^{0,h_{0,1}} \neq 0 \ \forall t \ge 0] & (\gamma^* = 0). \end{cases} \tag{2.2.44}$$

We now explain how Lemmas 2.23–2.25 imply Proposition 2.22. In doing so, it will be more convenient to work with weighted branching processes than with Poissonized branching processes. A little argument (which can be found in Lemma 2.79 below) shows that Lemmas 2.23–2.25 are equivalent to the next proposition.

Proposition 2.26 (Extinction versus unbounded growth) *Let $h_{1,1}$, $h_{0,0}$, and $h_{0,1}$ be as in Lemmas 2.23–2.25. For $\gamma^* \in [0,\infty)$, put $p^*_{1,1,\gamma^*}(x) := 1$, $p^*_{0,0,\gamma^*}(x) := 0$ $(x \in [0,1])$, and*

$$p^*_{0,1,\gamma^*}(0) := 0 \quad \textit{and} \quad p^*_{0,1,\gamma^*}(x) := h_{0,1}(x)\rho_{\gamma^*}(x) \qquad (x \in (0,1]), \tag{2.2.45}$$

with ρ_{γ^*} *as in (2.2.44). Then, for* $(l,r)=(1,1),(0,0),$ *and* $(0,1)$,

$$P^{-n,\delta_x}\big[\langle\mathcal{X}_0,h_{l,r}\rangle\in\cdot\,\big]\underset{n\to\infty}{\Longrightarrow} e^{-p^*_{l,r,\gamma^*}(x)}\delta_0+\big(1-e^{-p^*_{l,r,\gamma^*}(x)}\big)\delta_\infty, \tag{2.2.46}$$

uniformly for all $x\in[0,1]$.

Formula (2.2.46) says that the weighted branching process $\mathcal{X}^{h_{l,r}}$ exhibits a form of extinction versus unbounded growth. More precisely, for large n the total mass of $h_{l,r}\mathcal{X}_0$ is close to 0 or ∞ with high probability.

Proof of Proposition 2.22 By (2.2.15),

$$\mathcal{U}^{(n)}p(x)=-\log E^{-n,\delta_x}\big[e^{-\langle\mathcal{X}_0,p\rangle}\big]\qquad(p\in B_+[0,1],\ x\in[0,1]). \tag{2.2.47}$$

We first prove formula (2.2.39) (ii). For $(l,r)=(0,0)$, formula (2.2.46) says that

$$P^{-n,\delta_x}[\langle\mathcal{X}_0,h_{0,0}\rangle\in\cdot\,]\underset{n\to\infty}{\Longrightarrow}\delta_0 \tag{2.2.48}$$

uniformly for all $x\in[0,1]$. If $p\in\mathcal{H}_{0,0}$, then we can find $r>0$ such that $p\leq rh_{0,0}$. Therefore, (2.2.48) implies that for any $p\in\mathcal{H}_{0,0}$,

$$P^{-n,\delta_x}[\langle\mathcal{X}_0,p\rangle\in\cdot\,]\underset{n\to\infty}{\Longrightarrow}\delta_0. \tag{2.2.49}$$

By (2.2.47) it follows that

$$\mathcal{U}^{(n)}p(x)=-\log E^{-n,\delta_x}\big[e^{-\langle\mathcal{X}_0,p\rangle}\big]\underset{n\to\infty}{\longrightarrow}0, \tag{2.2.50}$$

where the limits in (2.2.49) and (2.2.50) are uniform in $x\in[0,1]$. This proves formula (2.2.39) (ii). To prove formula (2.2.39) (iii), note that for any $p\in\mathcal{H}_{0,1}$ we can choose $0<r_-<r_+$ such that $r_-h_{0,1}\leq p+h_{0,0}\leq r_+h_{0,1}$. Therefore, (2.2.46) implies that

$$P^{-n,\delta_x}[\langle\mathcal{X}_0,p\rangle+\langle\mathcal{X}_0,h_{0,0}\rangle\in\cdot\,]\underset{n\to\infty}{\Longrightarrow} e^{-p^*_{0,1,\gamma^*}(x)}\delta_0+\big(1-e^{-p^*_{0,1,\gamma^*}(x)}\big)\delta_\infty. \tag{2.2.51}$$

Using moreover (2.2.48), we see that

$$P^{-n,\delta_x}[\langle\mathcal{X}_0,p\rangle\in\cdot\,]\underset{n\to\infty}{\Longrightarrow} e^{-p^*_{0,1,\gamma^*}(x)}\delta_0+\big(1-e^{-p^*_{0,1,\gamma^*}(x)}\big)\delta_\infty. \tag{2.2.52}$$

By (2.2.47), it follows that

$$\mathcal{U}^{(n)}p(x)=-\log E^{-n,\delta_x}\big[e^{-\langle\mathcal{X}_0,p\rangle}\big]\underset{n\to\infty}{\longrightarrow}p^*_{0,1,\gamma^*}(x) \tag{2.2.53}$$

where all limits are uniform in $x\in[0,1]$. This proves (2.2.39) (iii). The proof of (2.2.39) (i) is similar but easier. ∎

2.2.8 Outline

In Section 2.3, we verify that $\mathcal{W}_{\mathrm{cat}}$ is a renormalization class, we prove Proposition 2.18, which connects the renormalization transformations F_c to the log-Laplace operators $\mathcal{U}_\gamma$, and we collect a number of technical properties of the operators $\mathcal{U}_\gamma$ that will be needed later on. In Section 2.4 we prove Theorem 2.19 about the convergence of the renormalization branching process to a time-homogeneous limit.

Sections 2.5–2.7 are devoted to the super-Wright-Fisher diffusio $\mathcal{Y}^0$, i.e., the limiting process from Theorem 2.19 (b). These sections have been written in such a way that they can be read independently of the rest of this chapter. In fact, we generalize a bit by allowing for an arbitrary positive constant to appear in front of the $u(1-u)$ term in (2.2.26). This generatization reveals that the case where this constant is one is in fact a critical case, marking the boundary between two types of long-time behavior. Section 2.5 gives an introduction to the super-Wright-Fisher diffusion, while Sections 2.6–2.7 contain proofs. The central tool in these proofs is a weighted superprocess, rather than embedded particle systems which are our main tool for studying the renormalization branching process $\mathcal{X}$

In Section 2.8, we take up the study of $\mathcal{X}$ and its embedded particle systems. In particular, we prove the statements from Section 2.2.7 about extinction versus unbounded growth of embedded particle systems, with the exception of Lemma 2.24, which is proved in Section 2.9. In Section 2.10, finally, we combine all results derived by that point to prove our main theorem.

Acknowledgements Work sponsored by the DFG. The authors thank Janos Engländer for answering our questions about his work and Jan Seidler for answering questions about the strong Feller property. Achim Klenke, Dmitry Turaev, and Anita Winter are thanked for useful discussions and comments. We than an anonymous referee for comments which lead to an improved exposition. We thank Anton Wakolbinger and Martin Möhle for pointing out reference [Ewe04] and the fact that the distribution in (2.3.17) is a β-distribution.

2.3 The renormalization class $\mathcal{W}_{\mathrm{cat}}$

In this section we prove Theorem 2.17 (a) and Proposition 2.18, as well as Lemmas 2.3–2.8 from Section 2.1.4, and Lemma 2.15. The section is organized according to the techniques used. Section 2.3.1 collects some facts that hold for general renormalization classes on compact sets. In Section 2.3.2 we use the SDE (2.2.4) to couple catalytic Wright-Fisher diffusions. In Section 2.3.3 we apply the moment duality for the Wright-Fisher diffusion to the catalyst and to the reactant conditioned on the catalyst. In Section 2.3.4 we prove that monotone concave catalyzing functions form a preserved class under renormalization.

2.3.1 Renormalization classes on compact sets

In this section, we prove the lemmas stated in Section 2.1.4, as well as Lemma 2.15. Recall that $D \subset \mathbb{R}^d$ is open, bounded, and convex, and that $\mathcal{W}$ is a prerenormalization class on $\overline{D}$, equipped with the topology of uniform convergence.

Proof of Lemma 2.3 To see that $(x, c, w) \mapsto \nu_x^{c,w}$ is continuous, let (x_n, c_n, w_n) be a sequence converging in $\overline{D} \times (0, \infty) \times \mathcal{W}$ to a limit (x, c, w). By the compactness of $\overline{D}$, the sequence

$(\nu_{x_n}^{c_n,w_n})_{n\geq 0}$ is tight, and each limit point ν^* satisfies

$$\langle \nu^*, A_x^{c,w} f\rangle = 0 \qquad (f \in \mathcal{C}^{(2)}(D)). \tag{2.3.1}$$

Therefore, by [EK86, Theorem 4.9.17], ν^* is an invariant law for the martingale problem associated with $A_x^{c,w}$. Since we are assuming uniqueness of the invariant law, $\nu^* = \nu_x^{c,w}$ and therefore $\nu_{x_n}^{c_n,w_n} \Rightarrow \nu_x^{c,w}$. The continuity of $F_c w(x)$ is a simple consequence of the continuity of $\nu_x^{c,w}$. ∎

Proof of Lemma 2.4 Formula (2.1.31) (i) follows from the fact that rescaling the time in solutions $(\mathbf{y}_t)_{t\geq 0}$ to the martingale problem for $A_x^{c,w}$ by a factor λ has no influence on the invariant law. Formula (2.1.31) (ii) is a direct consequence of formula (2.1.31) (i). ∎

Proof of Lemma 2.5 This follows by inserting the functions $f(x) = x_i$ and $f(x) = x_i x_j$ into the equilibrium equation (2.3.1). ∎

Proof of Lemma 2.6 If $x \in \partial_w D$, then $\mathbf{y}_t := x$ $(t \geq 0)$ is a stationary solution to the martingale problem for $A_x^{c,w}$, and therefore $\nu_x^{c,w} = \delta_x$ and $F_c w(x) = w(x) = 0$. On the other hand, if $x \notin \partial_w D$, then $\mathbf{y}_t := x$ $(t \geq 0)$ is not a stationary solution to the martingale problem for $A_x^{c,w}$ and therefore $\int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)|y-x|^2 > 0$. Let $\mathrm{tr}(w(y)) := \sum_i w_{ii}(y)$ denote the trace of $w(y)$. By (2.1.32) (ii), $\frac{1}{c}\mathrm{tr}(F_c w)(x) = \frac{1}{c}\int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)\mathrm{tr}(w(y)) = \int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)|y-x|^2 > 0$ and therefore $F_c w(x) \neq 0$. ∎

From now on assume that $\mathcal{W}$ is a renormalization class. Note that

$$K^{w,(n)} = \nu^{c_{n-1},F^{(n-1)}w} \cdots \nu^{c_0,w} \qquad (n \geq 1), \tag{2.3.2}$$

where we denote the composition of two probability kernels K, L on $\overline{D}$ by

$$(KL)_x(\mathrm{d}z) := \int_{\overline{D}} K_x(\mathrm{d}y) L_y(\mathrm{d}z). \tag{2.3.3}$$

Proof of Lemma 2.7 This is a direct consequence of Lemmas 2.3 and 2.5. In particular, the relations (2.1.36) follow by iterating the relations (2.1.32). ∎

Proof of Lemma 2.8 Recall that $\mathrm{tr}(w(y))$ denotes the trace of $w(y)$. Formulas (2.1.35) and (2.1.36) (ii) show that

$$\int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y)\, |y-x|^2 = s_n \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y)\, \mathrm{tr}(w(y)). \tag{2.3.4}$$

Since $\overline{D}$ is compact, the left-hand side of this equation is bounded uniformly in $x \in \overline{D}$ and $n \geq 1$, and therefore, since we are assuming $s_n \to \infty$,

$$\lim_{n\to\infty} \sup_{x \in D} \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y)\mathrm{tr}(w(y)) = 0. \tag{2.3.5}$$

Since w is symmetric and nonnegative definite, $\mathrm{tr}(w(y))$ is nonnegative, and zero if and only if $y \in \partial_w D$. If $f \in \mathcal{C}(\overline{D})$ satisfies $f = 0$ on $\partial_w D$, then, for every $\varepsilon > 0$, the sets $C_m := \{x \in \overline{D}:$

$|f(x)| \geq \varepsilon + m\,\mathrm{tr}(w(x))\}$ are compact with $C_m \downarrow \emptyset$ as $m \uparrow \infty$, so there exists an m (depending on ε) such that $|f| < \varepsilon + m\,\mathrm{tr}(w)$. Therefore,

$$\begin{aligned} \limsup_{n\to\infty} \sup_{x\in\overline{D}} \Big| \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y) f(y) \Big| &\leq \limsup_{n\to\infty} \sup_{x\in\overline{D}} \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y) |f(y)| \\ \leq \varepsilon + m \limsup_{n\to\infty} \sup_{x\in\overline{D}} \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y) \mathrm{tr}(w(y)) = \varepsilon. \end{aligned} \tag{2.3.6}$$

Since $\varepsilon > 0$ is arbitrary, (2.1.37) follows. ∎

Proof of Lemma 2.15 By (2.1.40), (2.1.42), and (2.1.43), $w^*_{\gamma^*} = \lim_{n\to\infty}(\overline{F}_{\gamma^*})^n w$ for each $w \in \mathcal{W}$. By Lemma 2.3 (b), $\overline{F}_{\gamma^*} : \mathcal{W} \to \mathcal{W}$ is continuous, so $w^*_{\gamma^*}$ is the unique fixed point of $\overline{F}_{\gamma^*}$. This proves part (a).

Now let $0 \neq w \in \mathcal{W}$ and assume that $\hat{\mathcal{W}} = \{\lambda w : \lambda > 0\}$ is a fixed shape. Then $\hat{\mathcal{W}} \ni s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*_{\gamma^*}$ whenever $s_n \to \infty$ and $s_{n+1}/s_n \to 1+\gamma^*$ for some $0 < \gamma^* < \infty$, which shows that $\hat{\mathcal{W}} = \{\lambda w^*_{\gamma^*} : \lambda > 0\}$. Thus, $\mathcal{W}$ can contain at most one fixed shape, and if it does, then the $w^*_{\gamma^*}$ for different values of γ^* must be constant multiples of each other. This proves part (c) and the uniqueness statement in part (b).

To complete the proof of part (b), note that if $w^* = w^*_{\gamma^*}$ does not depend on γ^*, then $w^* \in \mathcal{W}$ solves (2.1.46) (i) for all $0 < \gamma^* < \infty$, hence $F_c w^* = (1 + \frac{1}{c})^{-1} w^*$ for all $c > 0$, and therefore, by scaling (Lemma 2.4), $F_c(\lambda w^*) = \lambda F_{c/\lambda}(w^*) = \lambda(1 + \frac{\lambda}{c})^{-1} w^* = (\frac{1}{\lambda} + \frac{1}{c})^{-1} w^*$. ∎

2.3.2 Coupling of catalytic Wright-Fisher diffusions

In this section we verify condition (i) of Definition 2.1 for the class $\mathcal{W}_{\mathrm{cat}}$, and we prepare for the verification of conditions (ii)–(iv) in Section 2.3.3. In fact, we will show that the larger class $\overline{\mathcal{W}}_{\mathrm{cat}} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{C}_+[0,1]\}$ is also a renormalization class, and the equivalents of Theorem 2.17 (a) and Proposition 2.18 remain true for this larger class. (We do not know, however, if the convergence statements in Theorem 2.17 (b) also hold in this larger class; see the discussion in Section 2.2.2.)

For each $c \geq 0$, $w \in \overline{\mathcal{W}}_{\mathrm{cat}}$ and $x \in [0,1]^2$, the operator $A_x^{c,w}$ is a densely defined linear operator on $\mathcal{C}([0,1]^2)$ that maps the identity function into zero and, as one easily verifies, satisfies the positive maximum principle. Since $[0,1]^2$ is compact, the existence of a solution to the martingale problem for $A_x^{c,w}$, for each $[0,1]^2$-valued initial condition, now follows from general theory (see [RW87], Theorem 5.23.5, or [EK86, Theorem 4.5.4 and Remark 4.5.5]).

We are therefore left with the task of verifying uniqueness of solutions to the martingale problem for $A_x^{c,w}$. By [EK86, Problem 4.19, Corollary 5.3.4, and Theorem 5.3.6], it suffices to show that solutions to (2.2.4) are pathwise unique.

Lemma 2.27 (Monotone coupling of Wright-Fisher diffusions) *Assume that* $0 \leq x \leq \tilde{x} \leq 1$, $c \geq 0$ *and that* $(P_t)_{t\geq 0}$ *is a progressively measurable, nonnegative process such that* $\sup_{t\geq 0,\omega\in\Omega} P_t(\omega) < \infty$. *Let* $\mathbf{y}, \tilde{\mathbf{y}}$ *be* $[0,1]$*-valued solutions to the SDE's*

$$\begin{aligned} \mathrm{d}\mathbf{y}_t &= c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2P_t \mathbf{y}_t(1-\mathbf{y}_t)}\mathrm{d}B_t, \\ \mathrm{d}\tilde{\mathbf{y}}_t &= c\,(\tilde{x} - \tilde{\mathbf{y}}_t)\mathrm{d}t + \sqrt{2P_t \tilde{\mathbf{y}}_t(1-\tilde{\mathbf{y}}_t)}\mathrm{d}B_t, \end{aligned} \tag{2.3.7}$$

where in both equations B is the same Brownian motion. If $\mathbf{y}_0 \leq \tilde{\mathbf{y}}_0$ a.s., then

$$\mathbf{y}_t \leq \tilde{\mathbf{y}}_t \quad \forall t \geq 0 \quad a.s. \tag{2.3.8}$$

Proof This is an easy adaptation of a technique due to Yamada and Watanabe [YW71]. Since $\int_{0+} \frac{\mathrm{d}x}{x} = \infty$, it is possible to choose $\rho_n \in \mathcal{C}[0,\infty)$ such that $\int_0^\infty \rho_n(x)\mathrm{d}x = 1$ and

$$0 \leq \rho_n(x) \leq \frac{1}{nx} 1_{(0,1]}(x) \qquad (x \geq 0). \tag{2.3.9}$$

Define $\phi_n \in \mathcal{C}^{(2)}(\mathbb{R})$ by

$$\phi_n(x) := \int_0^{x \vee 0} \mathrm{d}y \int_0^y \mathrm{d}z\, \rho_n(z). \tag{2.3.10}$$

One easily verifies that $\phi_n(x)$, $x\phi_n'(x)$, and $x\phi_n''(x)$ are nonnegative and converge, as $n \to \infty$, to $x \vee 0$, $x \vee 0$, and 0, respectively. By Itô's formula:

$$\begin{aligned}
E[\phi_n(\mathbf{y}_t - \tilde{\mathbf{y}}_t)] = {} & E[\phi_n(\mathbf{y}_0 - \tilde{\mathbf{y}}_0)] && \text{(i)}\\
& + c\,(x - \tilde{x}) \int_0^t E[\phi_n'(\mathbf{y}_s - \tilde{\mathbf{y}}_s)]\mathrm{d}s - c \int_0^t E[(\mathbf{y}_s - \tilde{\mathbf{y}}_s)\phi_n'(\mathbf{y}_s - \tilde{\mathbf{y}}_s)]\mathrm{d}s && \text{(ii)}\\
& + \int_0^t E\Big[P_s\Big(\sqrt{\mathbf{y}_s(1-\mathbf{y}_s)} - \sqrt{\tilde{\mathbf{y}}_s(1-\tilde{\mathbf{y}}_s)}\Big)^2 \phi_n''(\mathbf{y}_s - \tilde{\mathbf{y}}_s)\Big]\mathrm{d}s. && \text{(iii)}
\end{aligned} \tag{2.3.11}$$

Here the terms in (ii) are nonpositive, and hence, letting $n \to \infty$ and using the elementary estimate

$$|\sqrt{y(1-y)} - \sqrt{\tilde{y}(1-\tilde{y})}| \leq |y - \tilde{y}|^{\frac{1}{2}} \qquad (y, \tilde{y} \in [0,1]), \tag{2.3.12}$$

the properties of ϕ_n, and the fact that the process P is uniformly bounded, we find that

$$E[0 \vee (\mathbf{y}_t - \tilde{\mathbf{y}}_t)] \leq E[0 \vee (\mathbf{y}_0 - \tilde{\mathbf{y}}_0)] = 0, \tag{2.3.13}$$

by our assumption that $\mathbf{y}_0 \leq \tilde{\mathbf{y}}_0$. This shows that $\mathbf{y}_t \leq \tilde{\mathbf{y}}_t$ a.s. for each fixed $t \geq 0$, and by the continuity of sample paths the statement holds for all $t \geq 0$ almost surely. ∎

Corollary 2.28 (Pathwise uniqueness) *For all $c \geq 0$, $\alpha > 0$, $p \in \mathcal{C}_+[0,1]$ and $x \in [0,1]$, solutions to the SDE (2.2.4) are pathwise unique.*

Proof Let $(\mathbf{y}^1, \mathbf{y}^2)$ and $(\tilde{\mathbf{y}}^1, \tilde{\mathbf{y}}^2)$ be solutions to (2.2.4) relative to the same pair (B^1, B^2) of Brownian motions, with $(\mathbf{y}_0^1, \mathbf{y}_0^2) = (\tilde{\mathbf{y}}_0^1, \tilde{\mathbf{y}}_0^2)$. Applying Lemma 2.27, with inequality in both directions, we see that $\mathbf{y}^1 = \tilde{\mathbf{y}}^1$ a.s. Applying Lemma 2.27 two more times, this time using that $\mathbf{y}^1 = \tilde{\mathbf{y}}^1$ a.s., we see that also $\mathbf{y}^2 = \tilde{\mathbf{y}}^2$ a.s. ∎

Corollary 2.29 (Exponential coupling) *Assume that $x \in [0,1]$, $c \geq 0$, and $\alpha > 0$. Let $\mathbf{y}, \tilde{\mathbf{y}}$ be solutions to the SDE*

$$\mathrm{d}\mathbf{y}_t = c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2\alpha\mathbf{y}_t(1-\mathbf{y}_t)}\mathrm{d}B_t, \tag{2.3.14}$$

relative to the same Brownian motion B. Then

$$E\big[|\tilde{\mathbf{y}}_t - \mathbf{y}_t|\big] = e^{-ct} E\big[|\tilde{\mathbf{y}}_0 - \mathbf{y}_0|\big]. \tag{2.3.15}$$

Proof If $\mathbf{y}_0 = y$ and $\tilde{\mathbf{y}}_0 = \tilde{y}$ are deterministic and $y \leq \tilde{y}$, then by Lemma 2.27 and a simple moment calculation

$$E\big[|\tilde{\mathbf{y}}_t - \mathbf{y}_t|\big] = E[\tilde{\mathbf{y}}_t - \mathbf{y}_t] = e^{-ct}|\tilde{y} - y|. \tag{2.3.16}$$

The same argument applies when $y \geq \tilde{y}$. The general case where $\mathbf{y}_0$ and $\tilde{\mathbf{y}}_0$ are random follows by conditioning on $(\mathbf{y}_0, \tilde{\mathbf{y}}_0)$. ∎

Corollary 2.30 (Ergodicity) *The Markov process defined by the SDE (2.2.17) has a unique invariant law Γ^γ_x and is ergodic, i.e, solutions to (2.2.17) started in an arbitrary initial law $\mathcal{L}(\mathbf{y}_0)$ satisfy $\mathcal{L}(\mathbf{y}_t) \underset{t\to\infty}{\Longrightarrow} \Gamma^\gamma_x$.*

Proof Since our process is a Feller diffusion on a compactum, the existence of an invariant law follows from a simple time averaging argument. Now start one solution $\tilde{\mathbf{y}}$ of (2.2.17) in this invariant law and let $\mathbf{y}$ be any other solution, relative to the same Brownian motion. Corollary 2.29 then gives ergodicity and, in particular, uniqueness of the invariant law. ∎

Remark 2.31 (Density of invariant law) It is well-known (see, for example [Ewe04, formula (5.70)]) that Γ^γ_x is a $\beta(\alpha_1, \alpha_2)$-distribution, where $\alpha_1 := x/\gamma$ and $\alpha_2 := (1-x)/\gamma$, i.e., $\Gamma^\gamma_x = \delta_x$ $(x \in \{0,1\})$ and

$$\Gamma^\gamma_x(\mathrm{d}y) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\, y^{\alpha_1-1}(1-y)^{\alpha_2-1}\mathrm{d}y \qquad (x \in (0,1)). \tag{2.3.17}$$

◊

We conclude this section with a lemma that prepares for the verification of condition (iv) in Definition 2.1 for the class $\mathcal{W}_{\rm cat}$.

Lemma 2.32 (Monotone coupling of stationary Wright-Fisher diffusions) *Assume that $c > 0$, $\alpha > 0$ and $0 \leq x \leq \tilde{x} \leq 1$. Then the pair of equations*

$$\begin{aligned} \mathrm{d}\mathbf{y}_t &= c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2\alpha\mathbf{y}_t(1-\mathbf{y}_t)}\mathrm{d}B_t, \\ \mathrm{d}\tilde{\mathbf{y}}_t &= c\,(\tilde{x} - \tilde{\mathbf{y}}_t)\mathrm{d}t + \sqrt{2\alpha\tilde{\mathbf{y}}_t(1-\tilde{\mathbf{y}}_t)}\mathrm{d}B_t \end{aligned} \tag{2.3.18}$$

has a unique stationary solution $(\mathbf{y}_t, \tilde{\mathbf{y}}_t)_{t\in\mathbb{R}}$. This stationary solution satisfies

$$\mathbf{y}_t \leq \tilde{\mathbf{y}}_t \quad \forall t \in \mathbb{R} \quad a.s. \tag{2.3.19}$$

Proof Let $(\mathbf{y}_t, \tilde{\mathbf{y}}_t)_{t\geq 0}$ be a solution of (2.3.18) and let $(\mathbf{y}'_t, \tilde{\mathbf{y}}'_t)_{t\geq 0}$ be another one, relative to the same Brownian motion B. Then, by Lemma 2.29, $E[|\mathbf{y}_t - \mathbf{y}'_t|] \to 0$ and also $E[|\tilde{\mathbf{y}}_t - \tilde{\mathbf{y}}'_t|] \to 0$ as $t \to \infty$. Hence we may argue as in the proof of Corollary 2.30 that (2.3.18) has a unique invariant law and is ergodic. Now start a solution of (2.3.18) in an initial condition such that $\mathbf{y}_0 \leq \tilde{\mathbf{y}}_0$. By ergodicity, the law of this solution converges as $t \to \infty$ to the invariant law of (2.3.18) and using Lemma 2.27 we see that this invariant law is concentrated on $\{(y,\tilde{y}) \in [0,1]^2 : y \leq \tilde{y}\}$. Now consider, on the whole real time axis, the stationary solution to (2.3.18) with this invariant law. Applying Lemma 2.27 once more, we see that (2.3.19) holds. ∎

2.3.3 Duality for catalytic Wright-Fisher diffusions

In this section we prove Theorem 2.17 (a) and Proposition 2.18. Moreover, we will show that their statements remain true if the renormalization class $\mathcal{W}_{\rm cat}$ is replaced by the larger class $\overline{\mathcal{W}}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{C}_+[0,1]\}$. We begin by recalling the usual moment duality for Wright-Fisher diffusions.

For $\gamma > 0$ and $x \in [0,1]$, let $\mathbf{y}$ be a solution to the SDE

$$\mathrm{d}\mathbf{y}(t) = \tfrac{1}{\gamma}\,(x - \mathbf{y}(t))\mathrm{d}t + \sqrt{2\mathbf{y}(t)(1-\mathbf{y}(t))}\mathrm{d}B(t), \tag{2.3.20}$$

i.e., $\mathbf{y}$ is a Wright-Fisher diffusion with a linear drift towards x. It is well-known that $\mathbf{y}$ has a moment dual. To be precise, let (ϕ,ψ) be a Markov process in $\mathbb{N}^2 = \{0,1,\ldots\}^2$ that jumps as:

$$\begin{array}{lll} (\phi_t,\psi_t) \to (\phi_t - 1, \psi_t) & \text{with rate} & \phi_t(\phi_t - 1) \\ (\phi_t,\psi_t) \to (\phi_t - 1, \psi_t + 1) & \text{with rate} & \frac{1}{\gamma}\phi_t. \end{array} \tag{2.3.21}$$

Then one has the following *duality relation* (see for example Lemma 2.3 in [Shi80a] or Proposition 1.5 in [GKW01])

$$E^y\big[\mathbf{y}_t^n x^m\big] = E^{(n,m)}\big[y^{\phi_t} x^{\psi_t}\big] \qquad (y \in [0,1],\ (n,m) \in \mathbb{N}^2), \tag{2.3.22}$$

where $0^0 := 1$. The duality in (2.3.22) has the following heuristic explanation. Consider a population containing a fixed, large number of organisms, that come in two genetic types, say I and II. Each pair of organisms in the population is *resampled* with rate 2. This means that one organism of the pair (chosen at random) dies, while the other organism produces one child of its own genetic type. Moreover, each organism is replaced with rate $\frac{1}{\gamma}$ by an organism chosen from an infinite reservoir where the frequency of type I has the fixed value x. In the limit that the number of organisms in the population is large, the relative frequency $\mathbf{y}_t$ of type I organisms follows the SDE (2.3.20). Now $E[\mathbf{y}_t^n]$ is the probability that n organisms sampled from the population at time t are all of type I. In order to find this probability, we follow the ancestors of these organisms back in time. Viewed backwards in time, these ancestors live for a while in the population, until, with rate $\frac{1}{\gamma}$, they jump to the infinite reservoir. Moreover, due to resampling, each pair of ancestors coalesces with rate 2 to one common ancestor. Denoting the number of ancestors that lived at time $t-s$ in the population and in the reservoir by ϕ_s and ψ_s, respectively, we see that the probability that all ancestors are of type I is $E^y[\mathbf{y}_t^n] = E^{(n,0)}[y^{\phi_t} x^{\psi_t}]$. This gives a heuristic explanation of (2.3.22).

Since eventually all ancestors of the process (ϕ,ψ) end up in the reservoir, we have $(\phi_t,\psi_t) \to (0,\psi_\infty)$ as $t \to \infty$ a.s. for some $\mathbb{N}$-valued random variable ψ_∞. Taking the limit $t \to \infty$ in (2.3.22), we see that the moments of the invariant law Γ_x^γ from Corollary 2.30 are given by:

$$\int \Gamma_x^\gamma(\mathrm{d}y) y^n = E^{(n,0)}[x^{\psi_\infty}] \qquad (n \geq 0). \tag{2.3.23}$$

It is not hard to obtain an inductive formula for the moments of Γ_x^γ, which can then be solved to yield the formula

$$\int \Gamma_x^\gamma(\mathrm{d}y) y^n = \prod_{k=0}^{n-1} \frac{x + k\gamma}{1 + k\gamma} \qquad (n \geq 1). \tag{2.3.24}$$

In particular, it follows that

$$\int \Gamma^\gamma_x(\mathrm{d}y)y(1-y) = \frac{1}{1+\gamma}x(1-x). \tag{2.3.25}$$

This is the important *fixed shape property* of the Wright-Fisher diffusion (see formula (2.1.58)).

We now consider catalytic Wright-Fisher diffusions $(\mathbf{y}^1,\mathbf{y}^2)$ as in (2.2.4) with $p\in\mathcal{C}_+[0,1]$ and apply duality to the catalyst $\mathbf{y}^2$ conditioned on the reactant $\mathbf{y}^1$. Let $(\mathbf{y}^1_t,\mathbf{y}^2_t)_{t\in\mathbb{R}}$ be a stationary solution to the SDE (2.2.4) with $c=1/\gamma$. Let $(\tilde\phi,\tilde\psi)$ be a $\mathbb{N}^2$-valued process, defined on the same probability space as $(\mathbf{y}^1,\mathbf{y}^2)$, such that conditioned on the past path $(\mathbf{y}^1_{-t})_{t\leq 0}$, the process $(\tilde\phi,\tilde\psi)$ is a (time-inhomogeneous) Markov process that jumps as:

$$\begin{array}{lll} (\tilde\phi_t,\tilde\psi_t)\to(\tilde\phi_t-1,\tilde\psi_t) & \text{with rate} & p(\mathbf{y}^1_{-t})\tilde\phi_t(\tilde\phi_t-1), \\ (\tilde\phi_t,\tilde\psi_t)\to(\tilde\phi_t-1,\tilde\psi_t+1) & \text{with rate} & \frac{1}{\gamma}\tilde\phi_t. \end{array} \tag{2.3.26}$$

Then, in analogy with (2.3.22),

$$E[(\mathbf{y}^2_0)^n x_2^m|(\mathbf{y}^1_{-t})_{t\leq 0}] = E^{(n,m)}[(\mathbf{y}^2_{-t})^{\tilde\phi_t}x_2^{\tilde\psi_t}|(\mathbf{y}^1_{-t})_{t\leq 0}] \qquad ((n,m)\in\mathbb{N}^2,\ t\geq 0). \tag{2.3.27}$$

We may interpret (2.3.26) by saying that pairs of ancestors in a finite population coalesce with time-dependent rate $2p(\mathbf{y}^1_{-t})$ and ancestors jump to an infinite reservoir with constant rate $\frac{1}{\gamma}$. Again, eventualy all ancestors end up in the reservoir, and therefore $(\tilde\phi_t,\tilde\psi_t)\to(0,\tilde\psi_\infty)$ as $t\to\infty$ a.s. for some $\mathbb{N}$-valued random variable $\tilde\psi_\infty$. Taking the limit $t\to\infty$ in (2.3.27) we find that

$$E[(\mathbf{y}^2_0)^n x_2^m|(\mathbf{y}^1_{-t})_{t\leq 0}] = E^{(n,m)}[x_2^{\tilde\psi_\infty}|(\mathbf{y}^1_{-t})_{t\leq 0}] \qquad ((n,m)\in\mathbb{N}^2,\ t\geq 0). \tag{2.3.28}$$

Lemma 2.33 (Uniqueness of invariant law) *For each $c>0$, $w\in\overline{\mathcal{W}}_{\rm cat}$, and $x\in[0,1]^2$, there exists a unique invariant law $\nu^{c,w}_x$ for the martingale problem for $A^{c,w}_x$.*

Proof Our process being a Feller diffusion on a compactum, the existence of an invariant law follows from time averaging. We need to show uniqueness. If $(\mathbf{y}^1,\mathbf{y}^2) = \mathbf{y}^1_t,\mathbf{y}^2_t)_{t\in\mathbb{R}}$ is a stationary solution, then $\mathbf{y}^1$ is an autonomous process, and $\mathcal{L}(\mathbf{y}^1_0)=\Gamma^{1/c}_x$, the unique invariant law from Corollary 2.30. Therefore, $\mathcal{L}((\mathbf{y}^1_t)_{t\in\mathbb{R}})$ is determined uniquely by the requirement that $(\mathbf{y}^1,\mathbf{y}^2)$ be stationary. By (2.3.28), the conditional distribution of $\mathbf{y}^2_0$ given $(\mathbf{y}^1_t)_{t\leq 0}$ is determined uniquely, and therefore the joint distribution of $\mathbf{y}^2_0$ and $(\mathbf{y}^1_t)_{t\leq 0}$ is determined uniquely. In particular, $\mathcal{L}(\mathbf{y}^1_0,\mathbf{y}^2_0)=\nu^{c,w}_x$ is determined uniquely. ∎

Remark 2.34 (Reversibility) It seems that the invariant law $\nu^{c,w}_x$ from Lemma 2.33 is reversible. In many cases (densities of) reversible invariant measures can be obtained in closed form by solving the equations of detailed balance. This is the case, for example, for the one-dimensional Wright-Fisher diffusion. We have not attempted this for the catalytic Wright-Fisher diffusion. ◊

The next proposition implies Proposition 2.18 and prepares for the proof of Theorem 2.17 (a).

Proposition 2.35 (Extended renormalization class) *The set $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class on $[0,1]^2$, and*

$$\overline{F}_\gamma w^{1,p} = w^{1,\mathcal{U}_\gamma p} \qquad (p \in \mathcal{C}_+[0,1],\ \gamma > 0). \tag{2.3.29}$$

Proof To see that $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class we need to check conditions (i)–(iv) from Definition 2.1. By Lemma 2.28, the martingale problem for $A^{c,w}_x$ is well-posed for all $c \geq 0$, $w \in \mathcal{W}_{\rm cat}$ and $x \in [0,1]^2$. By Lemma 2.33, the corresponding Feller process on $[0,1]^2$ has a unique invariant law $\nu^{c,w}_x$. This shows that conditions (i) and (ii) from Definition 2.1 are satisfied. Note that by the compactness of $[0,1]^2$, any continuous function on $[0,1]^2$ is bounded, so condition (iii) is automatically satisfied. Hence $\mathcal{W}$ is a prerenormalization class. As a consequence, for any $p \in \mathcal{C}_+[0,1]$, $\overline{F}_\gamma w^{1,p}$ is well-defined by (2.1.25) and (2.1.39). We will now first prove (2.3.29) and then show that $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class.

Fix $\gamma > 0$, $p \in \mathcal{C}_+[0,1]$, and $x \in [0,1]^2$. Let $(\mathbf{y}^1_t, \mathbf{y}^2_t)_{t\in\mathbb{R}}$ be a stationary solution to the SDE (2.2.4) with $\alpha = 1$ and $c = 1/\gamma$. Then

$$\overline{F}_\gamma w^{1,p}_{ij}(x) = (1+\gamma) E[w^{1,p}_{ij}(\mathbf{y}^1_0, \mathbf{y}^2_0)] \qquad (i,j = 1,2). \tag{2.3.30}$$

Since $w^{1,p}_{ij} = 0$ if $i \neq j$, it is clear that $\overline{F}_\gamma w^{1,p}_{ij}(x) = 0$ if $i \neq j$. Since $\mathcal{L}(\mathbf{y}^1_0) = \Gamma^\gamma_x$ it follows from (2.3.25) that $\overline{F}_\gamma w^{1,p}_{11}(x) = x_1(1-x_1)$. We are left with the task of showing that

$$\overline{F}_\gamma w^{1,p}_{22}(x) = \mathcal{U}_\gamma p(x_1) x_2(1-x_2). \tag{2.3.31}$$

Here, by (2.1.32) (ii),

$$\begin{aligned}\overline{F}_\gamma w^{1,p}_{22}(x) &= (1+\gamma) E[p(\mathbf{y}^1_0)\mathbf{y}^2_0(1-\mathbf{y}^2_0)] \\ &= (\tfrac{1}{\gamma}+1) E[(\mathbf{y}^2_0 - x_2)^2].\end{aligned} \tag{2.3.32}$$

By (2.3.28), using the fact that $E[\mathbf{y}^2_0] = x_2$ (which follows from (2.3.27) or more elementary from (2.1.36) (i)), we find that

$$E[(\mathbf{y}^2_0 - x_2)^2] = E[(\mathbf{y}^2_0)^2] - (x_2)^2 = E^{(2,0)}[x_2^{\tilde\psi_\infty}] - (x_2)^2 = P^{(2,0)}[\tilde\psi_\infty = 1] x_2(1-x_2) \qquad (t \geq 0). \tag{2.3.33}$$

Note that $P^{(2,0)}[\tilde\psi_\infty = 1]$ is the probability that the two ancestors coalesce before one of them leaves the population. The probability of *noncoalescence* is given by

$$P^{(2,0)}[\tilde\psi_\infty = 2] = E\big[e^{-\int_0^{\frac{1}{2}\tau_\gamma} 2p(y^1_{-t})\mathrm{d}t}\big], \tag{2.3.34}$$

where τ_γ is an exponentially distributed random variable with mean γ. Combining this with (2.3.32) and (2.3.33) we find that

$$\begin{aligned}\overline{F}_\gamma w^{1,p}_{22}(x) &= (\tfrac{1}{\gamma}+1) E\big[1 - e^{-\int_0^{\tau_\gamma} p(y^1_{-t/2})\mathrm{d}t}\big] x_2(1-x_2) \\ &= q_\gamma E\big[1 - e^{-\langle \mathcal{Z}^\gamma_x, p\rangle}\big] x_2(1-x_2) \\ &= \mathcal{U}_\gamma p(x_1) x_2(1-x_2),\end{aligned} \tag{2.3.35}$$

where we have used the definition of $\mathcal{U}_\gamma$.

We still have to show that $\overline{\mathcal{W}}_{\mathrm{cat}}$ satisfies condition (iv) from Definition 2.1. For any $\alpha > 0$ and $p \in \mathcal{C}_+[0,1]$, by scaling (Lemma 2.4) and (2.3.29),

$$F_c w^{\alpha,p} = \alpha F_{\frac{c}{\alpha}} w^{1,\frac{p}{\alpha}} = \alpha(1+\frac{\alpha}{c})^{-1}\overline{F}_{\frac{c}{\alpha}} w^{1,\frac{p}{\alpha}} = w^{(\frac{1}{\alpha}+\frac{1}{c})^{-1},(\frac{1}{\alpha}+\frac{1}{c})^{-1}\mathcal{U}_{\frac{c}{\alpha}}(\frac{p}{\alpha})}. \tag{2.3.36}$$

By Lemma 2.3, this diffusion matrix is continuous, which implies that $\mathcal{U}_{\frac{c}{\alpha}}(\frac{p}{\alpha})$ is continuous. ∎

Our proof of Propostion 2.35 has a corollary.

Corollary 2.36 (Continuity in parameters) *The map $(x,\gamma) \mapsto \mathcal{Q}_\gamma(x,\cdot)$ from $[0,1] \times (0,\infty)$ to $\mathcal{M}_1(\mathcal{M}[0,1])$ and the map $(x,\gamma,p) \mapsto \mathcal{U}_\gamma p(x)$ from $[0,1] \times (0,\infty) \times \mathcal{C}_+[0,1]$ to $\mathbb{R}$ are continuous.*

Proof By Lemma 2.3, the diffusion matrix in (2.3.36) is continuous in x, γ, and p, which implies the continuity of $\mathcal{U}_\gamma p(x)$. It follows that the map $(x,\gamma) \mapsto \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)e^{-\langle\chi,f\rangle}$ is continuous for all $f \in \mathcal{C}_+[0,1]$, so by [Kal76, Theorem 4.2], $(x,\gamma) \mapsto \mathcal{Q}_\gamma(x,\cdot)$ is continuous. ∎

Proof of Theorem 2.17 (a) We need to show that $\mathcal{W}_{\mathrm{cat}}$ is a renormalization class and that F_c maps the subclasses $\mathcal{W}^{l,r}_{\mathrm{cat}}$ into themselves. Since these classes correspond to the different possible effective boundaries of diffusion matrices in $\mathcal{W}_{\mathrm{cat}}$, this latter fact is in fact a consequence of Lemma 2.6. Since in Proposition 2.35 it has been shown that $\overline{\mathcal{W}}_{\mathrm{cat}}$ is a renormalization class, we are left with the task to show that F_c maps $\mathcal{W}_{\mathrm{cat}}$ into itself. By (2.3.29) and scaling, it suffices to show that $\mathcal{U}_\gamma$ maps $\mathcal{H}$ into itself.

Fix $0 \le x \le \tilde{x} \le 1$. By Lemma 2.32, we can couple the processes $\mathbf{y}^\gamma_x$ and $\mathbf{y}^\gamma_{\tilde{x}}$ from (2.2.17) such that

$$\mathbf{y}^\gamma_x(t) \le \mathbf{y}^\gamma_{\tilde{x}}(t) \quad \forall t \le 0 \quad \text{a.s.} \tag{2.3.37}$$

Since the function $z \mapsto 1 - e^{-z}$ on $[0,\infty)$ is Lipschitz continuous with Lipschitz constant 1,

$$\begin{aligned}
&\big|\mathcal{U}_\gamma p(\tilde{x}) - \mathcal{U}_\gamma p(x)\big| \\
&\quad = \Big|(\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} p(\mathbf{y}^\gamma_{\tilde{x}}(-t/2))\mathrm{d}t}\big] - (\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} p(\mathbf{y}^\gamma_x(-t/2))\mathrm{d}t}\big]\Big| \\
&\quad \le (\tfrac{1}{\gamma}+1)E\Big[\int_0^{\tau_\gamma} \big|p(\mathbf{y}^\gamma_{\tilde{x}}(-t/2)) - p(\mathbf{y}^\gamma_x(-t/2))\big|\mathrm{d}t\Big] \\
&\quad \le (\tfrac{1}{\gamma}+1)LE\Big[\int_0^{\tau_\gamma} \big|\mathbf{y}^\gamma_{\tilde{x}}(-t/2) - \mathbf{y}^\gamma_x(-t/2)\big|\mathrm{d}t\Big] \\
&\quad = (\tfrac{1}{\gamma}+1)L\gamma(\tilde{x}-x) = L(1+\gamma)|\tilde{x}-x|,
\end{aligned} \tag{2.3.38}$$

where L is the Lipschitz constant of p and we have used the same exponentially distributed τ_γ for $\mathbf{y}^\gamma_x$ and $\mathbf{y}^\gamma_{\tilde{x}}$. ∎

2.3.4 Monotone and concave catalyzing functions

In this section we prove that the log-Laplace operators $\mathcal{U}_\gamma$ from (2.2.20) map monotone functions into monotone functions, and monotone concave functions into monotone concave functions. We do not know if in general $\mathcal{U}_\gamma$ maps concave functions into concave functions.

Proposition 2.37 (Preservation of monotonicity and concavity) *Let $\gamma > 0$. Then:*

(a) *If $f \in \mathcal{C}_+[0,1]$ is nondecreasing, then $\mathcal{U}_\gamma f$ is nondecreasing.*

(b) *If $f \in \mathcal{C}_+[0,1]$ is nondecreasing and concave, then $\mathcal{U}_\gamma f$ is nondecreasing and concave.*

Proof Our proof of Proposition 2.37 is in part based on ideas from [BCGH97, Appendix A]. The proof is quite long and will depend on several lemmas. We remark that part (a) can be proved in a more elementary way using Lemma 2.32.

We recall some facts from Hille-Yosida theory. A linear operator A on a Banach space V is closable and its closure $\overline{A}$ generates a strongly continuous contraction semigroup $(S_t)_{t\geq 0}$ if and only if

$$\begin{array}{rl} \text{(i)} & \mathcal{D}(A) \text{ is dense,} \\ \text{(ii)} & A \text{ is dissipative,} \\ \text{(iii)} & \mathcal{R}(1-\alpha A) \text{ is dense for some, and hence for all } \alpha > 0. \end{array} \tag{2.3.39}$$

Here, for any linear operator B on V, $\mathcal{D}(B)$ and $\mathcal{R}(B)$ denote the domain and range of B, respectively. For each $\alpha > 0$, the operator $(1-\alpha\overline{A}) : \mathcal{D}(\overline{A}) \to V$ is a bijection and its inverse $(1-\alpha\overline{A})^{-1} : V \to \mathcal{D}(\overline{A})$ is a bounded linear operator, given by

$$(1-\alpha\overline{A})^{-1}u = \int_0^\infty S_t u\, \alpha^{-1} e^{-t/\alpha} \mathrm{d}t \qquad (u \in V,\ \alpha > 0). \tag{2.3.40}$$

If E is a compact metrizable space and $\mathcal{C}(E)$ is the Banach space of continuous real functions on E, equipped with the supremumnorm, then a linear operator A on $\mathcal{C}(E)$ is closable and its closure $\overline{A}$ generates a Feller semigroup if and only if (see [EK86, Theorem 4.2.2 and remarks on page 166])

$$\begin{array}{rl} \text{(i)} & 1 \in \mathcal{D}(\overline{A}) \text{ and } \overline{A}1 = 0, \\ \text{(ii)} & \mathcal{D}(A) \text{ is dense,} \\ \text{(iii)} & A \text{ satisfies the positive maximum principle,} \\ \text{(iv)} & \mathcal{R}(1-\alpha A) \text{ is dense for some, and hence for all } \alpha > 0. \end{array} \tag{2.3.41}$$

If $\overline{A}$ generates a Feller semigroup and $g \in \mathcal{C}(E)$, then the operator $\overline{A} + g$ (with domain $\mathcal{D}(\overline{A}+g) := \mathcal{D}(\overline{A})$) generates a strongly continuous semigroup $(S_t^g)_{t\geq 0}$ on $\mathcal{C}(E)$. If $g \leq 0$ then $(S_t^g)_{t\geq 0}$ is contractive. If $(\xi_t)_{t\geq 0}$ is the Feller process with generator $\overline{A}$, then one has the Feynman-Kac representation

$$S_t^g u(x) = E^x[u(\xi(t))\, e^{\int_0^t g(\xi(s))\mathrm{d}s}] \qquad (t \geq 0,\ x \in E,\ g, u \in \mathcal{C}(E)). \tag{2.3.42}$$

Let $\mathcal{C}^{(n)}([0,1]^2)$ denote the space of continuous real functions on $[0,1]^2$ whose partial derivatives up to n-th order exist and are continuous on $[0,1]^2$ (including the boundary), and put $\mathcal{C}^{(\infty)}([0,1]^2) := \bigcap_n \mathcal{C}^{(n)}([0,1]^2)$. Define a linear operator B on $\mathcal{C}([0,1]^2)$ with domain $\mathcal{D}(B) := \mathcal{C}^{(\infty)}([0,1]^2)$ by

$$Bu(x,y) := y(1-y)\tfrac{\partial^2}{\partial y^2}u(x,y) + \tfrac{1}{\gamma}(x-y)\tfrac{\partial}{\partial y}u(x,y). \tag{2.3.43}$$

Below, we will prove:

Lemma 2.38 (Feller semigroup) *The closure in $\mathcal{C}([0,1]^2)$ of the operator B generates a Feller semigroup on $\mathcal{C}([0,1]^2)$.*

Write

$$\begin{aligned} \mathcal{C}_+ &:= \big\{u \in \mathcal{C}([0,1]^2) : u \geq 0\big\}, \\ \mathcal{C}_{1+} &:= \big\{u \in \mathcal{C}^{(1)}([0,1]^2) : \tfrac{\partial}{\partial y}u, \tfrac{\partial}{\partial x}u \geq 0\big\}, \\ \mathcal{C}_{2+} &:= \big\{u \in \mathcal{C}^{(2)}([0,1]^2) : \tfrac{\partial^2}{\partial y^2}u, \tfrac{\partial^2}{\partial x \partial y}u, \tfrac{\partial^2}{\partial x^2}u \geq 0\big\}. \end{aligned} \tag{2.3.44}$$

Let $\overline{\mathcal{S}}$ denote the closure of a set $\mathcal{S} \subset \mathcal{C}([0,1]^2)$. We need the following lemma.

Lemma 2.39 (Preserved classes) *Let $g \in \mathcal{C}([0,1]^2)$ and let $(S^g_t)_{t\geq 0}$ be the strongly continuous semigroup with generator $\overline{B} + g$. Then, for each $t \geq 0$:*
(a) *If $g \in \overline{\mathcal{C}_{1+}}$, then S^g_t maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+}}$ into itself.*
(b) *If $g \in \overline{\mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$, then S^g_t maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$ into itself.*

To see why Lemma 2.39 implies Proposition 2.37, let $(\mathbf{x}(t), \mathbf{y}(t))_{t\geq 0}$ denote the Feller process in $[0,1]^2$ generated by $\overline{B}$. It is easy to see that $\mathbf{x}(t) = \mathbf{x}(0)$ a.s. for all $t \geq 0$. For fixed $\mathbf{x}(0) = x$, the process $(\mathbf{y}(t))_{t\geq 0}$ is the diffusion given by the SDE (2.3.20). Therefore, by Feynman-Kac, for each $g \in \mathcal{C}([0,1]^2)$,

$$E^y\big[e^{\int_0^t g(x, \mathbf{y}(s))\mathrm{d}s}\big] = S^g_t 1(x,y), \tag{2.3.45}$$

where 1 denotes the constant function $1 \in \mathcal{C}([0,1]^2)$. By (2.2.20),

$$\mathcal{U}_\gamma f(x) = (\tfrac{1}{\gamma} + 1)\Big(1 - \int \Gamma^\gamma_x(\mathrm{d}y) E^y\big[e^{-\int_0^{\tau_\gamma} f(\mathbf{y}_x(s))\mathrm{d}s}\big]\Big) \qquad (f \in \mathcal{C}_+[0,1]), \tag{2.3.46}$$

where Γ^γ_x is the invariant law of $(\mathbf{y}(t))_{t\geq 0}$ from Corollary 2.30 and τ_γ is an exponential time with mean γ, independent of $(\mathbf{y}(t))_{t\geq 0}$. Setting $g(x,y) := -f(y)$ in (2.3.45), using the ergodicity of $(\mathbf{y}(t))_{t\geq 0}$ (see Corollary 2.30), we find that for each $z \in [0,1]$ and $t \geq 0$,

$$\begin{aligned} \int \Gamma^\gamma_x(\mathrm{d}y) E^y\big[e^{-\int_0^t f(\mathbf{y}(s))\mathrm{d}s}\big] &= \lim_{r\to\infty} \int P^z[\mathbf{y}(r) \in \mathrm{d}y]\, E^y\big[e^{-\int_0^t g(x, \mathbf{y}(s))\mathrm{d}s}\big] \\ &= \lim_{r\to\infty} S^0_r S^g_t 1(x,z). \end{aligned} \tag{2.3.47}$$

It follows from Lemma 2.39 that for each fixed r, t, and z, the function $x \mapsto S^0_r S^g_t 1(x,z)$ is nondecreasing if f is nonincreasing, and nondecreasing and convex if f is nonincreasing and concave. Therefore, taking the expectation over the randomness of τ_γ, the claims follow from (2.3.46) and (2.3.47). ∎

We still need to prove Lemmas 2.38 and 2.39.

Proof of Lemma 2.38 It is easy to see that the operator B from (2.3.43) is densely defined, satisfies the positive maximum principle, and maps the constant function 1 into 0. Therefore, by Hille-Yosida (2.3.41), we must show that the range $\mathcal{R}(1 - \alpha B)$ is dense in $\mathcal{C}([0,1]^2)$ for some, and hence for all $\alpha > 0$. Let $\mathcal{P}_n$ denote the space of polynomials on $[0,1]^2$ of n-th and lower order, i.e., the space of functions $f : [0,1]^2 \to \mathbb{R}$ of the form

$$f(x,y) = \sum_{k,l\geq 0} a_{kl}\, x^k y^l \quad \text{with } a_{k,l} = 0 \text{ for } k + l > n. \tag{2.3.48}$$

Set $\mathcal{P}_\infty := \bigcup_n \mathcal{P}_n$. It is easy to see that B maps the space $\mathcal{P}_n$ into itself, for each $n \geq 0$. Since each $\mathcal{P}_n$ is finite-dimensional, a simple argument (see [EK86, Proposition 1.3.5]) shows that the image of $\mathcal{P}_\infty$ under $1 - \alpha B$ is dense in $\mathcal{C}([0,1]^2)$ for all but countably many, and hence for all $\alpha > 0$. ∎

As a first step towards proving Lemma 2.39, we prove:

Lemma 2.40 (Smooth solutions to Laplace equation) *Let $\alpha > 0$, $g \in \mathcal{C}^{(2)}([0,1])$, $g \leq 0$, $v \in \mathcal{C}([0,1]^2)$, and assume that $u \in \mathcal{C}^{(\infty)}([0,1]^2)$ solves the Laplace equation*

$$(1 - \alpha(B + g))u = v. \tag{2.3.49}$$

(a) *If $g \in \mathcal{C}_{1+}$, then $v \in \mathcal{C}_+ \cap \mathcal{C}_{1+}$ implies $u \in \mathcal{C}_+ \cap \mathcal{C}_{1+}$.*

(b) *If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then $v \in \mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$ implies $u \in \mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$.*

Proof Let $u^y := \frac{\partial}{\partial y}u$, $u^{xy} := \frac{\partial^2}{\partial x \partial y}u$, etc. denote the partial derivatives of u and similarly for v and g, whenever they exist. Set $c := \frac{1}{\gamma}$. Define linear operators B' and B'' on $\mathcal{C}([0,1]^2)$ with domains $\mathcal{D}(B') = \mathcal{D}(B'') := \mathcal{C}^{(\infty)}([0,1]^2)$ by

$$\begin{aligned} B' &:= y(1-y)\tfrac{\partial^2}{\partial y^2} + \big(c(x-y) + 2(\tfrac{1}{2} - y)\big)\tfrac{\partial}{\partial y}, \\ B'' &:= y(1-y)\tfrac{\partial^2}{\partial y^2} + \big(c(x-y) + 4(\tfrac{1}{2} - y)\big)\tfrac{\partial}{\partial y}. \end{aligned} \tag{2.3.50}$$

Then

$$\begin{aligned} \tfrac{\partial}{\partial y}Bu &= (B' - c)u^y, \quad & \tfrac{\partial}{\partial y}B'u &= (B'' - c - 2)u^y, \\ \tfrac{\partial}{\partial x}Bu &= Bu^x + cu^y, \quad & \tfrac{\partial}{\partial x}B'u &= B'u^x + cu^y. \end{aligned} \tag{2.3.51}$$

Therefore, it is easy to see that

$$\begin{array}{lrl} \text{(i)} & (1 - \alpha(B' - c + g))u^y &= v^y + \alpha g^y u, \\ \text{(ii)} & (1 - \alpha(B + g))u^x &= v^x + \alpha(cu^y + g^x u), \\ \text{(iii)} & (1 - \alpha(B'' - 2c - 2 + g))u^{yy} &= v^{yy} + \alpha(2g^y u^y + g^{yy}u), \\ \text{(iv)} & (1 - \alpha(B' - c + g))u^{xy} &= v^{xy} + \alpha(cu^{yy} + g^y u^x + g^{xy}u + g^x u^y), \\ \text{(v)} & (1 - \alpha(B + g))u^{xx} &= v^{xx} + \alpha(2cu^{xy} + 2g^x u^x + g^{xx}u), \end{array} \tag{2.3.52}$$

where in (i) and (ii) we assume that $v \in \mathcal{C}^{(1)}([0,1]^2)$ and in (iii)–(v) we assume that $v \in \mathcal{C}^{(2)}([0,1]^2)$. By Lemma 2.38, the closure of the operator B generates a Feller processes in $[0,1]^2$. Exactly the same proof shows that B' and B'' also generate Feller processes on $[0,1]^2$. Therefore, by Feynman-Kac, u is nonnegative if v is nonnegative and $u^y, \ldots, u^{xx}$ are nonnegative if the right-hand sides of the equations (i)–(v) are well-defined and nonnegative. (Instead of using Feynman-Kac, this follows more elementarily from the fact that B, B', and B'' satisfy the positive maximum principle.) In particular, if $g^y, g^x \geq 0$ and $v \in \mathcal{C}^{(1)}([0,1]^2)$, $v, v^y, v^x \geq 0$, then it follows that $u, u^y, u^x \geq 0$. If moreover $g^{yy}, g^{xy}, g^{xx} \geq 0$ and $v \in \mathcal{C}^{(2)}([0,1]^2)$, $v^{yy}, v^{xy}, v^{yy} \geq 0$, then also $u^{yy}, u^{xy}, u^{yy} \geq 0$. ∎

In order to prove Lemma 2.39, based on Lemma 2.40, we will show that the Laplace equation (2.3.49) has smooth solutions u for sufficiently many functions v. Here 'suffiently many' will mean dense in the topology of uniform convergence of functions and their derivatives up to second order. To this aim, we make $\mathcal{C}^{(2)}([0,1]^2)$ into a Banach space by equipping it with the norm

$$\|u\|_{(2)} := \|u\| + \|u^y\| + \|u^x\| + \|u^{yy}\| + 2\|u^{xy}\| + \|u^{xx}\|. \tag{2.3.53}$$

Here, to reduce notation, we denote the supremumnorm by $\|f\| := \|f\|_\infty$. Note the factor 2 in the second term from the right in (2.3.53), which is crucial for the next key lemma.

Lemma 2.41 (Semigroup on twice diffferentiable functions) *The closure in $\mathcal{C}^{(2)}([0,1]^2)$ of the operator B generates a strongly continuous contraction semigroup on $\mathcal{C}^{(2)}([0,1]^2)$.*

Proof We must check the conditions (i)–(iii) from (2.3.39). It is well-known (see for example [EK86, Proposition 7.1 from the appendix]) that the space $\mathcal{P}_\infty$ of polynomials is dense in $\mathcal{C}^{(2)}([0,1]^2)$. Therefore $\mathcal{D}(B) = \mathcal{C}^{(\infty)}([0,1]^2)$ is dense, and copying the proof of Lemma 2.38 we see that $\mathcal{R}(1-\alpha B)$ is dense for all but countably many α. To complete the proof, we must show that B is dissipative, i.e., that

$$\|(1-\varepsilon B)u\|_{(2)} \geq \|u\|_{(2)} \qquad (\varepsilon > 0,\ u \in \mathcal{C}^{(\infty)}([0,1]^2)). \tag{2.3.54}$$

Using (2.3.51), we calculate

$$\begin{aligned}
\tfrac{\partial}{\partial y}(1-\varepsilon B)u &= (1-\varepsilon(B'-c))u^y, \\
\tfrac{\partial}{\partial x}(1-\varepsilon B)u &= (1-\varepsilon B)u^x - \varepsilon c u^y, \\
\tfrac{\partial^2}{\partial y^2}(1-\varepsilon B)u &= (1-\varepsilon(B''-2c-2))u^{yy}, \\
\tfrac{\partial^2}{\partial x \partial y}(1-\varepsilon B)u &= (1-\varepsilon(B'-c))u^{xy} - \varepsilon c u^{yy}, \\
\tfrac{\partial^2}{\partial x^2}(1-\varepsilon B)u &= (1-\varepsilon B)u^{xx} - 2\varepsilon c u^{xy}.
\end{aligned} \tag{2.3.55}$$

Using the disipativity of B, B', and B'' with respect to the supremumnorm (which follows from the positive maximum principle) we see that $\|(1-\varepsilon(B'-c))u^y\| = (1+\varepsilon c)\|(1-\frac{\varepsilon}{1+\varepsilon c}B)u^y\| \geq (1+\varepsilon c)\|u^y\|$ etc. We conclude therefore from (2.3.55) that

$$\begin{aligned}
\|(1-\varepsilon B)u\|_{(2)} \geq\, & \|(1-\varepsilon B)u\| + \|(1-\varepsilon(B'-c))u^y\| + \|(1-\varepsilon B)u^x\| - \varepsilon c\|u^y\| \\
& + \|(1-\varepsilon(B''-2c-2))u^{yy}\| + 2\|(1-\varepsilon(B'-c))u^{xy}\| - 2\varepsilon c\|u^{yy}\| \\
& + \|(1-\varepsilon B)u^{xx}\| - 2\varepsilon c\|u^{xy}\| \\
\geq\, & \|u\| + (1+\varepsilon c)\|u^y\| + \|u^x\| - \varepsilon c\|u^y\| \\
& + (1+\varepsilon(2c+2))\|u^{yy}\| + 2(1+\varepsilon c)\|u^{xy}\| - 2\varepsilon c\|u^{yy}\| \\
& + \|u^{xx}\| - 2\varepsilon c\|u^{xy}\| \geq \|u\|_{(2)}
\end{aligned} \tag{2.3.56}$$

for each $\varepsilon > 0$, which shows that B is dissipative with respect to the norm $\|\cdot\|_{(2)}$. ∎

Proof of Lemma 2.39 Let $g \in \mathcal{C}^{(2)}([0,1]^2)$. Then $u \mapsto gu$ is a bounded operator on both $\mathcal{C}([0,1]^2)$ and $\mathcal{C}^{(2)}([0,1]^2)$, so we can choose a $\lambda > 0$ such that

$$\|gu\| \leq \lambda\|u\| \quad \text{and} \quad \|gu\|_{(2)} \leq \lambda\|u\|_{(2)} \tag{2.3.57}$$

for all u in $\mathcal{C}([0,1]^2)$ and $\mathcal{C}^{(2)}([0,1]^2)$, respectively. Put $\tilde{g} := g - \lambda$. By Lemma 2.38, $\overline{B} + \tilde{g}$ generates a strongly continuous contraction semigroup $(S_t^{\tilde{g}})_{t\geq 0} = (e^{-\lambda t} S_t^g)_{t\geq 0}$ on $\mathcal{C}([0,1]^2)$. Note that $\mathcal{R}(1 - \alpha(B + \tilde{g}))$ is the space of all $v \in \mathcal{C}([0,1]^2)$ for which the Laplace equation $(1 - \alpha(B + \tilde{g}))u = v$ has a solution $u \in \mathcal{C}^{(\infty)}([0,1]^2)$. Therefore, by Lemma 2.40, for each $\alpha > 0$:

(i) If $g \in \mathcal{C}_{1+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\mathcal{R}(1 - \alpha(B + \tilde{g})) \cap \mathcal{C}_+ \cap \mathcal{C}_{1+}$ into $\mathcal{C}_+ \cap \mathcal{C}_{1+}$.

(ii) If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\mathcal{R}(1 - \alpha(B + \tilde{g})) \cap \mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$ into $\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$.

(2.3.58)

By Lemma 2.41, the restriction of the semigroup $(S_t^{\tilde{g}})_{t\geq 0}$ to $\mathcal{C}^{(2)}([0,1]^2)$ is strongly continuous and contractive in the norm $\|\cdot\|_{(2)}$. Therefore, by Hille-Yosida (2.3.39), $\mathcal{R}(1 - \alpha(B + \tilde{g}))$ is dense in $\mathcal{C}^{(2)}([0,1]^2)$ for each $\alpha > 0$. It follows that $\mathcal{R}(1 - \alpha(B + \tilde{g})) \cap \mathcal{C}_+ \cap \mathcal{C}_{1+}$ is dense in $\mathcal{C}_+ \cap \mathcal{C}_{1+}$ and likewise $\mathcal{R}(1 - \alpha(B + \tilde{g})) \cap \mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$ is dense in $\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, both in the norm $\|\cdot\|_{(2)}$. Note that we need density in the norm $\|\cdot\|_{(2)}$ here: if we would only know that $\mathcal{R}(1 - \alpha(B + \tilde{g}))$ is a dense subset of $\mathcal{C}([0,1]^2)$ in the norm $\|\cdot\|$, then $\mathcal{R}(1 - \alpha(B + \tilde{g})) \cap \mathcal{C}_+ \cap \mathcal{C}_{1+}$ might be empty. By approximation in the norm $\|\cdot\|_{(2)}$ it follows from (2.3.58) that:

(i) If $g \in \mathcal{C}_{1+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\mathcal{C}_+ \cap \mathcal{C}_{1+}$ into itself.

(ii) If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$ into itself. (2.3.59)

Using also continuity in the norm $\|\cdot\|$ we find that:

(i) If $g \in \mathcal{C}_{1+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+}}$ into itself.

(ii) If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then $(1 - \alpha(\overline{B} + \tilde{g}))^{-1}$ maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$ into itself. (2.3.60)

For $\varepsilon > 0$ let

$$G_\varepsilon := \varepsilon^{-1}\big((1 - \varepsilon(\overline{B} + \tilde{g}))^{-1} - 1\big) \tag{2.3.61}$$

be the Yosida approximation to $\overline{B} + \tilde{g}$. Then

$$e^{G_\varepsilon t} = e^{-\varepsilon^{-1}t} \sum_{n=0}^{\infty} \frac{t^n}{n!}(1 - \varepsilon(\overline{B} + \tilde{g}))^{-n} \qquad (t \geq 0), \tag{2.3.62}$$

and therefore, by (2.3.60), for each $t \geq 0$:

(i) If $g \in \mathcal{C}_{1+}$, then $e^{G_\varepsilon t}$ maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+}}$ into itself.

(ii) If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then $e^{G_\varepsilon t}$ maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$ into itself. (2.3.63)

Finally

$$e^{-\lambda t} S_t^g u = S_t^{\tilde{g}} u = \lim_{\varepsilon \to 0} e^{G_\varepsilon t} u \qquad (t \geq 0,\ u \in \mathcal{C}([0,1]^2)), \tag{2.3.64}$$

so (2.3.63) implies that for each $t \geq 0$:

(i) If $g \in \mathcal{C}_{1+}$, then S_t^g maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+}}$ into itself.

(ii) If $g \in \mathcal{C}_{1+} \cap \mathcal{C}_{2+}$, then S_t^g maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$ into itself. (2.3.65)

Using the continuity of S_t^g in g (which follows from Feynman-Kac (2.3.42)) we arrive at the statements in Lemma 2.39. ∎

2.4 Convergence to a time-homogeneous process

2.4.1 Convergence of certain Markov chains

Section 2.4 is devoted to the proof of Theorem 2.19. In the present subsection, we start by formulating a theorem about the convergence of certain Markov chains to continuous-time processes. In Section 2.4.2 we specialize to Poisson-cluster branching processes and superprocesses. In Section 2.4.3, finally, we carry out the necessary calculations for the specific processes from Theorem 2.19.

Let E be a *compact* metrizable space. We equip the space $\mathcal{C}(E)$ of continuous real functions on E with the supremumnorm $\|\cdot\|_\infty$. By definition, $\mathcal{D}_E[0,\infty)$ is the space of cadlag functions $w:[0,\infty)\to E$, equipped with the Skorohod topology. Let $A:\mathcal{D}(A)\to\mathcal{C}(E)$ be an operator defined on a domain $\mathcal{D}(A)\subset\mathcal{C}(E)$. We say that a process $\mathbf{y}=(\mathbf{y}_t)_{t\geq 0}$ solves the martingale problem for A if $\mathbf{y}$ has sample paths in $\mathcal{D}_E[0,\infty)$ and for each $f\in\mathcal{D}(A)$, the process $(M^f_t)_{t\geq 0}$ given by

$$M^f_t := f(\mathbf{y}_t) - \int_0^t Af(\mathbf{y}_s)\mathrm{d}s \qquad (t\geq 0) \tag{2.4.1}$$

is a martingale with respect to the filtration generated by $\mathbf{y}$. We say that existence (uniqueness) holds for the martingale problem for A if for each probability measure μ on E there is at least one (at most one (in law)) solution $\mathbf{y}$ to the martingale problem for A with initial law $\mathcal{L}(\mathbf{y}_0)=\mu$. If both existence and uniqueness hold we say that the martingale problem is well-posed. For each $n\geq 0$, let $X^{(n)}=(X^{(n)}_0,\ldots,X^{(n)}_{m(n)})$ (with $1\leq m(n)<\infty$) be a (time-inhomogeneous) Markov process in E with k-th step transition probabilities

$$P_k(x,\mathrm{d}y) = P\big[X^{(n)}_k\in\mathrm{d}y\big|X^{(n)}_{k-1}=x\big] \qquad (1\leq k\leq m(n)). \tag{2.4.2}$$

We assume that the P_k are continuous probability kernels on E. Let $(\varepsilon^{(n)}_k)_{1\leq k\leq m(n)}$ be positive constants. Set

$$A^{(n)}_k f(x) := (\varepsilon^{(n)}_k)^{-1}\Big(\int_E P_k(x,\mathrm{d}y)f(y) - f(x)\Big) \qquad (1\leq k\leq m(n),\ f\in\mathcal{C}(E)). \tag{2.4.3}$$

Define $t^{(n)}_0 := 0$ and

$$t^{(n)}_k := \sum_{l=1}^k \varepsilon^{(n)}_l \qquad (1\leq k\leq m(n)), \tag{2.4.4}$$

and put

$$k^{(n)}(t) := \max\big\{k\ :\ 0\leq k\leq m(n),\ t^{(n)}_k\leq t\big\} \qquad (t\geq 0). \tag{2.4.5}$$

Define processes $\mathbf{y}^{(n)}=(\mathbf{y}^{(n)}_t)_{t\geq 0}$ with sample paths in $\mathcal{D}_E[0,\infty)$ by

$$\mathbf{y}^{(n)}_t := X^{(n)}_{k^{(n)}(t)} \qquad (t\geq 0). \tag{2.4.6}$$

By definition, a space $\mathcal{A}$ of real functions is called an algebra if $\mathcal{A}$ is a linear space and $f,g\in\mathcal{A}$ implies $fg\in\mathcal{A}$.

Theorem 2.42 (Convergence of Markov chains) *Assume that $\mathcal{L}(X_0^{(n)}) \Rightarrow \mu$ as $n \to \infty$ for some probability law μ on E. Suppose that there exists at most one (in law) solution to the martingale problem for A with initial law μ. Assume that the linear span of $\mathcal{D}(A)$ contains an algebra that separates points. Assume that*

$$\text{(i)} \lim_{n\to\infty} \sum_{k=1}^{m(n)} \varepsilon_k^{(n)} = \infty, \qquad \text{(ii)} \lim_{n\to\infty} \sup_{k:\, t_k^{(n)} \leq T} \varepsilon_k^{(n)} = 0, \tag{2.4.7}$$

and

$$\lim_{n\to\infty} \sup_{k:\, t_k^{(n)} \leq T} \|A_k^{(n)} f - Af\|_\infty = 0 \qquad (f \in \mathcal{D}(A)) \tag{2.4.8}$$

for each $T > 0$. Then there exists a unique solution $\mathbf{y}$ to the martingale problem for A with initial law μ and moreover $\mathcal{L}(\mathbf{y}^{(n)}) \Rightarrow \mathcal{L}(\mathbf{y})$, where $\Rightarrow$ denotes weak convergence of probability measures on $\mathcal{D}_E[0,\infty)$.

Proof We apply [EK86, Corollary 4.8.15]. Fix $f \in \mathcal{D}(A)$. We start by observing that

$$f(X_k^{(n)}) - \sum_{i=1}^{k} \varepsilon_i^{(n)} A_i^{(n)} f(X_{i-1}^{(n)}) \qquad (0 \leq k \leq m(n)) \tag{2.4.9}$$

is a martingale with respect to the filtration generated by $X^{(n)}$ and therefore,

$$f(\mathbf{y}_t^{(n)}) - \sum_{i=1}^{k^{(n)}(t)} \varepsilon_i^{(n)} A_i^{(n)} f(\mathbf{y}_{t_{i-1}^{(n)}}^{(n)}) \qquad (t \geq 0) \tag{2.4.10}$$

is a martingale with respect to the filtration generated by $\mathbf{y}^{(n)}$. Put

$$\lfloor t \rfloor^{(n)} := t_{k^{(n)}(t)}^{(n)} \qquad (t \geq 0) \tag{2.4.11}$$

and set

$$\phi_t^{(n)} := A_{k^{(n)}(t)+1}^{(n)} f(\mathbf{y}_{\lfloor t \rfloor^{(n)}}^{(n)}) 1_{\{t < t_{m(n)}^{(n)}\}} \qquad (t \geq 0) \tag{2.4.12}$$

and

$$\xi_t^{(n)} := f(\mathbf{y}_t^{(n)}) + \int_{\lfloor t \rfloor^{(n)}}^{t} \phi_s^{(n)} \mathrm{d}s \qquad (t \geq 0). \tag{2.4.13}$$

Then we can rewrite the martingale in (2.4.10) as

$$\xi_t^{(n)} - \int_0^t \phi_s^{(n)} \mathrm{d}s. \tag{2.4.14}$$

By [EK86, Corollary 4.8.15] and the compactness of the state space, it suffices to check the following conditions on $\phi^{(n)}$ and $\xi^{(n)}$:

$$\begin{aligned}
&\text{(i)} && \sup_{n\geq N}\ \sup_{t\leq T} E\big[|\xi_t^{(n)}|\big] < \infty, \\
&\text{(ii)} && \sup_{n\geq N}\ \sup_{t\leq T} E\big[|\phi_t^{(n)}|\big] < \infty, \\
&\text{(iii)} && \lim_{n\to\infty} E\Big[\big(\xi_T^{(n)} - f(\mathbf{y}_T^{(n)})\big)\prod_{i=1}^r h_i(\mathbf{y}_{s_i}^{(n)})\Big] = 0, \\
&\text{(iv)} && \lim_{n\to\infty} E\Big[\big(\phi_T^{(n)} - Af(\mathbf{y}_T^{(n)})\big)\prod_{i=1}^r h_i(\mathbf{y}_{s_i}^{(n)})\Big] = 0, \\
&\text{(v)} && \lim_{n\to\infty} E\Big[\sup_{t\in\mathbb{Q}\cap[0,T]}\big|\xi_t^{(n)} - f(\mathbf{y}_t^{(n)})\big|\Big] = 0, \\
&\text{(vi)} && \sup_{n\geq N} E\big[\|\phi^{(n)}\|_{p,T}\big] < \infty \qquad \text{for some } p\in(1,\infty],
\end{aligned} \tag{2.4.15}$$

for some $N \geq 0$ and for each $T > 0$, $r \geq 1$, $0 \leq s_1 < \cdots < s_r \leq T$, and $h_1, \ldots, h_r \in \mathcal{H} \subset \mathcal{C}(E)$. Here $\mathcal{H}$ is separating, i.e., $\int h\mathrm{d}\mu = \int h\mathrm{d}\nu$ for all $h \in \mathcal{H}$ implies $\mu = \nu$ whenever μ, ν are probability measures on E. In (vi):

$$\|g\|_{p,T} := \Big(\int_0^T |g(t)|^p \mathrm{d}t\Big)^{1/p} \qquad (1 \leq p < \infty) \tag{2.4.16}$$

and $\|g\|_{\infty,T}$ denotes the essential supremum of g over $[0,T]$.

The conditions (2.4.15) (i)–(vi) are implied by the stronger conditions

$$\begin{aligned}
&\text{(i)} && \lim_{n\to\infty}\ \sup_{0\leq t\leq T}\big\|\xi_t^{(n)} - f(\mathbf{y}_t^{(n)})\big\|_\infty = 0, \\
&\text{(ii)} && \lim_{n\to\infty}\ \sup_{0\leq t\leq T}\big\|\phi_t^{(n)} - Af(\mathbf{y}_t^{(n)})\big\|_\infty = 0,
\end{aligned} \tag{2.4.17}$$

where we denote the essential supremumnorm of a real-valued random variable X by $\|X\|_\infty := \inf\{K \geq 0 : |X| \leq K \text{ a.s.}\}$. Condition (2.4.17) (ii) is implied by (2.4.7) (i) and (2.4.8). To see that also (2.4.17) (i) holds, set

$$M_n := \sup_{0\leq t\leq T}\big\|\phi_t^{(n)}\big\|_\infty, \tag{2.4.18}$$

and estimate

$$\sup_{0\leq t\leq T}\big\|\xi_t^{(n)} - f(\mathbf{y}_t^{(n)})\big\|_\infty \leq M_n \sup\{\varepsilon_k^{(n)} \,:\, 1 \leq k \leq m(n),\ t_k^{(n)} \leq T\}. \tag{2.4.19}$$

Condition (2.4.17) (ii) implies that $\limsup_n M_n < \infty$ and therefore the right-hand side of (2.4.19) tends to zero by assumption (2.4.7) (ii). ∎

2.4.2 Convergence of certain branching processes

In this section we apply Theorem 2.42 to certain branching processes and superprocesses.

Throughout this section, E is a compact metrizable space and $A : \mathcal{D}(A) \to \mathcal{C}(E)$ is a linear operator on $\mathcal{C}(E)$ such that the closure $\overline{A}$ of A generates a Feller process $\xi = (\xi_t)_{t\geq 0}$ in E with Feller semigroup $(P_t)_{t\geq 0}$ given by $P_t f(x) := E^x[f(\xi_t)]$ $(t \geq 0,\ f \in \mathcal{C}(E))$.

Let $\alpha \in \mathcal{C}_+(E)$ and $\beta, f \in \mathcal{C}(E)$. By definition, a function $t \mapsto u_t$ from $[0,\infty)$ into $\mathcal{C}(E)$ is a *classical* solution to the semilinear Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} u_t = \overline{A} u_t + \beta u_t - \alpha u_t^2 & (t \geq 0), \\ u_0 = f \end{cases} \tag{2.4.20}$$

if $t \mapsto u_t$ is continuously differentiable (in $\mathcal{C}(E)$), $u_t \in \mathcal{D}(\overline{A})$ for all $t \geq 0$, and (2.4.20) holds. We say that u is a *mild* solution to (2.4.20) if $t \mapsto u_t$ is continuous and

$$u_t = P_t f + \int_0^t P_{t-s}(\beta u_s - \alpha u_s^2)\mathrm{d}s \qquad (t \geq 0). \tag{2.4.21}$$

Lemma 2.43 (Mild and classical solutions) *Equation (2.4.20) has a unique $\mathcal{C}_+(E)$-valued mild solution u for each $f \in \mathcal{C}_+(E)$, and $f > 0$ implies that $u_t > 0$ for all $t \geq 0$. If moreover $f \in \mathcal{D}(\overline{A})$ then u is a classical solution. For each $t \geq 0$, u_t depends continuously on $f \in \mathcal{C}_+(E)$.*

Proof It follows from [Paz83, Theorems 6.1.2, 6.1.4, and 6.1.5] that for each $f \in \mathcal{C}(E)$, (2.4.20) has a unique solution $(u_t)_{0\leq t<T}$ up to an explosion time T, and that this is a classical solution if $f \in \mathcal{D}(\overline{A})$. Moreover, u_t depends continuously on f. Using comparison arguments based on the fact that $\overline{A}$ satisfies the positive maximum principle (which follows from Hille-Yosida (2.3.41)) one easily proves the other statements; compare [FS04, Lemmas 23 and 24]. ∎

We denote the (mild or classical) solution of (2.4.20) by $\mathcal{U}_t f := u_t$; then $\mathcal{U}_t : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$ are continuous operators and $\mathcal{U} = (\mathcal{U}_t)_{t\geq 0}$ is a (nonlinear) semigroup on $\mathcal{C}_+(E)$.

Since E is compact, the spaces $\{\mu \in \mathcal{M}(E) : \mu(E) \leq M\}$ are compact for each $M \geq 0$. In particular, $\mathcal{M}(E)$ is locally compact. We denote its one-point compactification by $\mathcal{M}(E)_\infty = \mathcal{M}(E) \cup \{\infty\}$. We define functions $F_f \in \mathcal{C}(\mathcal{M}(E)_\infty)$ by $F_f(\infty) := 0$ and

$$F_f(\mu) := e^{-\langle \mu, f \rangle} \qquad (f \in \mathcal{C}_+(E),\ f > 0,\ \mu \in \mathcal{M}(E)). \tag{2.4.22}$$

We introduce an operator $\mathcal{G}$ with domain

$$\mathcal{D}(\mathcal{G}) := \{F_f : f \in \mathcal{D}(A),\ f > 0\}, \tag{2.4.23}$$

given by $\mathcal{G}F_f(\infty) := 0$ and

$$\mathcal{G}F_f(\mu) := -\langle \mu, Af + \beta f - \alpha f^2 \rangle\, e^{-\langle \mu, f \rangle} \qquad (\mu \in \mathcal{M}(E)). \tag{2.4.24}$$

Note that $\mathcal{G}F_f \in \mathcal{C}(\mathcal{M}(E)_\infty)$ for all $F_f \in \mathcal{D}(\mathcal{G})$.

Proposition 2.44 ($(\overline{A},\alpha,\beta)$-superprocesses) *The martingale problem for the operator $\mathcal{G}$ is well-posed. The solutions to this martingale problem define a Feller process $\mathcal{Y}=(\mathcal{Y}_t)_{t\geq 0}$ in $\mathcal{M}(E)_\infty$ with continuous sample paths, called the $(\overline{A},\alpha,\beta)$-superprocess. If $\mathcal{Y}_0=\infty$ then $\mathcal{Y}_t=\infty$ for all $t\geq 0$. If $\mathcal{Y}_0=\mu\in\mathcal{M}(E)$ then*

$$E^\mu\big[e^{-\langle \mathcal{Y}_t, f\rangle}\big]=e^{-\langle \mu, \mathcal{U}_t f\rangle}\qquad (f\in\mathcal{C}_+(E)). \tag{2.4.25}$$

Proof Results of this type are well-known, see for example [EK86, Theorem 9.4.3], [Fit88], and [ER91, Théorème 7]. Since, however, it is not completely straightforward to derive the proposition above from these references, we give a concise autonomous proof of most of our statements. Only for the continuity of sample paths we refer the reader to [Fit88, Corollary (4.7)] or [ER91, Corollaire 9].

We are going to extend $\mathcal{G}$ to an operator $\hat{\mathcal{G}}$ that is linear and satisfies the conditions of the Hille-Yosida Theorem (2.3.41). For any $\gamma\in\mathcal{C}_+(E)$ and $\mu\in\mathcal{M}(E)$, let $\mathrm{Clust}_\gamma(\mu)$ denote a random measure such that on $\{\gamma=0\}$, $\mathrm{Clust}_\gamma(\mu)$ is equal to μ, and on $\{\gamma>0\}$, $\mathrm{Clust}_\gamma(\mu)$ is a Poisson cluster measure with intensity $\frac{1}{\gamma}\mu$ and cluster mechanism $\mathcal{Q}(x,\cdot)=\mathcal{L}(\tau_{\gamma(x)}\delta_x)$, where $\tau_{\gamma(x)}$ is exponentially distributed with mean $\gamma(x)$. It is not hard to see that

$$E\big[e^{-\langle \mathrm{Clust}_\gamma(\mu), f\rangle}\big]=e^{-\langle \mu, \mathcal{V}_\gamma f\rangle}\qquad (f\in\mathcal{C}(E),\ f>0), \tag{2.4.26}$$

where $\mathcal{V}_\gamma f(x):=(\frac{1}{f(x)}+\gamma(x))^{-1}$. Note that since $\mathcal{V}_\gamma 1$ is bounded, the previously mentioned Poisson cluster measure mentioned above is well-defined. By definition, we put $\mathrm{Clust}_\gamma(\infty):=\infty$.

Define a linear operator $\mathcal{G}_\alpha$ on $\mathcal{C}(\mathcal{M}(E))_\infty)$ by

$$\mathcal{G}_\alpha F(\mu):=\lim_{\varepsilon\to 0}\varepsilon^{-1}\big(E[F(\mathrm{Clust}_{\varepsilon\alpha}(\mu))]-F(\mu)\big) \tag{2.4.27}$$

with as domain $\mathcal{D}(\mathcal{G}_\alpha)$ the space of all $F\in\mathcal{C}(\mathcal{M}(E)_\infty)$ for which the limit exists. Define a linear operator $\mathcal{G}_\beta$ by

$$\mathcal{G}_\beta F(\mu):=\lim_{\varepsilon\to 0}\varepsilon^{-1}\big(F((1+\varepsilon\beta)\mu)-F(\mu)\big) \tag{2.4.28}$$

with domain $\mathcal{D}(\mathcal{G}_\beta):=\mathcal{C}(\mathcal{M}(E))_\infty)$. Define $P^*_t:\mathcal{M}(E)_\infty\to\mathcal{M}(E)_\infty$ by $\langle P^*_t\mu,f\rangle:=\langle\mu,P_tf\rangle$ $(t\geq 0,\ f\in\mathcal{C}(E),\ \mu\in\mathcal{M}(E))$ and $P^*_t\infty:=\infty$ $(t\geq 0)$. Finally, let $\mathcal{G}_{\overline{A}}$ be the linear operator on $\mathcal{C}(\mathcal{M}(E))_\infty)$ defined by

$$\mathcal{G}_{\overline{A}}F(\mu):=\lim_{\varepsilon\to 0}\varepsilon^{-1}\big(F(P^*_\varepsilon\mu)-F(\mu)\big), \tag{2.4.29}$$

with as domain $\mathcal{D}(\mathcal{G}_{\overline{A}})$ the space of all F for which the limit exists. Define an operator $\hat{\mathcal{G}}$ by

$$\hat{\mathcal{G}}:=\mathcal{G}_\alpha+\mathcal{G}_\beta+\mathcal{G}_{\overline{A}}, \tag{2.4.30}$$

with domain $\mathcal{D}(\hat{\mathcal{G}}):=\mathcal{D}(\mathcal{G}_\alpha)\cap\mathcal{D}(\mathcal{G}_{\overline{A}})$. If $f\in\mathcal{D}(\overline{A})$, $f>0$, and F_f is as in (2.4.22), then it is not hard to see that $\hat{\mathcal{G}}F_f(\infty)=0$ and

$$\hat{\mathcal{G}}F_f(\mu):=-\langle\mu,\overline{A}f+\beta f-\alpha f^2\rangle\,e^{-\langle\mu, f\rangle}\qquad (\mu\in\mathcal{M}(E)). \tag{2.4.31}$$

In particular, $\hat{\mathcal{G}}$ extends the operator $\mathcal{G}$ from (2.4.24). Since $\mathcal{D}(\overline{A})$ is dense in $\mathcal{C}(E)$, it is easy to see that $\{F_f : f \in \mathcal{D}(\overline{A}),\ f > 0\}$ is dense in $\mathcal{C}(\mathcal{M}(E)_\infty)$. Hence $\mathcal{D}(\hat{\mathcal{G}})$ is dense. Using (2.4.27)–(2.4.29) it is not hard to show that $\hat{\mathcal{G}}$ satisfies the positive maximum principle. Moreover, by Lemma 2.43, for $f \in \mathcal{D}(\overline{A})$ with $f > 0$, the function $t \mapsto F_{\mathcal{U}_t f}$ from $[0,\infty)$ into $\mathcal{C}(\mathcal{M}(E)_\infty)$ is continuously differentiable, satisfies $F_{\mathcal{U}_t f} \in \mathcal{D}(\hat{\mathcal{G}})$ for all $t \geq 0$, and

$$\tfrac{\partial}{\partial t} F_{\mathcal{U}_t f} = \hat{\mathcal{G}} F_{\mathcal{U}_t f} \qquad (t \geq 0). \tag{2.4.32}$$

From this it is not hard to see that $\hat{\mathcal{G}}$ also satisfies condition (2.3.41) (ii), so the closure of $\hat{\mathcal{G}}$ generates a Feller semigroup $(S_t)_{t\geq 0}$ on $\mathcal{C}(\mathcal{M}(E)_\infty)$. It is easy to see that $S_t F_f = F_{\mathcal{U}_t f}$ $(t \geq 0)$. By [EK86, Theorem 4.2.7], this semigroup corresponds to a Feller process $\mathcal{Y}$ with cadlag sample paths in $\mathcal{M}(E)_\infty$. This means that $E^\mu[F_f(\mathcal{Y}_t)] = F_{\mathcal{U}_t f}(\mu)$ for all $f \in \mathcal{D}(\overline{A})$ with $f > 0$. If $\mu = \infty$ this shows that $\mathcal{Y}_t = \infty$ for all $t \geq 0$. If $\mu \in \mathcal{M}(E)$ we obtain (2.4.25) for $f \in \mathcal{D}(\overline{A})$, $f > 0$; the general case follows by approximation. ∎

Now let $(q_\varepsilon)_{\varepsilon>0}$ be continuous weight functions and let $(\mathcal{Q}_\varepsilon)_{\varepsilon>0}$ be continuous cluster mechanisms on E. Assume that

$$Z_\varepsilon(x) := \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle < \infty \qquad (x \in E) \tag{2.4.33}$$

and define probability kernels K_ε on E by

$$\int K_\varepsilon(x,\mathrm{d}y) f(y) := \frac{1}{Z_\varepsilon(x)} \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f\rangle \qquad (f \in B(E)). \tag{2.4.34}$$

For each $n \geq 0$, let $(\varepsilon^{(n)}_k)_{1\leq k\leq m(n)}$ (with $1 \leq m(n) < \infty$) be positive constants. Let $\mathcal{X}^{(n)} = (\mathcal{X}^{(n)}_0, \ldots, \mathcal{X}^{(n)}_{m(n)})$ be a Poisson-cluster branching process with weight functions $q_{\varepsilon^{(n)}_1}, \ldots, q_{\varepsilon^{(n)}_{m(n)}}$ and cluster mechanisms $\mathcal{Q}_{\varepsilon^{(n)}_1}, \ldots, \mathcal{Q}_{\varepsilon^{(n)}_{m(n)}}$. Define $t^{(n)}_k$ and $k^{(n)}(t)$ as in (2.4.4)–(2.4.5). Define processes $\mathcal{Y}^{(n)}$ by

$$\mathcal{Y}^{(n)}_t := \mathcal{X}^{(n)}_{k^{(n)}(t)} \qquad (t \geq 0). \tag{2.4.35}$$

Theorem 2.45 (Convergence of Poisson-cluster branching processes) *Assume that $\mathcal{L}(\mathcal{X}^{(n)}_0) \Rightarrow \rho$ as $n \to \infty$ for some probability law ρ on $\mathcal{M}(E)$. Suppose that the constants $\varepsilon^{(n)}_k$ fulfill (2.4.7). Assume that*

$$\begin{aligned}
&\text{(i)} & q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle &= 1 + \varepsilon\beta(x) + o(\varepsilon),\\
&\text{(ii)} & q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 &= \varepsilon\, 2\alpha(x) + o(\varepsilon),\\
&\text{(iii)} & q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}} &= o(\varepsilon)
\end{aligned} \tag{2.4.36}$$

for each $\delta > 0$, and

$$\int K_\varepsilon(x,\mathrm{d}y) f(y) = f(x) + \varepsilon A f(x) + o(\varepsilon) \tag{2.4.37}$$

for each $f \in \mathcal{D}(A)$, uniformly in x as $\varepsilon \to 0$. Then $\mathcal{L}(\mathcal{Y}^{(n)}) \Rightarrow \mathcal{L}(\mathcal{Y})$, where $\mathcal{Y}$ is the $(\overline{A},\alpha,\beta)$-superprocess with initial law ρ.

Here $\Rightarrow$ denotes weak convergence of probability measures on $\mathcal{D}_{\mathcal{M}(E)}[0,\infty)$.

Proof We apply Theorem 2.42 to the operator $\mathcal{G}$, where we use the fact that if we view $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)}[0,\infty))$ as a subspace of $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)_\infty}[0,\infty))$ (note the compactification), equipped with the topology of weak convergence, then the induced topology on $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)}[0,\infty))$ is again the topology of weak convergence.

By Proposition 2.44, solutions to the martingale problem for $\mathcal{G}$ are unique. Since $F_fF_g = F_{f+g}$ and $\mathcal{D}(A)$ is a linear space, the linear span of the domain of $\mathcal{G}$ is an algebra. Using the fact that $\mathcal{D}(A)$ is dense in $\mathcal{C}(E)$ we see that this algebra separates points. Therefore, we are left with the task to check (2.4.8).

Define $\mathcal{U}_\varepsilon : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$ by

$$\mathcal{U}_\varepsilon f(x) := q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\big(1 - e^{-\langle \chi, f\rangle}\big) \qquad (x \in E,\ f \in \mathcal{C}_+[0,1],\ f > 0,\ \varepsilon > 0), \quad (2.4.38)$$

and define transition probabilities $P_\varepsilon(\mu,\mathrm{d}\nu)$ on $\mathcal{M}(E)_\infty$ by $P_\varepsilon(\infty,\,\cdot\,) := \delta_\infty$ and

$$\int P_\varepsilon(\mu,\mathrm{d}\nu) e^{-\langle \nu, f\rangle} = e^{-\langle \mu, \mathcal{U}_\varepsilon f\rangle}. \quad (2.4.39)$$

We will show that

$$\lim_{\varepsilon\to 0} \big\| \varepsilon^{-1}(\mathcal{U}_\varepsilon f - f) - (Af + \beta f - \alpha f^2)\big\|_\infty = 0 \qquad (f \in \mathcal{D}(A),\ f > 0). \quad (2.4.40)$$

Together with (2.4.39) this implies that

$$\int P_\varepsilon(\mu,\mathrm{d}\nu)F_f(\nu) = F_f(\mu) + \varepsilon\mathcal{G}F_f(\mu) + o(\varepsilon) \qquad (f \in \mathcal{D}(A),\ f > 0), \quad (2.4.41)$$

uniformly in $\mu \in \mathcal{M}(E)_\infty$ as $\varepsilon \to 0$. Therefore, the result follows from Theorem 2.42.

It remains to prove (2.4.40). Set $g(z) := 1 - z + \frac{1}{2}z^2 - e^{-z}$ $(z \geq 0)$ and write

$$\mathcal{U}_\varepsilon f(x) = q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\big(\langle \chi, f\rangle - \tfrac{1}{2}\langle \chi, f\rangle^2 + g(\langle \chi, f\rangle)\big). \quad (2.4.42)$$

Since

$$g(z) = \int_0^z \mathrm{d}y \int_0^y \mathrm{d}x \int_0^x \mathrm{d}t\, e^{-t} \qquad (z \geq 0), \quad (2.4.43)$$

it is easy to see that g is nondecreasing on $[0,\infty)$ and (since $0 \leq e^{-t} \leq 1$ and $\int_0^x \mathrm{d}t\, e^{-t} \leq 1$)

$$0 \leq g(z) \leq \tfrac{1}{2}z^2 \wedge \tfrac{1}{6}z^3 \qquad (z \geq 0). \quad (2.4.44)$$

Using these facts and (2.4.36) (ii) and (iii), we find that

$$\begin{aligned}
&q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi) g(\langle \chi, f\rangle) \\
&\leq \|f\|_\infty q_\varepsilon(x) \Big\{ \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi) g(\langle \chi, 1\rangle) 1_{\{\langle \chi,1\rangle \leq \delta\}} + \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi) g(\langle \chi, 1\rangle) 1_{\{\langle \chi,1\rangle > \delta\}} \Big\} \\
&\quad \leq \|f\|_\infty q_\varepsilon(x) \Big\{ \tfrac{1}{6}\delta \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi) \langle \chi, 1\rangle^2 1_{\{\langle \chi,1\rangle \leq \delta\}} + \tfrac{1}{2} \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi) \langle \chi, 1\rangle^2 1_{\{\langle \chi,1\rangle > \delta\}} \Big\} \\
&\quad = \tfrac{1}{6}\delta \|f\|_\infty \big(\varepsilon\, 2\alpha(x) + o(\varepsilon)\big) + o(\varepsilon).
\end{aligned} \quad (2.4.45)$$

Since this holds for any $\delta > 0$, we conclude that

$$q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) g(\langle \chi, f \rangle) = o(\varepsilon) \tag{2.4.46}$$

uniformly in x as $\varepsilon \to 0$. By (2.4.36) (i) and (2.4.37),

$$\begin{aligned} q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f \rangle &= \Big(q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, 1 \rangle \Big) \Big(\int K_\varepsilon(x, \mathrm{d}y) f(y) \Big) \\ &= \big(1 + \varepsilon\beta(x) + o(\varepsilon)\big)\big(f(x) + \varepsilon A f(x) + o(\varepsilon)\big) \\ &= f(x) + \varepsilon\beta(x) f(x) + \varepsilon A f(x) + o(\varepsilon). \end{aligned} \tag{2.4.47}$$

Finally, write

$$\begin{aligned} & q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f \rangle^2 \\ & = q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \big(\langle \chi, f(x) \rangle^2 + 2\langle \chi, f(x) \rangle \langle \chi, f - f(x) \rangle + \langle \chi, f - f(x) \rangle^2 \big). \end{aligned} \tag{2.4.48}$$

Then, by (2.4.36) (ii),

$$q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f(x) \rangle^2 = f(x)^2 \big(\varepsilon\, 2\alpha(x) + o(\varepsilon) \big). \tag{2.4.49}$$

We will prove that

$$q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f - f(x) \rangle^2 = o(\varepsilon). \tag{2.4.50}$$

Then, by Hölder's inequality, (2.4.36) (ii), and (2.4.50),

$$\begin{aligned} & \Big| q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f - f(x) \rangle \langle \chi, f(x) \rangle \Big| \\ & \leq \Big(q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f - f(x) \rangle^2 \Big)^{1/2} \Big(q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f(x) \rangle^2 \Big)^{1/2} \\ & \leq \big(o(\varepsilon)(2\alpha(x)\varepsilon + o(\varepsilon)) \big)^{1/2} = o(\varepsilon). \end{aligned} \tag{2.4.51}$$

Inserting (2.4.49), (2.4.50) and (2.4.51) into (2.4.48) we find that

$$q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f \rangle^2 = \varepsilon\, 2\alpha(x) f(x)^2 + o(\varepsilon). \tag{2.4.52}$$

Inserting (2.4.46), (2.4.47) and (2.4.52) into (2.4.42), we arrive at (2.4.40). We still need to prove (2.4.50). To this aim, we estimate, using (2.4.47),

$$\begin{aligned} & q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f - f(x) \rangle^2 1_{\{\langle \chi, 1 \rangle \leq \delta\}} \\ & \leq \delta \| f - f(x) \|_\infty q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi) \langle \chi, f - f(x) \rangle \\ & = \delta \| f - f(x) \|_\infty \big(\varepsilon A f(x) + o(\varepsilon) \big) \end{aligned} \tag{2.4.53}$$

and, using (2.4.36) (iii),

$$\begin{aligned} q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi, f-f(x)\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}} \\ \leq \|f-f(x)\|_\infty q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}} = o(\varepsilon). \end{aligned} \tag{2.4.54}$$

It follows that

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi, f-f(x)\rangle^2 \leq \delta\varepsilon\|f-f(x)\|_\infty Af(x) + o(\varepsilon) \tag{2.4.55}$$

for any $\delta > 0$. This implies (2.4.50) and completes the proof of (2.4.40). ∎

2.4.3 Application to the renormalization branching process

Proof of Theorem 2.19 (a) For any $f_0,\ldots,f_k \in \mathcal{C}_+[0,1]$ one has

$$\begin{aligned} &E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k},f_k\rangle}\big] \\ &= E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k-1},f_{k-1}+\mathcal{U}_{\gamma_{n-k}}f_k\rangle}\big] \\ &= \cdots = \quad E\big[e^{-\langle\mathcal{X}_{-n},g_k\rangle}\big], \end{aligned} \tag{2.4.56}$$

where we define inductively

$$g_0 := f_k \quad\text{and}\quad g_{m+1} := f_{k-m-1} + \mathcal{U}_{\gamma_{n-k+m}} g_m. \tag{2.4.57}$$

By the compactness of $[0,1]$ and Corollary 2.36, the map $(\gamma,f)\mapsto\mathcal{U}_\gamma f$ from $(0,\infty)\times\mathcal{C}_+[0,1]$ to $\mathcal{C}_+[0,1]$ (equipped with the supremumnorm) is continuous. Using this fact and (2.4.56) we find that

$$E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k},f_k\rangle}\big] \underset{n\to\infty}{\longrightarrow} E\big[e^{-\langle\mathcal{Y}^{\gamma^*}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{Y}^{\gamma^*}_{-n+k},f_k\rangle}\big]. \tag{2.4.58}$$

Since $f_1,\ldots,f_k$ are arbitrary, (2.2.23) follows. ∎

Proof of Theorem 2.19 (b) We apply Theorem 2.45 to the weight functions q_γ and cluster mechanisms $\mathcal{Q}_\gamma$ from (2.2.19) and to $A_{\rm WF} = x(1-x)\frac{\partial^2}{\partial x^2}$ with domain $\mathcal{D}(A_{\rm WF}) = \mathcal{C}^{(2)}[0,1]$, and $\alpha=\beta=1$. It is well-known that $\overline{A}_{\rm WF}$ generates a Feller semigroup [EK86, Theorem 8.2.8]. We observe that

$$\int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle\chi,f\rangle = E\big[2\int_0^{\tau_\gamma} f(\mathbf{y}^\gamma_x(-t))\big] = 2E[\tau_\gamma]E\big[f(\mathbf{y}^\gamma_x(0))\big] = \gamma\int\Gamma^\gamma_x(\mathrm{d}y)f(y), \tag{2.4.59}$$

where Γ^γ_x is the equilibrium law of the process $\mathbf{y}^\gamma_x$ from Corollary 2.30. It follows from (2.3.24) that

$$\begin{aligned} &\text{(i)} && \int\Gamma^\gamma_x(\mathrm{d}y)(y-x) = 0, \\ &\text{(ii)} && \int\Gamma^\gamma_x(\mathrm{d}y)(y-x)^2 = \frac{\gamma x(1-x)}{1+\gamma}, \\ &\text{(iii)} && \int\Gamma^\gamma_x(\mathrm{d}y)(y-x)^4 = O(\gamma^2), \end{aligned} \tag{2.4.60}$$

uniformly in x as $\gamma \to 0$. Therefore, for any $\delta > 0$,

$$
\begin{aligned}
&\text{(i)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x) = 0,\\
&\text{(ii)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x)^2 = \gamma x(1-x) + o(\gamma),\\
&\text{(iii)} && \int \Gamma^\gamma_x(\mathrm{d}y) 1_{\{|y-x|>\delta\}} = o(\gamma),
\end{aligned}
\tag{2.4.61}
$$

uniformly in x as $\gamma \to 0$. Consequently, a Taylor expansion of f around x yields

$$
\int \Gamma^\gamma_x(\mathrm{d}y) f(x) = f(x) + \gamma \tfrac{1}{2} x(1-x)\tfrac{\partial^2}{\partial x^2} f(x) + o(\gamma) \qquad (f \in \mathcal{C}^{(2)}[0,1]),
\tag{2.4.62}
$$

uniformly in x as $\gamma \to 0$. (For details, in particular the uniformity in x, see for example [Swa99, Proposition B.1.1].) This shows that condition (2.4.37) is satisfied. Moreover,

$$
\begin{aligned}
\int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle &= E[2\tau_\gamma] = \gamma,\\
\int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^2 &= E[(2\tau_\gamma)^2] = \int_0^\infty z^2 \tfrac{1}{\gamma} e^{-z/\gamma}\mathrm{d}z = 2\gamma^2,\\
\int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^3 &= E[(2\tau_\gamma)^3] = \int_0^\infty z^3 \tfrac{1}{\gamma} e^{-z/\gamma}\mathrm{d}z = 6\gamma^3,
\end{aligned}
\tag{2.4.63}
$$

which, using the fact that $q_\gamma = (\frac{1}{\gamma} + 1)$, gives

$$
\begin{aligned}
q_\gamma \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle &= 1 + \gamma,\\
q_\gamma \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^2 &= 2\gamma + o(\gamma),\\
q_\gamma \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^3 &= o(\gamma).
\end{aligned}
\tag{2.4.64}
$$

This shows that (2.4.36) is fulfilled. In particular,

$$
q_\gamma \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^2 1_{\{\langle \chi,1\rangle > \delta\}} \leq \delta^{-1} q_\gamma \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)\langle \chi, 1\rangle^3 = o(\gamma)
\tag{2.4.65}
$$

for all $\delta > 0$. ∎

2.5 The super-Wright-Fisher diffusion: introduction

2.5.1 Superprocesses and binary splitting particle systems

Let E be a compact metrizable space, G the generator of a Feller process $\xi = (\xi_t)_{t\geq 0}$ in E, and $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$. Then, for each $f \in B_+(E)$, the semilinear Cauchy problem in $B_+(E)$

$$
\begin{cases}
\frac{\partial}{\partial t} u_t = G u_t + \beta u_t - \alpha u_t^2 \qquad (t \geq 0),\\
u_0 = f,
\end{cases}
\tag{2.5.1}
$$

has a unique mild solution $u_t =: \mathcal{U}_t f$. Moreover, there exists a unique (in law) Markov process $\mathcal{Y}$ with continuous sample paths in the space $\mathcal{M}(E)$ of finite measures on E, defined by its Laplace functionals

$$E^\mu[e^{-\langle \mathcal{Y}_t, f\rangle}] = e^{-\langle \mu, \mathcal{U}_t f\rangle} \qquad (t \geq 0,\ \mu \in \mathcal{M}(E),\ f \in B_+(E)). \tag{2.5.2}$$

The process $\mathcal{Y}$ is called the *superprocess* in E with *underlying motion generator* G, *activity* α and *growth parameter* β (the last two terms are our terminology), or in short the (G, α, β)-*superprocess*. The operators $(\mathcal{U}_t)_{t\geq 0} = \mathcal{U} = \mathcal{U}(G, \alpha, \beta)$ form a semigroup, called the *log-Laplace semigroup* of $\mathcal{Y}$.

The process $\mathcal{Y}$ can be constructed in several ways and is nowadays standard. We outlined one such construction in Section 2.4.2; see also, e.g., [Fit88, Fit91, Fit92]. We can think of $\mathcal{Y}$ as describing a population where mass flows with generator G, and during a time interval $\mathrm{d}t$ a bit of mass $\mathrm{d}m$ at position x produces offspring with mean $(1 + \beta(x)\mathrm{d}t)\mathrm{d}m$ and finite variance $2\alpha(x)\mathrm{d}t\,\mathrm{d}m$. For basic facts on superprocesses we refer to [Daw93, Eth00, Dyn02].

Similarly, when G is (again) the generator of a Feller process on a compact metrizable space E and $\alpha \in \mathcal{C}_+(E)$, then, for any $f \in B_{[0,1]}(E)$, the semilinear Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} u_t = G u_t + \alpha u_t(1 - u_t) & (t \geq 0), \\ u_0 = f, \end{cases} \tag{2.5.3}$$

has a unique mild solution $u_t =: U_t f$ in $B_{[0,1]}(E)$. Moreover, there exists a unique Markov process Y with cadlag sample paths in the space $\mathcal{N}(E)$ of finite counting measures on E, defined by its generating functionals

$$E^\nu\big[(1 - f)^{Y_t}\big] = (1 - U_t f)^\nu \qquad (t \geq 0,\ \nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)). \tag{2.5.4}$$

Here if $\nu = \sum_{i=1}^n \delta_{x_i}$ is a finite counting measure and $g \in B_{[0,1]}(E)$, then $g^\nu := \prod_{i=1}^n g(x_i)$. We call Y the *binary splitting particle system* in E with *underlying motion generator* G and *splitting rate* α, or in short the (G, α)-*bin-split-process*. The semigroup $(U_t)_{t\geq 0} = U = U(G, \alpha)$ is called the *generating semigroup* of Y. The process Y consists of particles that independently move according to the generator G, and additionally split with local rate α into two new particles, created at the position of the old one.

2.5.2 Statement of the problem and motivation

Let $\overline{A}$ be the closure in $\mathcal{C}[0,1]$ (equipped with the supremum norm) of the operator

$$A = \tfrac{1}{2} x(1 - x) \tfrac{\partial^2}{\partial x^2}. \tag{2.5.5}$$

It is well-known that $\overline{A}$ is the generator of a Feller process ξ on $[0,1]$, called the (standard) *Wright-Fisher diffusion*, see [EK86, Theorem 8.2.8]. We are interested in mild solutions to the Cauchy equation

$$\begin{cases} \frac{\partial}{\partial t} u_t = \overline{A} u_t + \alpha u_t(1 - u_t) & (t \geq 0), \\ u_0 = f, \end{cases} \tag{2.5.6}$$

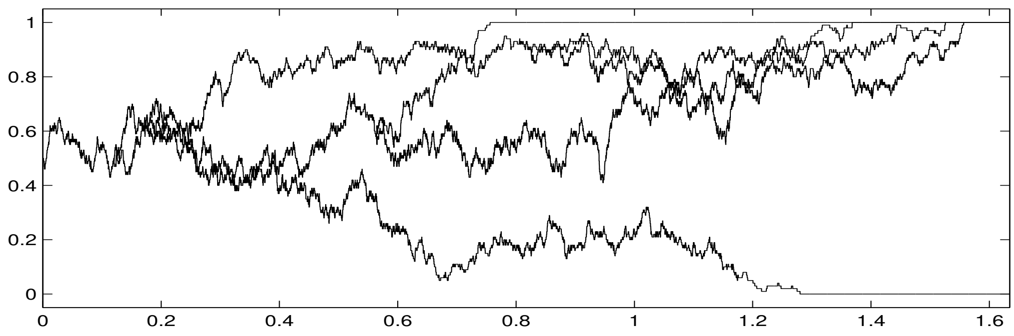


Figure 2.3: A system of binary splitting Wright-Fisher diffusions with splitting rate $\alpha = 1$.

where $\alpha > 0$ is a constant. We wish to find all fixed points of (2.5.6) and determine their domains of attraction.

For $f \in B_+[0,1]$, the mild solution of (2.5.6) is given by $u_t = \mathcal{U}_t f$, where $\mathcal{U} = \mathcal{U}(\overline{A}, \alpha, \alpha)$ is the log-Laplace semigroup of a superprocess $\mathcal{Y}$ in $[0,1]$ with underlying motion generator $G = \overline{A}$, and activity and growth parameter both equal to α. We call $\mathcal{Y}$ the *super-Wright-Fisher diffusion* (with activity and growth parameter $\alpha > 0$).[4]

Our main interest is in the case $\alpha = 1$. In this case, we have proved in Theorem 2.19 (b) above that a suitably rescaled version of the renormalization branching process converges to $\mathcal{Y}$. In particular, we will need Proposition 2.47 below for $\alpha = 1$ in our proof of Lemmas 2.24 and 2.25 (see Propositions 2.82 (b) and 2.83 (b) below). We will generalize a bit and treat general $\alpha > 0$. This will not be much more work and will give a more complete picture. In particular, we will see that the case $\alpha = 1$ is a critical case, since $\mathcal{Y}$ dies out on the interior if and only if $\alpha \leq 1$, and the weighted process $\mathcal{Y}^v$ from (2.5.19) is critical for $\alpha = 1$.

If $f \in B_{[0,1]}[0,1]$, then the solution of (2.5.6) is also given by $u_t = U_t f$, where $U = U(\overline{A}, \alpha)$ is the generating semigroup of a system Y of *binary splitting Wright-Fisher diffusions*, with splitting rate α. The process Y can be obtained from $\mathcal{Y}$ by Poissonization with the constant function 1 (compare Proposition 2.21). In fact, Y is the *trimmed tree* of $\mathcal{Y}$, i.e., the particles in Y correspond to those infinitesimal bits of mass in $\mathcal{Y}$, that have offspring at all later times. For a precise statement of this fact we refer the reader to [FS04].

See Figure 2.3 for a simulation of Y for $\alpha = 1$. The points $0, 1$ are accessible traps for the Wright-Fisher diffusion, and therefore a natural question is whether eventually all particles of Y end up in 0 or 1. This question will be answered for all $\alpha > 0$ in Proposition 2.48 below.

Binary splitting Wright-Fisher diffusions have been studied before in [GKW01]. In particular, the authors of that paper investigated the function p, which is defined in terms of the system Y of binary splitting Wright-Fisher diffusions with splitting rate $\alpha = 1$, as

$$p(x) := \lim_{t\to\infty} P^{\delta_x}[Y_t(\{1\}) > 0] = \lim_{t\to\infty} P^{\delta_x}[Y_t((0,1]) > 0] \qquad (x \in [0,1]). \tag{2.5.7}$$

In order to show that the two expressions for p in (2.5.7) are identical, in [GKW01] the authors

[4]More generally, if $\mathcal{Z}$ is the $(\overline{A}, \alpha', \alpha)$-superprocess, with $\alpha', \alpha > 0$ constants, then $\frac{\alpha}{\alpha'}\mathcal{Z} = \mathcal{Y}$ in law, and therefore this more general case can be reduced to the case $\alpha' = \alpha$.

note that both expressions correspond to a fixed point p of the generating semigroup $U(\overline{A},1)$ with boundary conditions $p(0)=0$ and $p(1)=1$. Assuming that p is sufficiently smooth, the fixed point property means that p solves the equation

$$\tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p(x)+\alpha p(x)(1-p(x))=0 \qquad (x\in[0,1]). \tag{2.5.8}$$

Though stated only for the case $\alpha=1$, the proof of Lemma 1.13 in [GKW01] shows that equation (2.5.8) has at most one solution with boundary conditions $p(0)=0$ and $p(1)=1$ when $\alpha<z_0^2/8\cong 1.836$, where z_0 is the smallest non-trivial zero of the Bessel function of the first kind with parameter 1. The authors do not answer the question whether solutions to (2.5.8) with these boundary condions are unique for $\alpha\geq z_0^2/8$, or what solutions may exist for other boundary conditions. Proposition 2.47 below settles these questions. We show moreover that all fixed points of $U(\overline{A},\alpha)$ are smooth, a fact tacitly assumed in [GKW01].

2.5.3 Results

The following theorem is our main result. We write 'eventually' behind an event, depending on t, to denote the existence of a (random) time $\tau<\infty$ such that the event holds for all $t\geq\tau$.

Theorem 2.46 (Long-time behavior of the super-Wright-Fisher diffusion) *Let $\mathcal{Y}$ be the super-Wright-Fisher diffusion with activity and growth parameter equal to the same constant $\alpha>0$, started in $\mu\in\mathcal{M}[0,1]$. Set*

$$v(x):=6x(1-x) \qquad (x\in[0,1]). \tag{2.5.9}$$

Then there exist nonnegative random variables W_0, W_1, $W_{(0,1)}$ (depending on μ) such that

$$\begin{array}{lll}\text{(i)} & \lim_{t\to\infty} e^{-\alpha t}\langle\mathcal{Y}_t,1_{\{r\}}\rangle=W_r \quad \text{a.s.} & (r=0,1),\\ \text{(ii)} & \lim_{t\to\infty} e^{-(\alpha-1)t}\langle\mathcal{Y}_t,v\rangle=W_{(0,1)} \quad \text{a.s.} & \end{array} \tag{2.5.10}$$

and

$$\begin{array}{lll}\text{(i)} & \{W_r=0\}=\{\mathcal{Y}_t(\{r\})=0 \textit{ eventually}\} \quad \text{a.s.} & (r=0,1),\\ \text{(ii)} & \{W_{(0,1)}=0\}=\{\mathcal{Y}_t((0,1))=0 \textit{ eventually}\} \quad \text{a.s.} & \end{array} \tag{2.5.11}$$

Moreover,

$$\{W_{(0,1)}>0\}\subset\{W_0>0\}\cap\{W_1>0\} \quad \text{a.s.} \tag{2.5.12}$$

If $\alpha\leq 1$, then

$$W_{(0,1)}=0 \qquad \text{a.s.} \tag{2.5.13}$$

If $\alpha>1$, then $W_{(0,1)}$ satisfies

$$E^\mu(W_{(0,1)})=\langle\mu,v\rangle \quad \textit{and} \quad \mathrm{Var}^\mu(W_{(0,1)})\leq 3\tfrac{\alpha}{\alpha-1}\langle\mu,v\rangle \tag{2.5.14}$$

as well as

$$\lim_{t\to\infty}E^\mu\Big[\big|e^{-(\alpha-1)t}\langle\mathcal{Y}_t,vf\rangle-W_{(0,1)}\langle\ell,vf\rangle\big|^2\Big]=0 \quad \forall f\in B[0,1], \tag{2.5.15}$$

where ℓ denotes the Lebesgue measure on $(0,1)$.

Except for the statement about smoothness (of the functions $p_{0,0},\ldots,p_{1,1}$ below) and the uniformity of the limit in (2.5.16), the following result about the log-Laplace semigroup $\mathcal{U}(\overline{A},\alpha,\alpha)$ is an immediate consequence of Theorem 2.46.

Proposition 2.47 (Long-time behavior of $\mathcal{U}(\overline{A},\alpha,\alpha)$) *Let $\mathcal{Y}$, $W_0,W_1,W_{(0,1)}$ be as in Theorem 2.46 and let $\mathcal{U}=\mathcal{U}(\overline{A},\alpha,\alpha)$. Then, for all $f\in B_+[0,1]$, uniformly on $[0,1]$,*

$$\lim_{t\to\infty}\mathcal{U}_t f=\begin{cases}0 & \text{if } \ f(0)=f(1)=\langle\ell,f\rangle=0,\\ p_{0,0} & \text{if } \ f(0)=f(1)=0,\ \langle\ell,f\rangle>0,\\ p_{1,0} & \text{if } \ f(0)>0,\ f(1)=0,\\ p_{0,1} & \text{if } \ f(0)=0,\ f(1)>0,\\ p_{1,1} & \text{if } \ f(0)>0,\ f(1)>0,\end{cases}\tag{2.5.16}$$

where the constant function 0 and

$$\left.\begin{array}{l} p_{0,0}(x):=-\log P^{\delta_x}[W_{(0,1)}=0],\\ p_{1,0}(x):=-\log P^{\delta_x}[W_0=0]=P^{\delta_x}[W_0=W_{(0,1)}=0],\\ p_{0,1}(x):=-\log P^{\delta_x}[W_1=0]=P^{\delta_x}[W_1=W_{(0,1)}=0],\\ p_{1,1}(x):=-\log P^{\delta_x}[W_0=W_1=0]=P^{\delta_x}[W_0=W_1=W_{(0,1)}=0]\end{array}\right\}\quad(x\in[0,1]).\tag{2.5.17}$$

are all fixed points of the log-Laplace semigroup $\mathcal{U}(\overline{A},\alpha,\alpha)$. Here $p_{0,0}=0$ if $\alpha\leq 1$, and $p_{0,0}>0$ on $(0,1)$ if $\alpha>1$. The functions $p_{l,r}$ ($l,r\in\{0,1\}$ satisfy $p_{l,r}(0)=l$ and $p_{l,r}(1)=r$, are twice continuously differentiable on $[0,1]$, and solve (2.5.8).

Since conversely, every nonnegative twice continuously differentiable solution to (2.5.8) is a fixed point of $\mathcal{U}(\overline{A},\alpha,\alpha)$, we see that (2.5.8) has precisely four solutions when $\alpha\leq 1$ and precisely five solutions when $\alpha>1$. The functions $p_{0,0},\ldots,p_{1,1}$ are $[0,1]$-valued and therefore fixed points of the generating semigroup $U(\overline{A},\alpha)$ as well. Our final result describes $p_{0,0},\ldots,p_{1,1}$ in terms of the system Y of binary splitting Wright-Fisher diffusions with splitting rate α.

Proposition 2.48 (Fixed points of $U(\overline{A},\alpha)$) *The functions $p_{0,0},\ldots,p_{1,1}$ in (2.5.17) satisfy*

$$\left.\begin{array}{l} p_{0,0}(x)=P^{\delta_x}[Y_t((0,1))>0 \textit{ eventually}],\\ p_{1,0}(x)=P^{\delta_x}[Y_t(\{0\})>0 \textit{ eventually}]=P^{\delta_x}[Y_t([0,1))>0 \textit{ eventually}],\\ p_{0,1}(x)=P^{\delta_x}[Y_t(\{1\})>0 \textit{ eventually}]=P^{\delta_x}[Y_t((0,1])>0 \textit{ eventually}],\\ p_{1,1}(x)=1\end{array}\right\}\quad(x\in[0,1]).\tag{2.5.18}$$

See Figure 2.4 for a plot of the functions $p_{0,0}$ and $p_{0,1}$ (for $\alpha=2$).

2.5.4 Methods and related work

An essential tool in the proof of Theorem 2.46 is the *weighted super-Wright-Fisher diffusion* $\mathcal{Y}^v$, defined as

$$\mathcal{Y}^v_t(\mathrm{d}x):=v(x)\mathcal{Y}_t(\mathrm{d}x)\qquad(t\geq 0),\tag{2.5.19}$$

where v is defined in (2.5.9). Note that v is an eigenfunction of the operator $\overline{A}$, with eigenvalue -1. For convenience, we have normalized v such that $\langle\ell,v\rangle=1$.

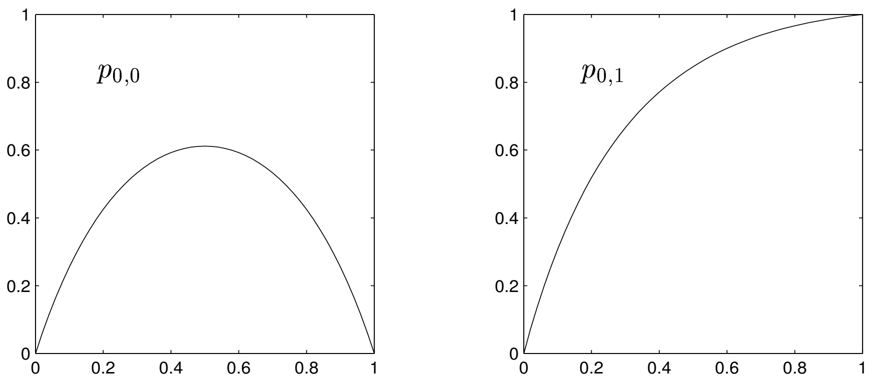


Figure 2.4: Two solutions to the differential equation $\frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p(x) + 2p(x)(1-p(x)) = 0$.

When a superprocess is weighted with a sufficiently smooth density, the result is a new superprocess, with a new activity and growth parameter and a new underlying motion, which is a compensated h-transform of the old one. For the case that the underlying motion is a locally uniformly elliptic diffusion on a open domain $D \subset \mathbb{R}^d$, weighted superprocesses were developed by [EP99]. In our case, where uniform ellepticity does not hold, the following can be proved without too much effort.

Lemma 2.49 (Weighted super-Wright-Fisher diffusion) *Let $\mathcal{Y}$ be the super-Wright-Fisher diffusion with $\alpha > 0$ and let $\mathcal{Y}^v$ be defined as in (2.5.19). Then $\mathcal{Y}^v$ is the $(\overline{A^v}, \alpha v, \alpha - 1)$-superprocess in $[0,1]$, where $\overline{A^v}$ is the closure of the operator*

$$A^v := \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2} + 2(\tfrac{1}{2} - x)\tfrac{\partial}{\partial x}. \tag{2.5.20}$$

Indeed, $\overline{A^v}$ generates a Feller process ξ^v in $[0,1]$, see [EK86, Theorem 8.2.1]. The diffusion ξ^v is a compensated h-transform (with $h = v$) of the Wright-Fisher diffusion ξ. This compensated v-transformed Wright-Fisher diffusion ξ^v is ergodic with invariant law $v\ell$ (Lemma 2.65 below). For $\alpha > 1$, the $(\overline{A^v}, \alpha v, \alpha - 1)$-superprocess is supercritical, and in this case one expects $e^{-(\alpha-1)t}\mathcal{Y}^v_t$ to converge, in some way, to a random multiple of $v\ell$. This is the idea behind formula (2.5.15).

Recently, [ET02], have shown for a certain class of superdiffusions $\mathcal{Y}$ in $\mathbb{R}^d$ with underlying motion generator G, growth parameter β and activity α, the convergence in law

$$e^{-\lambda_c t}\langle \mathcal{Y}, g\rangle \Rightarrow W\langle \rho, g\rangle \quad \text{as } t \to \infty, \tag{2.5.21}$$

where W is a nonnegative random variable, λ_c is the generalized principal eigenvalue of $G + \beta$ (which is assumed to be positive), ρ is a measure on $\mathbb{R}^d$, defined in terms of $G + \beta$, and g is any compactly supported continuous function on $\mathbb{R}^d$. In their work, the weighted superprocess $\mathcal{Y}^\phi_t(\mathrm{d}x) := \phi(x)\mathcal{Y}_t(\mathrm{d}x)$ plays a central role, where ϕ is the principal eigenfunction of the operator $G + \beta$. Their dynamical system methods are based on a result on the existence of an invariant curve of the log-Laplace semigroup of their superprocess. Using this invariant curve, they give an expression for the Laplace-transform of the law of the random variable W in (2.5.21). Their results are in line with our results for the super-Wright-Fisher diffusion restricted to $(0,1)$, where in our case $\lambda_c = \alpha - 1$ and $\phi = v$. However, their methods use in an essential way the fact that their underlying space is $\mathbb{R}^d$ (and not an open subset of $\mathbb{R}^d$, like $(0,1)$), and therefore their results are not applicable to our situation. It is stated as an open problem by [ET02] whether the random variable W in (2.5.21) in general satisfies $P[W=0] = P[\mathcal{Y}_t = 0 \text{ eventually}]$. For a recent result on *local* extinction versus *local* exponential growth of superdiffusions on open domains $D \subset \mathbb{R}^d$, we refer to [EK04].

In our set-up, we can prove that $\{W_{(0,1)} = 0\} = \{\mathcal{Y}_t((0,1)) = 0 \text{ eventually}\}$ because of the following property of the weighted super-Wright-Fisher diffusion $\mathcal{Y}^v$.

Lemma 2.50 (Finite ancestry) *For all $\alpha > 0$, the weighted super-Wright-Fisher diffusion $\mathcal{Y}^v$ satisfies*

$$\inf_{x\in[0,1]} P^{\delta_x}[\mathcal{Y}^v_t = 0] > 0 \qquad \forall t > 0. \tag{2.5.22}$$

Formula (2.5.22) has been called the *finite ancestry property* (of $\mathcal{Y}^v$); for a justification of this terminology we refer the reader to [FS04]. A sufficient condition for a superprocess to enjoy the finite ancestry property is that the activity be bounded away from zero (see Lemma 2.55 below). This condition is not necessary. In fact, the activity of $\mathcal{Y}^v$ is αv, which is zero on $\{0,1\}$. Our proof of Lemma 2.50 is quite long. It is not clear whether the weighted superprocesses $\mathcal{Y}^\phi$ occurring in [ET02] will in general satisfy a formula of the form (2.5.22). Therefore, we mention as an open problem:

> How to check, in a practical way, whether a given superprocess has the finite ancestry property (2.5.22)?

Another problem that is left open in here, is whether the L_2-convergence in (2.5.15) can be replaced by almost sure convergence. In fact, we suspect that (2.5.15) can be strengthened to

$$\lim_{t\to\infty} e^{-(\alpha-1)t}\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle = W_{(0,1)}\langle \ell, f\rangle \quad \forall f \in B[0,1] \quad \text{a.s.}, \tag{2.5.23}$$

but we do not have a proof.

The following sections are organized as follows. Sections 2.6.1 and 2.6.2 contain some general facts about (G,α,β)-superprocesses and on (G,α,β)-superprocesses enjoying the finite ancestry property, respectively. After some preparatory work in Sections 2.6.3 and 2.6.4, we prove Lemmas 2.49 and 2.50 in Section 2.6.5. In Sections 2.7.1 and 2.7.2 we derive some properties of the weighted super-Wright-Fisher diffusion $\mathcal{Y}^v$, culminating in the proof of Theorem 2.46 in Section 2.7.3. Finally, Sections 2.7.4–2.7.5 contain the proofs of Propositions 2.47 and 2.48.

2.6 The super-Wright-Fisher diffusion: preparatory results

2.6.1 Some general facts about log-Laplace semigroups

Let E be a compact metrizable space and let $\mathcal{C}(E)$ be the space of continuous real functions on E, equipped with the supremum norm $\|\cdot\|_\infty$. Let $\xi = (\xi_t)_{t\geq 0}$ be a Feller process in E with semigroup $S_t f(x) := E^x[f(\xi_t)]$ $(t \geq 0,\ x \in E, f \in B(E))$. By definition, the (full) generator G of ξ is the linear operator on $\mathcal{C}(E)$ given by $Gf := \lim_{t\to 0} t^{-1}(S_t f - f)$ where the domain $\mathcal{D}(G)$ of G is the space of all functions $f \in \mathcal{C}(E)$ for which the limit exists in $\mathcal{C}(E)$.

Let $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$, and $f \in \mathcal{C}_+(E)$. By definition, we call u a *classical solution* of the Cauchy problem (2.5.1) if $u : [0,\infty) \to \mathcal{C}_+(E) \cap \mathcal{D}(G)$ is continuously differentiable in $\mathcal{C}(E)$ (i.e., the derivative $\frac{\partial}{\partial t}u_t := \lim_{s\to t} s^{-1}(u_{t+s} - u_t)$ exists in $\mathcal{C}(E)$ for all $t \geq 0$ and the map $\frac{\partial}{\partial t}u : [0,\infty) \to \mathcal{C}(E)$ is continuous) and (2.5.1) holds. A measurable function $u : [0,\infty) \times E \to [0,\infty)$ is called a *mild solution* of (2.5.1) if u is bounded on finite time intervals and solves (pointwise)

$$u_t = S_t f + \int_0^t S_{t-s}\big(\beta u_s - \alpha u_s^2\big)\mathrm{d}s \qquad (t \geq 0). \tag{2.6.1}$$

Equation (2.5.1) has a unique mild solution for all $f \in B_+(E)$, see [Fit88] and this solution is a classical solution if $f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)$. (See [Paz83], Theorems 6.1.4 and 6.1.5. The fact that f is nonnegative and $\alpha \geq 0$ implies that solutions cannot explode. Our definition of a classical solution is slightly stronger than the one used in [Paz83], since we require u to be continuously differentiable on $[0,\infty)$ instead of $(0,\infty)$. However, the proof of Theorem 6.1.5 in [Paz83] shows that u is continuously differentiable on $[0,\infty)$ if $f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)$.)

The (G,α,β)-superprocess $\mathcal{Y}$ is defined as the unique strong Markov process with continuous sample paths in $\mathcal{M}(E)$, equipped with the topology of weak convergence, such that (2.5.2) holds for all $f \in B_+(E)$; see [Fit88, Fit91, Fit92].

Note the following elementary properties of the log-Laplace semigroup $\mathcal{U}(G,\alpha,\beta)$. Here, we write bp-$\lim_{n\to\infty} f_n = f$ if f is the bounded pointwise limit of the sequence $(f_n)_{n\geq 0}$.

Lemma 2.51 (Continuity and monotonicity of log-Laplace semigroups) *For each $t \geq 0$, $\mathcal{U}_t : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$ is continuous. Moreover, if* bp-$\lim_{n\to\infty} f_n = f$ *for some sequence $f_n \in B_+(E)$, then* bp-$\lim_{n\to\infty} \mathcal{U}_t f_n = \mathcal{U}_t f$. *Finally, $f \leq g$ implies $\mathcal{U}_t f \leq \mathcal{U}_t g$ $(f,g \in B_+(E))$.*

Proof The continuity of $\mathcal{U}_t : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$ follows from [Paz83, Theorem 6.1.2] and the fact that solutions do not explode. Continuity of $\mathcal{U}_t$ with respect to bounded pointwise limits is obvious from (2.5.2), and the same formula also makes clear that $\mathcal{U}_t : B_+(E) \to B_+(E)$ is monotone. ∎

Recall that (2.5.1) has a classical solution for $f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)$. Because of the following, for many purposes it suffices to work with classical solutions.

Lemma 2.52 (Closure and bp-closure) *For $t \geq 0$ fixed, $\{(f,\mathcal{U}_t f) : f \in \mathcal{C}_+(E)\}$ is the closure in $\mathcal{C}(E)$ of $\{(f,\mathcal{U}_t f) : f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)\}$, and $\{(f,\mathcal{U}_t f) : f \in B_+(E)\}$ is the bp-closure of $\{(f,\mathcal{U}_t f) : f \in \mathcal{C}_+(E)\}$.*

Here, the bp-closure of a set B is the smallest set $\overline{B}$ such that $B \subset \overline{B}$ and $f \in \overline{B}$ whenever bp-$\lim_{n\to\infty} f_n = f$ for some sequence $f_n \in \overline{B}$.

Proof of Lemma 2.52 It follows from the Hille-Yosida Theorem, see [EK86, Theorem 1.2.6] that $\mathcal{D}(G)$ is dense in $\mathcal{C}(E)$. Since $\mathcal{D}(G)$ is a linear space and $1 \in \mathcal{D}(G)$, it is not hard to see that $\mathcal{C}_+(E) \cap \mathcal{D}(G)$ is dense in $\mathcal{C}_+(E)$. The fact that $\{(f,\mathcal{U}_t f) : f \in \mathcal{C}_+(E)\}$ is the closure in $\mathcal{C}(E)$ of $\{(f,\mathcal{U}_t f) : f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)\}$ now follows from the continuity of $\mathcal{U}_t : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$.

In [EK86, Proposition 3.4.2], it is proved that $\mathcal{C}(E)$ is bp-dense in $B(E)$; the argument can easily be adapted to show that $\mathcal{C}_+(E)$ is bp-dense in $B_+(E)$. Therefore Lemma 2.52 follows from the continuity of $\mathcal{U}_t$ with respect to bounded pointwise limits. ∎

$\mathcal{U}_t f$ may be defined unambiguously such that (2.5.2) holds also for functions f that are not bounded, or even infinite.

Lemma 2.53 (Extension of $\mathcal{U}$ to unbounded functions) *For each measurable $f : E \to [0,\infty]$ and $t \geq 0$ there exists a unique measurable $\mathcal{U}_t f : E \to [0,\infty]$ such that (2.5.2) holds for all $\mu \in \mathcal{M}(E)$, where we put $e^{-\infty} := 0$.*

Proof Define $\mathcal{U}_t f$ by $\mathcal{U}_t f(x) := -\log E^{\delta_x}[e^{-\langle \mathcal{Y}_t, f\rangle}]$ where $\log 0 := -\infty$. To see that (2.5.2) holds again for all $\mu \in \mathcal{M}(E)$, choose $B_+(E) \ni f_n \uparrow f$, note that $\mathcal{U}_t f_n \uparrow \mathcal{U}_t f$, and take the limit in (2.5.2). ∎

We will often need the following comparison result, compare [Smo83, Theorem 10.1].

Lemma 2.54 (Sub- and supersolutions) *Assume that $T > 0$ and that $\tilde{u} : [0,T] \to \mathcal{C}_+(E) \cap \mathcal{D}(G)$ is continuously differentiable in $\mathcal{C}(E)$ and solves*

$$\tfrac{\partial}{\partial t}\tilde{u}_t \leq G\tilde{u}_t + \beta\tilde{u}_t - \alpha\tilde{u}_t^2 \qquad (t \in [0,T]). \tag{2.6.2}$$

Then $\tilde{u}_T \leq \mathcal{U}_T \tilde{u}_0$. The same holds with both inequality signs reversed.

Proof Let $g : [0,T] \to \mathcal{C}_+(E)$ be defined by the formula

$$\tfrac{\partial}{\partial t}\tilde{u}_t = G\tilde{u}_t + \beta\tilde{u}_t - \alpha\tilde{u}_t^2 - g_t \qquad (t \in [0,T]). \tag{2.6.3}$$

Set $u_t := \mathcal{U}_t \tilde{u}_0$. Then $u : [0,T] \to \mathcal{C}_+(E)$ is the classical solution of

$$\begin{cases} \tfrac{\partial}{\partial t}u_t = Gu_t + \beta u_t - \alpha u_t^2 & (t \in [0,T]), \\ u_0 = \tilde{u}_0. \end{cases} \tag{2.6.4}$$

Put $\Delta_t := u_t - \tilde{u}_t$ $(t \in [0,T])$. Then Δ solves

$$\begin{cases} \tfrac{\partial}{\partial t}\Delta_t = G\Delta_t + \beta\Delta_t - \alpha\,(u_t + \tilde{u}_t)\Delta_t + g_t & (t \in [0,T]), \\ \Delta_0 = 0. \end{cases} \tag{2.6.5}$$

The generator G satisfies the positive maximum principle, see [EK86, Theorem 4.2.2] and therefore (2.6.5) implies that $\Delta \geq 0$. For imagine that $\Delta_t(x) < 0$ somewhere on $[0,T] \times E$. Let R be a constant such that $\beta - \alpha\,(u_t + \tilde{u}_t) + R < 0$. Then $\tilde{\Delta}_t := e^{Rt}\Delta_t$ solves

$$\begin{cases} \tfrac{\partial}{\partial t}\tilde{\Delta}_t = G\tilde{\Delta}_t + \{\beta - \alpha\,(u_t + \tilde{u}_t) + R\}\tilde{\Delta}_t + g_t e^{Rt} & (t \in [0,T]), \\ \tilde{\Delta}_0 = 0. \end{cases} \tag{2.6.6}$$

If $\tilde{\Delta}_t(x) < 0$ for some $(t,x) \in [0,T] \times E$, then $\tilde{\Delta}$ must assume a negative minimum over $[0,T] \times E$ in some point (s,y), with $s > 0$ since $\tilde{\Delta}_0 = 0$. But in such a point one would have $\frac{\partial}{\partial s}\tilde{\Delta}_s(y) \leq 0$ while $G\tilde{\Delta}_s(y) + \{\beta(y) - \alpha(y)\,(u_s(y) + \tilde{u}_s(y)) + R\}\tilde{\Delta}_s(y) + g_s(y)e^{Rs} > 0$, in contradiction with (2.6.6).

The same argument applies when both inequality signs are reversed. ∎

Lemma 2.54 has the following application.

Lemma 2.55 (Bounds on log-Laplace semigroups) *Let* $\mathcal{U} = \mathcal{U}(G, \alpha, \beta)$, $\overline{\mathcal{U}} = \mathcal{U}(G, \underline{\alpha}, \overline{\beta})$, *where* $\alpha, \underline{\alpha} \in \mathcal{C}_+(E)$ *and* $\beta, \overline{\beta} \in \mathcal{C}(E)$ *satisfy*

$$\alpha \geq \underline{\alpha} \quad \text{and} \quad \beta \leq \overline{\beta}. \tag{2.6.7}$$

Then

$$\mathcal{U}_t f \leq \overline{\mathcal{U}}_t f \text{ for all measurable } f : E \to [0, \infty] \quad (t \geq 0). \tag{2.6.8}$$

In particular, if $\underline{\alpha}, \overline{\beta}$ *are constants and* $\underline{\alpha} > 0$, *then, for* $t > 0$,

$$\overline{\mathcal{U}}_t \infty = \frac{\overline{\beta}}{\underline{\alpha}\,(1 - e^{-\overline{\beta} t})} \quad (\overline{\beta} \neq 0) \quad \text{and} \quad \overline{\mathcal{U}}_t \infty = \frac{1}{\underline{\alpha}\, t} \quad (\overline{\beta} = 0), \tag{2.6.9}$$

and (2.6.8) with $f = \infty$ *gives*

$$P^\mu[\mathcal{Y}_t = 0] \geq e^{-\langle \mu, \overline{\mathcal{U}}_t \infty \rangle} \qquad (t > 0). \tag{2.6.10}$$

Proof For each $f \in \mathcal{C}_+(E) \cap \mathcal{D}(G)$, the function $\tilde{u}_t := \mathcal{U}_t f$ solves

$$\tfrac{\partial}{\partial t} \tilde{u}_t = G \tilde{u}_t + \beta \tilde{u}_t - \alpha \tilde{u}_t^2 \leq G \tilde{u}_t + \overline{\beta} \tilde{u}_t - \underline{\alpha} \tilde{u}_t^2 \qquad (t \geq 0), \tag{2.6.11}$$

and therefore $\mathcal{U}_t f = \tilde{u}_t \leq \overline{\mathcal{U}}_t f$ by Lemma 2.54. Using Lemmas 2.52 and 2.53 this is easily extended to measurable $f : E \to [0, \infty]$, giving (2.6.8). Define $\overline{u}$ by the right-hand side of the equations in (2.6.9). Then it is easy to check that $\overline{u}$ solves $\frac{\partial}{\partial t} \overline{u}_t = \overline{\beta}\, \overline{u}_t - \underline{\alpha}\, \overline{u}_t^2$ $(t > 0)$ with $\lim_{t \to 0} \overline{u}_t = \infty$, and therefore (2.6.10) follows from the fact that

$$P^\mu[\mathcal{Y}_t = 0] = E^\mu[e^{-\langle \mathcal{Y}_t, \infty \rangle}] = e^{-\langle \mu, \mathcal{U}_t \infty \rangle} \qquad (t \geq 0,\ \mu \in \mathcal{M}(E)), \tag{2.6.12}$$

and a little approximation argument. ∎

2.6.2 Some consequences of the finite ancestry property

Let $\mathcal{Y}$ be a (G, α, β)-superprocess as in the last section. In line with Lemma 2.50, we say that $\mathcal{Y}$ has the *finite ancestry property* if

$$\inf_{x \in E} P^{\delta_x}[\mathcal{Y}_t = 0] > 0 \qquad (t > 0). \tag{2.6.13}$$

Note that by (2.6.12), property (2.6.13) is equivalent to $\|\mathcal{U}_t \infty\|_\infty < \infty$ $(t > 0)$. In this section we prove three simple consequences of the finite ancestry property.

Lemma 2.56 (Extinction versus unbounded growth) *Assume that the* (G, α, β)*-superprocess* $\mathcal{Y}$ *has the finite ancestry property. Then, for any* $\mu \in \mathcal{M}(E)$,

$$P^\mu\big[\mathcal{Y}_t = 0 \text{ eventually or } \lim_{t \to \infty} \langle \mathcal{Y}_t, 1 \rangle = \infty\big] = 1. \tag{2.6.14}$$

Proof We use a general fact about tail events of strong Markov processes, the statement and proof of which can be found in Section 2.6.6. Consider the tail event $A := \{\mathcal{Y}_t = 0 \text{ eventually}\}$. By Lemma 2.64 below,

$$\lim_{t\to\infty} P^{\mathcal{Y}_t}(A) = 1_A \qquad \text{a.s.} \tag{2.6.15}$$

For any fixed $T > 0$, by (2.6.12),

$$P^\mu(A) \geq P^\mu[\mathcal{Y}_T = 0] = e^{-\langle \mu, \mathcal{U}_T \infty\rangle} \geq e^{-\langle \mu, 1\rangle \|\mathcal{U}_T\infty\|_\infty} \qquad (\mu \in \mathcal{M}(E)). \tag{2.6.16}$$

Hence (2.6.15) implies that

$$\liminf_{t\to\infty} e^{-\langle \mathcal{Y}_t, 1\rangle \|\mathcal{U}_T\infty\|_\infty} \leq 1_A \qquad \text{a.s.} \tag{2.6.17}$$

By the finite ancestry property, $\|\mathcal{U}_T\infty\|_\infty < \infty$ and therefore $\lim_{t\to\infty}\langle \mathcal{Y}_t, 1\rangle = \infty$ a.s. on $A^{\rm c}$. ∎

The following is a simple consequence of Lemma 2.56.

Lemma 2.57 (Extinction of (sub-) critical processes) *Assume that the (G, α, β)-superprocess $\mathcal{Y}$ has the finite ancestry property and that $\beta \leq 0$. Then, for any $\mu \in \mathcal{M}(E)$,*

$$P^\mu\big[\mathcal{Y}_t = 0 \text{ eventually}\big] = 1. \tag{2.6.18}$$

Proof Since $E^\mu[\langle \mathcal{Y}_t, 1\rangle] \leq \langle \mu, 1\rangle$, $P^\mu[\lim_{t\to\infty}\langle \mathcal{Y}_t, 1\rangle = \infty] = 0$. Now the claim follows from Lemma 2.56. ∎

Our final result of this section is the following.

Lemma 2.58 (Extinction versus exponential growth) *Assume that the (G, α, β)-superprocess $\mathcal{Y}$ has the finite ancestry property and that $\beta > 0$ is a constant. Then, for any $\mu \in \mathcal{M}(E)$, there exists a nonnegative random variable W, depending on μ, such that*

$$\begin{array}{rl}
\text{(i)} & \lim_{t\to\infty} e^{-\beta t}\langle \mathcal{Y}_t, 1\rangle = W \qquad P^\mu\text{-a.s.},\\
\text{(ii)} & \lim_{t\to\infty} E^\mu\big[|e^{-\beta t}\langle \mathcal{Y}_t, 1\rangle - W|^2\big] = 0,\\
\text{(iii)} & E^\mu(W) = \langle \mu, 1\rangle,\\
\text{(iv)} & \mathrm{Var}^\mu(W) \leq 2\beta^{-1}\|\alpha\|_\infty \langle \mu, 1\rangle,\\
\text{(v)} & \{W = 0\} = \{\mathcal{Y}_t = 0 \text{ eventually}\} \qquad P^\mu\text{-a.s.}
\end{array} \tag{2.6.19}$$

Proof Put $\mathcal{V}_t f := e^{\beta t} S_t$. The mean and covariance of $\mathcal{Y}$ are given by the following formulas, see, for example, [Fit88]:

$$\left.\begin{array}{rl}
\text{(i)} & E^\mu[\langle \mathcal{Y}_t, f\rangle] = \langle \mu, \mathcal{V}_t f\rangle\\
\text{(ii)} & \mathrm{Cov}^\mu(\langle \mathcal{Y}_t, f\rangle, \langle \mathcal{Y}_t, g\rangle) = 2\int_0^t \mathrm{d}s\, \langle \mu, \mathcal{V}_s(\alpha\, (\mathcal{V}_{t-s}f)(\mathcal{V}_{t-s}g))\rangle
\end{array}\right\} \quad (t \geq 0,\ f, g \in B(E)). \tag{2.6.20}$$

Therefore,

$$E^\mu[\langle \mathcal{Y}_t, f\rangle] = e^{\beta t}\langle \mu, S_t f\rangle \qquad (t \geq 0,\ f \in B(E)), \tag{2.6.21}$$

and

$$\begin{aligned}\mathrm{Var}^\mu(\langle\mathcal{Y}_t,f\rangle) &= 2\int_0^t \mathrm{d}s\, e^{\beta s}e^{2\beta(t-s)}\langle\mu,S_s(\alpha(S_{t-s}f)^2)\rangle \\ &\le 2\|\alpha\|_\infty\|f\|_\infty^2\langle\mu,1\rangle e^{\beta t}\int_0^t \mathrm{d}s\, e^{\beta(t-s)} \\ &\le 2\beta^{-1}\|\alpha\|_\infty\|f\|_\infty^2\langle\mu,1\rangle e^{2\beta t} \qquad (t\ge 0,\ f\in B(E)).\end{aligned} \tag{2.6.22}$$

Let $(\mathcal{F}_t)_{t\ge 0}$ be the filtration generated by $\mathcal{Y}$ and put

$$\tilde{\mathcal{Y}}_t := e^{-\beta t}\mathcal{Y}_t \qquad (t\ge 0). \tag{2.6.23}$$

Then (2.6.21) and (2.6.22) show that for any $0\le s\le t$ and $f\in B(E)$,

$$\begin{aligned}&\text{(i)} \quad E^\mu\big[\langle\tilde{\mathcal{Y}}_t,f\rangle\big|\mathcal{F}_s\big] = \langle\tilde{\mathcal{Y}}_s,S_{t-s}f\rangle \quad \text{a.s.},\\ &\text{(ii)} \quad \mathrm{Var}^\mu\big[\langle\tilde{\mathcal{Y}}_t,f\rangle\big|\mathcal{F}_s\big] \le 2\beta^{-1}\|\alpha\|_\infty\|f\|_\infty^2\langle\tilde{\mathcal{Y}}_s,1\rangle e^{-\beta s} \quad \text{a.s.}\end{aligned} \tag{2.6.24}$$

Since $S_{t-s}1=1$, formula (2.6.24) (i) shows that $(\langle\tilde{\mathcal{Y}}_t,1\rangle)_{t\ge 0}$ is a nonnegative martingale, and hence there exists a nonnegative random variable W such that (2.6.19) (i) holds. Setting $s=0$ in (2.6.24) (ii), we see that

$$\mathrm{Var}^\mu\big[\langle\tilde{\mathcal{Y}}_t,1\rangle\big] \le 2\beta^{-1}\|\alpha\|_\infty\langle\mu,1\rangle \qquad (t\ge 0). \tag{2.6.25}$$

This implies (2.6.19) (ii), and, using Fatou, (2.6.19) (iv). Moreover, by (2.6.25) the random variables $\langle\mathcal{Y}_t,1\rangle_{t\ge 0}$ are uniformly integrable, and therefore (2.6.19) (iii) holds.

We are left with the task to prove (2.6.19) (v). The inclusion $\supset$ is trivial. Formulas (2.6.19) (iii) and (2.6.19) (iv) imply that

$$\langle\mu,1\rangle^2P^\mu[W=0] \le \mathrm{Var}^\mu(W) \le 2\beta^{-1}\|\alpha\|_\infty\langle\mu,1\rangle, \tag{2.6.26}$$

and therefore

$$P^\mu[W>0] \ge 1-2\beta^{-1}\|\alpha\|_\infty\langle\mu,1\rangle^{-1} \qquad (\mu\neq 0). \tag{2.6.27}$$

Note that $\{W>0\}$ is a tail event. Thus, by Lemma 2.64,

$$\lim_{t\to\infty} P^{\mathcal{Y}_t}[W>0] = 1_{\{W>0\}} \quad \text{a.s.} \tag{2.6.28}$$

Formula (2.6.27) shows that

$$\liminf_{t\to\infty} P^{\mathcal{Y}_t}[W>0] \ge 1_{\{\lim_{t\to\infty}\langle\mathcal{Y}_t,1\rangle=\infty\}}. \tag{2.6.29}$$

Combining Lemma 2.56 with formulas (2.6.28) and (2.6.29) we see that $\{\mathcal{Y}_t=0 \text{ eventually}\}^{\mathrm{c}} \subset \{\lim_{t\to\infty}\langle\mathcal{Y}_t,1\rangle=\infty\} \subset \{W>0\}$ a.s. ∎

2.6.3 Smoothness of two log-Laplace semigroups

We return to the special situation $E = [0,1]$ and $G = \overline{A}$ or $G = \overline{A^v}$, where $\overline{A}$ and $\overline{A^v}$ are the closures in $\mathcal{C}(E)$ of the operators A in (2.5.5) and A^v in (2.5.20), respectively, with domains $\mathcal{D}(A) = \mathcal{D}(A^v) := \mathcal{C}^{(2)}[0,1]$, the space of real functions on $[0,1]$ that are twice continuously differentiable. Let $\mathcal{U} = \mathcal{U}(\overline{A}, \alpha, \alpha)$ and $\mathcal{U}^v = \mathcal{U}(\overline{A^v}, \alpha v, \alpha - 1)$ denote the log-Laplace semigroups of the super-Wright-Fisher diffusion $\mathcal{Y}$ and the weighted super-Wright-Fisher diffusion $\mathcal{Y}^v$, respectively, where $\alpha > 0$ is constant. In this section we prove:

Lemma 2.59 (Smoothing property of $\mathcal{U}$ and $\mathcal{U}^v$) *One has $\mathcal{U}_t(B_+[0,1]) \subset \mathcal{C}_+[0,1]$ and $\mathcal{U}^v_t(B_+[0,1]) \subset \mathcal{C}_+[0,1]$ for all $t > 0$. Moreover, if $\text{bp-lim}_{n\to\infty} f_n = f$ for some $f_n, f \in B_+[0,1]$, then $\lim_{n\to\infty} \|\mathcal{U}_t f_n - \mathcal{U}_t f\|_\infty = 0$ and $\lim_{n\to\infty} \|\mathcal{U}^v_t f_n - \mathcal{U}^v_t f\|_\infty = 0$ for all $t > 0$.*

To prepare for the proof, we start with the following elementary property of the semigroups S and S^v generated by $\overline{A}$ and $\overline{A^v}$, respectively (recall (2.5.5) and (2.5.20)).

Lemma 2.60 (Strong Feller property) *The semigroups S and S^v have the strong Feller property, i.e., $S_t(B[0,1]) \subset \mathcal{C}[0,1]$ and $S^v_t(B[0,1]) \subset \mathcal{C}[0,1]$ for all $t > 0$.*

Proof Couple two realizations ξ^x, ξ^y of the process with generator $\overline{A}$, started in $x, y \in [0,1]$, in such a way that ξ^x and ξ^y move independently up to the random time $\tau := \inf\{t \geq 0 : \xi^x_t = \xi^y_t\}$, and such that $\xi^x_t = \xi^y_t$ for all $t \geq \tau$. (Here the superscript in ξ^x refers to the initial condition, and not, like elsewhere, to a compensated h-transform.) Then it is not hard to see that

$$P[\xi^y_t = \xi^x_t] \to 1 \quad \text{as} \quad y \to x \quad \forall t > 0. \tag{2.6.30}$$

In particular, (2.6.30) holds also for $x \in \{0,1\}$ since the boundary is attainable. Since $|S_t f(x) - S_t f(y)| \leq 2\|f\|_\infty P[\xi^x_t \neq \xi^y_t]$, formula (2.6.30) shows that $S_t f \in \mathcal{C}[0,1]$ for all $f \in B[0,1]$ and $t > 0$. For the process with generator $\overline{A^v}$ the argument is similar but easier, since in this case $\{0,1\}$ is an entrance boundary. ∎

Proof of Lemma 2.59 For each $f \in B[0,1]$, the function $u_t := \mathcal{U}_t f$ is a mild solution of (2.5.6), i.e., (see (2.6.1))

$$\mathcal{U}_t f = S_t f + \int_0^t S_{t-s}\big(\alpha \mathcal{U}_s f(1 - \mathcal{U}_s f)\big)\, \mathrm{d}s \qquad (t \geq 0). \tag{2.6.31}$$

By the strong Feller property of $(S_t)_{t\geq 0}$ (Lemma 2.60), the functions $S_t f$ and $S_{t-s}(\alpha \mathcal{U}_s f(1 - \mathcal{U}_s f))$ are continuous for each $0 \leq s < t$, and therefore $\mathcal{U}_t f$ is continuous.

Now let $f_n \to f$ in a bounded pointwise way for some $f_n, f \in B_+[0,1]$, and let $t > 0$. By Lemma 2.51, $\mathcal{U}_t f_n \to \mathcal{U}_t f$ in a bounded pointwise way. By the strong Feller property of $(S_t)_{t\geq 0}$ and [Rev84, Prop. 1.5.8 and Thm. 1.5.9], $S_t f_n$ converges uniformly to $S_t f$ and the function $(x,s) \mapsto S_{t-s}\big(\alpha \mathcal{U}_s f_n(1 - \mathcal{U}_s f_n)\big)(x)$ converges uniformly on $[0,1] \times [0, t-\varepsilon]$ to $S_{t-s}\big(\alpha \mathcal{U}_s f(1 - \mathcal{U}_s f)\big)(x)$, for all $\varepsilon > 0$. By (2.6.31), it follows that $\mathcal{U}_t f_n \to \mathcal{U}_t f$ uniformly on $[0,1]$.

The same arguments apply to $\mathcal{U}^v_t f$. ∎

2.6.4 Bounds on the absorption probability

Let $\mathcal{U} = \mathcal{U}(\overline{A}, \alpha, \alpha)$. Since the points $0, 1$ are traps for the Wright-Fisher diffusion, $f(r) = 0$ implies $\mathcal{U}_t f(r) = 0$ $(r = 0, 1)$. We have already seen (Lemma 2.59) that $\mathcal{U}_t f$ is continuous for each $t > 0$. The following lemma shows that if $f(r) = 0$, then $\mathcal{U}_t f$ has a finite slope at $r = 0, 1$, for all $t > 0$. By symmetry, it suffices to consider the case $r = 0$.

Lemma 2.61 (Absorption of the super-Wright-Fisher diffusion) *Let $\mathcal{U} = \mathcal{U}(\overline{A}, \alpha, \alpha)$, with $\alpha > 0$. Then*

$$\mathcal{U}_t(\infty 1_{(0,1]})(x) \leq K_t\, x \qquad (t > 0,\ x \in [0, 1]), \tag{2.6.32}$$

with

$$K_t := \frac{e^{\alpha t/2}}{1 - e^{-\alpha t/2}}\Big(\frac{8}{t} + 2\Big) \qquad (t > 0). \tag{2.6.33}$$

Note that (2.6.32) implies that

$$P^{\delta_x}[\mathcal{Y}_t((0, 1]) > 0] \leq 1 - e^{-K_t\, x} \leq K_t\, x \qquad (t > 0,\ x \in [0, 1]). \tag{2.6.34}$$

We begin with a preparatory lemma.

Lemma 2.62 (Absorption of the Wright-Fisher diffusion) *For the Wright-Fisher diffusion ξ,*

$$P^x[\xi_t > 0] \leq \Big(\frac{4}{t} + 2\Big)x \qquad (t > 0,\ x \in [0, 1]). \tag{2.6.35}$$

Proof For $x \geq 0$ put

$$f_0(x) := 1_{\{0\}}(x) \quad \text{and} \quad f_t(x) := (1 - 2x)e^{-\frac{4x}{t}} 1_{[0,\frac{1}{2}]}(x) \qquad (t > 0). \tag{2.6.36}$$

A little calculation shows that for $t > 0$ and $x \geq 0$,

$$\begin{aligned} \tfrac{\partial}{\partial t} f_t(x) &= 4x(1 - 2x)t^{-2} e^{-\frac{4x}{t}} 1_{[0,\frac{1}{2}]}(x) \\ \tfrac{1}{2}x(1 - x)D_x^2 f_t(x) &= \big(8x(1 - x)(1 - 2x)t^{-2} e^{-\frac{4x}{t}} + 8x(1 - x)t^{-1} e^{-\frac{4x}{t}}\big) 1_{[0,\frac{1}{2}]}(x) \\ &\quad + 2e^{-\frac{2}{t}} \delta_{\frac{1}{2}}(x), \end{aligned} \tag{2.6.37}$$

where D_x^2 denotes the generalized second derivative with respect to x and $\delta_{\frac{1}{2}}$ is the delta-function at $\frac{1}{2}$. Since $4x \leq 8x(1 - x)$ for all $x \in [0, \frac{1}{2}]$, it follows that

$$\tfrac{\partial}{\partial t} f_t(x) \leq \tfrac{1}{2}x(1 - x)D_x^2 f_t(x) \qquad (t > 0,\ x \geq 0). \tag{2.6.38}$$

If f_t were contained in $\mathcal{D}(\overline{A})$, then (2.6.38) would mean that $\frac{\partial}{\partial t} f_t \leq \overline{A} f_t$ for $t > 0$, and a standard argument (compare Lemma 2.54) would tell us that $f_t \leq S_t f_0$, where S is the semigroup of ξ. In the present case, we need a little approximation argument.

Let $\phi_n \geq 0$ $(n \geq 0)$ denote $\mathcal{C}^{(\infty)}$-functions defined on $[0, \infty)$ with support contained in $[0, \frac{1}{3}]$, say, such that $\phi_n(x)\mathrm{d}x$ are probability measures converging weakly to the δ-measure δ_0 as $n \to \infty$. Put

$$f_t^n(x) := \int_0^\infty \mathrm{d}y\, \phi_n(y) f_t(x + y) =: \phi_n * f_t(x) \qquad (t > 0,\ x \geq 0). \tag{2.6.39}$$

Then

$$
\begin{aligned}
\tfrac{\partial}{\partial t} f^n_t(x) &= \phi_n * \tfrac{\partial}{\partial t} f_t(x) \\
\tfrac{\partial^2}{\partial x^2} f^n_t(x) &= \phi_n * D_x^2 f_t(x),
\end{aligned}
\tag{2.6.40}
$$

and therefore (2.6.38) shows that

$$
\tfrac{\partial}{\partial t} f^n_t(x) \leq \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2} f^n_t(x) \qquad (t>0,\ x\geq 0,\ n\geq 0). \tag{2.6.41}
$$

Since $f^n_t \in \mathcal{D}(\overline{A})$ for all $t>0$, the argument mentioned above gives

$$
f^n_{t+\varepsilon} \leq S_t f^n_\varepsilon \qquad (t\geq 0,\ \varepsilon>0). \tag{2.6.42}
$$

Letting $n\to\infty$ and afterwards $\varepsilon\to 0$ we find that

$$
f_t(x) \leq S_t f_0(x) = P^x[\xi_t = 0] \qquad (t\geq 0,\ x\in[0,1]). \tag{2.6.43}
$$

Note that $\frac{\partial}{\partial x}(1-f_t(x)) = (1-2x)4t^{-1}e^{-\frac{4x}{t}} + 2e^{-\frac{4x}{t}} \leq (\frac{4}{t}+2)$ for $x\in[0,\frac{1}{2}]$. Therefore (2.6.43) implies (2.6.35). (Note that (2.6.35) is trivial for $x\in[\frac{1}{2},1]$.) ∎

Proof of Lemma 2.61 Fix $f\in B_+[0,1]$ satisfying $f(0)=0$ and write $\mathcal{U}_t f = \mathcal{U}_{t/2}\mathcal{U}_{t/2}f$. By (2.6.10) from Lemma 2.55, $\mathcal{U}_{t/2}f \leq (1-e^{-\alpha t/2})^{-1}$. Since moreover $\mathcal{U}_{t/2}f(0)=0$ because of absorption at zero, we have

$$
\mathcal{U}_t f \leq \mathcal{U}_{t/2}((1-e^{-\alpha t/2})^{-1}1_{(0,1]}) \qquad (t>0). \tag{2.6.44}
$$

Using (2.6.8) from Lemma 2.55, we may estimate $\mathcal{U}(\overline{A},\alpha,\alpha)$ in terms of $\mathcal{U}(\overline{A},0,\alpha)$, which is just the linear semigroup $(e^{\alpha t}S_t)_{t\geq 0}$. Thus, by Lemma 2.62,

$$
\begin{aligned}
\mathcal{U}_t f(x) &\leq e^{\alpha t/2} S_{t/2}((1-e^{-\alpha t/2})^{-1}1_{(0,1]})(x) \\
&\leq e^{\alpha t/2}(1-e^{-\alpha t/2})^{-1}(\tfrac{8}{t}+2)x \qquad (t>0,\ x\in[0,1]).
\end{aligned}
\tag{2.6.45}
$$

Letting $f\uparrow\infty$, by monotonicity we arrive at (2.6.32). ∎

2.6.5 The weighted super-Wright-Fisher diffusion

In this section we prove Lemmas 2.49 and 2.50. Recall that ξ,ξ^v are the diffusions in $[0,1]$ with generators $\overline{A},\overline{A^v}$ defined in (2.5.5) and (2.5.20), and associated semigroups S,S^v, respectively, and that $\mathcal{U}=\mathcal{U}(\overline{A},\alpha,\alpha)$ and $\mathcal{U}^v=\mathcal{U}(\overline{A^v},\alpha v,\alpha-1)$.

Lemma 2.63 (v-transformed log-Laplace semigroup) *If $f\in\mathcal{D}(\overline{A^v})$, then $vf\in\mathcal{D}(\overline{A})$ and*

$$
\overline{A}(vf) = v\,(\overline{A^v}-1)f. \tag{2.6.46}
$$

Moreover,

$$
\mathcal{U}_t(vf) = v\,\mathcal{U}^v_t f \qquad (t\geq 0,\ f\in B_+[0,1]). \tag{2.6.47}
$$

Proof For any $f \in \mathcal{C}^{(2)}[0,1]$, it is easy to check that

$$A(vf) = v\,(A^v - 1)f. \tag{2.6.48}$$

Fix $f \in \mathcal{D}(\overline{A^v})$ and choose $f_n \in \mathcal{C}^{(2)}[0,1]$ such that $f_n \to f$ in $\mathcal{C}[0,1]$. Then (2.6.48) shows that $A(vf_n) \to v\,(\overline{A^v} - 1)f$, which implies that $vf \in \mathcal{D}(\overline{A})$ and that (2.6.46) holds.

Now fix $f \in \mathcal{C}_+[0,1] \cap \mathcal{D}(\overline{A^v})$ and put $u^v_t := \mathcal{U}^v_t f$ $(t \geq 0)$. Then u^v is the classical solution of the Cauchy equation

$$\begin{cases} \frac{\partial}{\partial t} u^v_t = \overline{A^v} u^v_t + (\alpha - 1) u^v_t - \alpha v\,(u^v_t)^2 \qquad (t \geq 0), \\ u^v_0 = f. \end{cases} \tag{2.6.49}$$

It follows from (2.6.46) that

$$\begin{aligned} \frac{\partial}{\partial t} v u^v_t &= v \frac{\partial}{\partial t} u^v_t = v\overline{A^v} u^v_t + (\alpha - 1) v u^v_t - \alpha\,(v u^v_t)^2 \\ &= \overline{A}(v u^v_t) + \alpha v u^v_t - \alpha\,(v u^v_t)^2 \qquad (t \geq 0), \end{aligned} \tag{2.6.50}$$

i.e., $u_t := v u^v_t$ is the classical solution to the Cauchy equation

$$\begin{cases} \frac{\partial}{\partial t} u_t = \overline{A} u_t + \alpha u_t - \alpha u_t^2 \qquad (t \geq 0), \\ u_0 = v f. \end{cases} \tag{2.6.51}$$

This proves that $\mathcal{U}_t(vf) = u_t = v u^v_t = v \mathcal{U}^v_t f$ for all $f \in \mathcal{C}_+[0,1] \cap \mathcal{D}(\overline{A^v})$. The general case follows from Lemma 2.52 and the fact that the class of $f \in B_+[0,1]$ for which (2.6.47) holds is closed under bounded pointwise limits. ∎

Proof of Lemma 2.49 Set $\mathcal{F}_t := \sigma(\mathcal{Y}_s : 0 \leq s \leq t)$. Then by (2.6.47), for all $0 \leq s \leq t$ and $f \in B_+[0,1]$,

$$\begin{aligned} &E\big[e^{-\langle v\mathcal{Y}_t, f\rangle} \big| \mathcal{F}_s\big] = E\big[e^{-\langle \mathcal{Y}_t, vf\rangle} \big| \mathcal{F}_s\big] = e^{-\langle \mathcal{Y}_s, \mathcal{U}_{t-s}(vf)\rangle} \\ &= e^{-\langle \mathcal{Y}_s, v\mathcal{U}^v_{t-s} f\rangle} = e^{-\langle v\mathcal{Y}_s, \mathcal{U}^v_{t-s} f\rangle}. \end{aligned} \tag{2.6.52}$$

It follows that $(v\mathcal{Y}_t)_{t \geq 0}$ is a Markov process and that its transition probabilities coincide with those of the $(\overline{A^v}, \alpha v, \alpha - 1)$-superprocess. Since $\mathcal{Y}$ has continuous sample paths, so has $v\mathcal{Y}$. ∎

Proof of Lemma 2.50 We need to prove (2.5.22), which by (2.6.12) is equivalent to the statement that $\|\mathcal{U}^v_t \infty\|_\infty < \infty$ for all $t > 0$. Assume that $f \in B_+[0,1]$ satisfies $f(0) = f(1) = 0$. By Lemma 2.61, $\mathcal{U}_t f(x) \leq K_t x$ for the constant K_t mentioned there. By symmetry, one also has $\mathcal{U}_t f(x) \leq K_t (1 - x)$ and, since $x \wedge (1 - x) \leq \frac{1}{3} v(x)$, $\mathcal{U}_t f(x) \leq \frac{1}{3} K_t\, v(x)$. Let $g \in B_+[0,1]$. By formula (2.6.47) and the fact that $(vg)(0) = (vg)(1) = 0$, we see that $\mathcal{U}^v_t g(x) = \frac{1}{v(x)} \mathcal{U}_t(vg)(x) \leq \frac{1}{3} K_t$ for all $x \in (0,1)$. By Lemma 2.60, $\mathcal{U}^v_t g$ is continuous on $[0,1]$ and therefore $\mathcal{U}^v_t g(x) \leq \frac{1}{3} K_t$ holds also for $x = 0, 1$. Taking the limit $g \uparrow \infty$ we see that $\|\mathcal{U}^v_t \infty\|_\infty \leq \frac{1}{3} K_t < \infty$ for all $t > 0$. ∎

2.6.6 A zero-one law for Markov processes

Let E be a Polish space and let $(P^x)^{x\in E}$ be a family of probability measures on $\mathcal{D}_E[0,\infty)$ (the space of cadlag functions $w:[0,\infty)\to E$) such that under $(P^x)^{x\in E}$, the coordinate projections $\{w\mapsto w_t =: \xi_t(w) : t\geq 0\}$ form a Borel right process in the sense of [Sha88]. This is true, for example, if $(P^x)^{x\in E}$ are the laws of a Feller process on a locally compact Polish space, or a (G,α,β)-superprocess as introduced in Section 2.6.1, see [Fit88]. Let $\mathcal{T} := \bigcap_{t\geq 0}\sigma(\xi_s : s\geq t)$ denote the tail-σ-field of ξ. Let $(\theta_t w)_s := w_{t+s}$ $(t,s\geq 0)$ be the time-shift on $\mathcal{D}_E[0,\infty)$. Then the following holds.

Lemma 2.64 (Zero-one law for Markov processes) *Assume that $A\in\mathcal{T}$. Then for each $x\in E$,*

$$\lim_{t\to\infty} P^{\xi_t}(\theta_t^{-1}(A)) = 1_A \quad P^x\text{−a.s.} \tag{2.6.53}$$

Proof Let $\mathcal{F}_t := \sigma(\xi_s : 0\leq s\leq t)$ $(t\geq 0)$ be the filtration generated by ξ and set $\mathcal{F}_\infty := \sigma(\xi_s : s\geq 0)$. Since ξ is a Markov process, $P^{\xi_t}(\theta_t^{-1}(A)) = P[A|\mathcal{F}_t]$ a.s. For any sequence of times $t_n\uparrow\infty$ one has $\mathcal{F}_{t_n}\uparrow\mathcal{F}_\infty$ and therefore $P[A|\mathcal{F}_{t_n}]\to P[A|\mathcal{F}_\infty] = 1_A$ a.s., see [Loe63, § 29, Complement 10 (b)]. Since ξ is a right process, the function $t\mapsto P^{\xi_t}(\theta_t^{-1}(A))$ is a.s. right-continuous, see [Sha88, Theorem (7.4.viii)], and we conclude that (2.6.53) holds. ∎

2.7 The super-Wright-Fisher diffusion: long-time behavior

2.7.1 Ergodicity of the compensated v-transformed Wright-Fisher diffusion

Recall that ξ^v is the diffusion on $[0,1]$ with generator $\overline{A^v}$ defined in (2.5.20) and associated semigroup S^v. As in Theorem 2.46, ℓ denotes the Lebesgue measure on $(0,1)$ and v is defined by (2.5.9). In this section we prove:

Lemma 2.65 (Ergodicity of the compensated v-transformed Wright-Fisher diffusion) *The Markov process ξ^v has the unique invariant law $v\ell$ and is ergodic:*

$$\lim_{t\to\infty}\|S^v_t f - \langle v\ell, f\rangle\|_\infty = 0 \qquad \forall f\in B[0,1]. \tag{2.7.1}$$

Proof Since

$$\tfrac{\partial}{\partial x}\big[\tfrac12 x(1-x)v(x)\big] = 2(\tfrac12 - x)v(x) \qquad (x\in[0,1]), \tag{2.7.2}$$

$v\ell$ is a (reversible) invariant law for the process with generator $\overline{A^v}$, see [EK86, Proposition 4.9.2]. Fix $x\in[0,1]$. Let ξ^v be the process started in x and let $\tilde\xi^v$ be the process started in the invariant law $v\ell$. Then $\xi^v,\tilde\xi^v$ may represented as solutions to the SDE

$$\mathrm{d}\xi^v_t = 2(\tfrac12 - \xi^v_t)\mathrm{d}t + \sqrt{\xi^v_t(1-\xi^v_t)}\,\mathrm{d}B_t, \tag{2.7.3}$$

relative to the same Brownian motion B. Using the technique of Yamada & Watanabe (see [YW71] or, for example, [EK86, Theorem 5.3.8]), it is easy to prove that

$$E[|\xi^v_t - \tilde\xi^v_t|] = e^{-2t}E[|\xi^v_0 - \tilde\xi^v_0|] \leq e^{-2t} \qquad (t\geq 0). \tag{2.7.4}$$

It follows that for any function f satisfying $|f(y)-f(z)| \le |y-z| \quad (y,z \in [0,1])$,

$$\big|E[f(\xi^v_t)] - \langle v\ell, f\rangle\big| \le E[|f(\xi^v_t) - f(\tilde\xi^v_t)|] \le e^{-2t}. \tag{2.7.5}$$

This implies that the function $x \mapsto \mathcal{L}^x(\xi^v_t)$ from $[0,1]$ into the space $\mathcal{M}_1[0,1]$ of probability measures on $[0,1]$, converges as $t\to\infty$ uniformly to the constant function $v\ell$. This shows that (2.7.1) holds for all $f \in \mathcal{C}[0,1]$. Since ξ^v has the strong Feller property (Lemma 2.60), (2.7.1) holds for all $f \in B[0,1]$. ∎

2.7.2 Long-time behavior of the weighted super-Wright-Fisher diffusion

The following lemma prepares for the proof of formula (2.5.15) in Theorem 2.46.

Lemma 2.66 (Mean square convergence) *Assume that $\alpha > 1$. Let $\mathcal{Y}^v$ be the $(\overline{A^v}, \alpha v, \alpha-1)$-superprocess started in $\mathcal{Y}^v_0 = \mu \in \mathcal{M}[0,1]$. Then there exists a nonnegative random variable W, depending on μ, such that*

$$\begin{aligned} &\text{(i)} \quad \lim_{t\to\infty} e^{-(\alpha-1)t}\langle \mathcal{Y}^v_t, 1\rangle = W \quad \text{a.s.}\\ &\text{(ii)} \quad \lim_{t\to\infty} E^\mu\Big[\big|e^{-(\alpha-1)t}\langle \mathcal{Y}^v_t, f\rangle - W\langle v\ell, f\rangle\big|^2\Big] = 0 \quad \forall f \in B[0,1]. \end{aligned} \tag{2.7.6}$$

Moreover,

$$E^\mu(W) = \langle \mu, 1\rangle \quad \textit{and} \quad \mathrm{Var}^\mu(W) \le 3\tfrac{\alpha}{\alpha-1}\langle \mu, 1\rangle, \tag{2.7.7}$$

and

$$\{W=0\} = \{\mathcal{Y}^v_t = 0 \ \textit{eventually}\} \qquad \text{a.s.} \tag{2.7.8}$$

Proof Except for formula (2.7.6) (ii), all statements are direct consequences of the fact that $\mathcal{Y}^v$ has the finite ancestry property (Lemma 2.50) and of Lemma 2.58 (note that $\|\alpha v\|_\infty = \frac{3}{2}\alpha$).

Fix $f \in B[0,1]$. Let $(\mathcal{F}_t)_{t\ge0}$ be the filtration generated by $\mathcal{Y}^v$ and put $\tilde{\mathcal{Y}}^v_t := e^{-(\alpha-1)t}\mathcal{Y}^v_t$ $(t \ge 0)$. Pick $1 \le s_n \le t_n$ such that $s_n \to \infty$ and $t_n - s_n \to \infty$. Then, by (2.6.24),

$$E^\mu\Big[\big|\langle \tilde{\mathcal{Y}}^v_{t_n}, f\rangle - \langle \tilde{\mathcal{Y}}^v_{s_n}, S^v_{t_n-s_n} f\rangle\big|^2 \Big| \mathcal{F}_{s_n}\Big] \le 3\tfrac{\alpha}{\alpha-1}\|f\|^2_\infty \langle \tilde{\mathcal{Y}}^v_{s_n}, 1\rangle e^{-(\alpha-1)s_n} \quad \text{a.s.} \tag{2.7.9}$$

Taking expectations on both sides in (2.7.9), one finds that

$$E^\mu\Big[\big|\langle \tilde{\mathcal{Y}}^v_{t_n}, f\rangle - \langle \tilde{\mathcal{Y}}^v_{s_n}, S^v_{t_n-s_n} f\rangle\big|^2\Big] \le 3\tfrac{\alpha}{\alpha-1}\|f\|^2_\infty \langle \mu, 1\rangle e^{-(\alpha-1)s_n}. \tag{2.7.10}$$

By (2.6.19) (ii),

$$\lim_{t\to\infty} E^\mu\big[|\langle \tilde{\mathcal{Y}}^v_t, 1\rangle - W|^2\big] = 0. \tag{2.7.11}$$

Using Lemma 2.65 (about the ergodicity of ξ^v) and (2.7.11), it is easy to show that

$$\lim_{n\to\infty} E^\mu\Big[\big|\langle \tilde{\mathcal{Y}}^v_{s_n}, S^v_{t_n-s_n} f\rangle - W\langle v\ell, f\rangle\big|^2\Big] = 0. \tag{2.7.12}$$

Combining this with (2.7.10), we see that

$$\lim_{n\to\infty} E^\mu\Big[\big|\langle \tilde{\mathcal{Y}}^v_{t_n}, f\rangle - W\langle v\ell, f\rangle\big|^2\Big] = 0. \tag{2.7.13}$$

Since this is true for any $t_n \to \infty$, (2.7.6) (ii) follows. ∎

2.7.3 Long-time behavior of the super-Wright-Fisher diffusion

Proof of Theorem 2.46 Using Lemma 2.49, we can translate our results on the weighted super-Wright-Fisher diffusion $\mathcal{Y}^v$ to the super-Wright-Fisher diffusion $\mathcal{Y}$. Thus, Lemma 2.66 proves formulas (2.5.10) (ii), (2.5.11) (ii), and (2.5.14)–(2.5.15), where $W_{(0,1)}$ is the random variable W from Lemma 2.66. Formula (2.5.13) follows from Lemma 2.57. To finish the proof of Theorem 2.46, it suffices to prove (2.5.10) (i), (2.5.11) (i) and (2.5.12).

1°. Proof of formula (2.5.10) (i) One has $E^\mu[\langle \mathcal{Y}_t, f\rangle] = e^{\alpha t}\langle \mu, S_t f\rangle$ for all $t \geq 0$, $f \in B[0,1]$ by (2.6.21). Since the points $r = 0, 1$ are traps for the Wright-Fisher diffusion, $E^\mu[\langle \mathcal{Y}_t, 1_{\{r\}}\rangle] = e^{\alpha t}\langle \mu, S_t 1_{\{r\}}\rangle \geq e^{\alpha t}\langle \mu, 1_{\{r\}}\rangle$ for all $t \geq 0$, $r = 0, 1$. Thus, the processes $(e^{-\alpha t}\langle \mathcal{Y}_t, 1_{\{r\}}\rangle)_{t\geq 0}$ ($r = 0, 1$) are nonnegative submartingales, and hence there exist random variables W_r ($r = 0, 1$) such that (2.5.10) (i) holds.

2°. Proof of formula (2.5.12) For $\alpha \leq 1$ the statement is trivial by (2.5.13), so assume $\alpha > 1$. By symmetry it suffices to consider the case $r = 0$. From the L_2-convergence formula (2.5.15) we have, for any $K > 0$,

$$\{W_{(0,1)} > 0\} \subset \big\{\forall T < \infty\ \exists t \geq T \text{ such that } \mathcal{Y}_t([\tfrac{1}{4}, \tfrac{1}{3}]) \geq K\big\} \quad \text{a.s.} \tag{2.7.14}$$

Assume for the moment that for some $t > 0$ and (sufficiently large) K,

$$\inf_{\mu:\, \mu([\frac{1}{4},\frac{1}{3}])\geq K} P^\mu[W_0 > 0] > 0. \tag{2.7.15}$$

Then we see from (2.7.14) and (2.7.15) that

$$\{W_{(0,1)} > 0\} \subset \big\{\lim_{t\to\infty} P^{\mathcal{Y}_t}[W_0 > 0] = 0\big\}^{\mathrm{c}} \subset \{W_0 > 0\} \quad \text{a.s.}, \tag{2.7.16}$$

where the second inclusion follows from the fact that, by Lemma 2.64,

$$\lim_{t\to\infty} P^{\mathcal{Y}_t}[W_0 > 0] = 1_{\{W_0>0\}} \quad \text{a.s.} \tag{2.7.17}$$

Thus, we are done if we can prove (2.7.15). By the branching property, it suffices to prove (2.7.15) for measures μ that are concentrated on $[\frac{1}{4}, \frac{1}{3}]$. Fix any $t > 0$. Formulas (2.6.21) and (2.6.22) give

$$\begin{aligned} &\text{(i)} \quad E^\mu\big[\langle \mathcal{Y}_t, 1_{\{0\}}\rangle\big] = \langle \mu, S_t 1_{\{0\}}\rangle e^{\alpha t}, \\ &\text{(ii)} \quad \mathrm{Var}^\mu\big[\langle \mathcal{Y}_t, 1_{\{0\}}\rangle\big] \leq 2\langle \mu, 1\rangle e^{2\alpha t}. \end{aligned} \tag{2.7.18}$$

It follows from formula (2.6.43) (recall (2.6.36)) that

$$\inf_{x\in[\frac{1}{4},\frac{1}{3}]} S_t 1_{\{0\}}(x) > 0. \tag{2.7.19}$$

Denoting the infimum by ε, we get the bounds

$$\begin{aligned} &\text{(i)} \quad E^\mu\big[\langle \mathcal{Y}_t, 1_{\{0\}}\rangle\big] \geq \varepsilon\langle \mu, 1\rangle e^{\alpha t}, \\ &\text{(ii)} \quad \mathrm{Var}^\mu\big[\langle \mathcal{Y}_t, 1_{\{0\}}\rangle\big] \leq 2\langle \mu, 1\rangle e^{2\alpha t}. \end{aligned} \tag{2.7.20}$$

These formulas show that for large $\langle \mu, 1\rangle$, the standard deviation of $\langle \mathcal{Y}_t, 1_{\{0\}}\rangle$ is small compared to its mean. Therefore, using Chebyshev's inequality, it is easy to show that for every $M > 0$ there exists a $K > 0$ such that

$$\inf_{\mu\in\mathcal{M}[\frac{1}{4},\frac{1}{3}]:\,\langle\mu,1\rangle\geq K} P^\mu[\langle \mathcal{Y}_t, 1_{\{0\}}\rangle \geq M] > 0. \tag{2.7.21}$$

Hence, by the Markov property, in order to prove (2.7.15) it suffices to show that for M sufficiently large,

$$\inf_{\mu:\,\mu(\{0\})\geq M} P^\mu[W_0 > 0] > 0. \tag{2.7.22}$$

By the branching property, it suffices to prove (2.7.22) for measures μ that are concentrated on $\{0\}$. In that case, $\mathcal{Y}_t(\{0\})_{t\geq 0}$ is an autonomous supercritical Feller's branching diffusion (a superprocess in a single-point space is just a Feller's branching diffusion). Applying Lemma 2.58 to this Feller's branching diffusion, again using Chebyshev, it is not hard to prove (2.7.22). Since the arguments are very similar to those we have already seen, we skip the details.

3°. Proof of formula (2.5.11) (i) The inclusion $\{W_r = 0\} \supset \{\mathcal{Y}_t(\{r\}) = 0 \text{ eventually}\}$ a.s. is trivial. By (2.5.12) and (2.5.11) (ii), $\{W_r = 0\} \subset \{W_{(0,1)} = 0\} \subset \{\mathcal{Y}_t((0,1)) = 0 \text{ eventually}\}$ a.s. Therefore, by the strong Markov property, it suffices to prove $\{W_r = 0\} \subset \{\mathcal{Y}_t(\{r\}) = 0 \text{ eventually}\}$ a.s. for the process started in μ with $\mu((0,1)) = 0$. In this case, $(\mathcal{Y}_t(\{r\}))_{t\geq 0}$ is an autonomous supercritical Feller's branching diffusion, and the statement is easy (see the previous parapraph). ∎

2.7.4 Long-time behavior of the log-Laplace semigroup

Proof of Proposition 2.47 We start by proving that for all $\mu \in \mathcal{M}[0,1]$ and $f \in B_+[0,1]$,

$$\begin{aligned}
&\lim_{t\to\infty} e^{-\langle \mu, \mathcal{U}_t f\rangle} \\
&= P^\mu\Big[\big\{f(0) = 0 \text{ or } W_0 = 0\big\} \cap \big\{f(1) = 0 \text{ or } W_1 = 0\big\} \cap \big\{\langle \ell, f\rangle = 0 \text{ or } W_{(0,1)} = 0\big\}\Big] \\
&= \begin{cases}
1 & \text{if } \; f(0) = f(1) = \langle \ell, f\rangle = 0, \\
P^\mu\big[W_{(0,1)} = 0\big] & \text{if } \; f(0) = f(1) = 0,\ \langle \ell, f\rangle > 0, \\
P^\mu\big[W_0 = 0\big] = P^\mu\big[W_0 = W_{(0,1)} = 0\big] & \text{if } \; f(0) > 0,\ f(1) = 0, \\
P^\mu\big[W_1 = 0\big] = P^\mu\big[W_1 = W_{(0,1)} = 0\big] & \text{if } \; f(0) = 0,\ f(1) > 0, \\
P^\mu\big[W_0 = W_1 = 0\big] = P^\mu\big[W_0 = W_1 = W_{(0,1)} = 0\big] & \text{if } \; f(0) > 0,\ f(1) > 0,
\end{cases}
\end{aligned} \tag{2.7.23}$$

where $P^\mu[W_{(0,1)} = 0] < 1$ if and only if $\alpha > 1$ and $\langle \mu, v\rangle > 0$.

Indeed, by formula (2.5.2),

$$e^{-\langle \mu, \mathcal{U}_t f\rangle} = E^\mu\big[e^{-f(0)\mathcal{Y}_t(\{0\})} e^{-f(1)\mathcal{Y}_t(\{1\})} e^{-\langle \mathcal{Y}_t, 1_{(0,1)} f\rangle}\big]. \tag{2.7.24}$$

By (2.5.10) (i) and (2.5.11) (i) in Theorem 2.46,

$$\lim_{t\to\infty} e^{-f(r)\mathcal{Y}_t(\{r\})} = 1_{\{f(r)=0 \text{ or } W_r=0\}} \quad \text{a.s.} \quad (r = 0, 1). \tag{2.7.25}$$

Now, if $\langle \ell, f\rangle = 0$ for some $f \in B_+[0,1]$, then $e^{-\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle} = 1$ a.s. for each $t > 0$. To see this, note that by (2.6.20), $E^{\delta_x}[\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle] = e^{\alpha t}\langle \delta_x, S_t 1_{(0,1)}f\rangle = e^{\alpha t}E^x[1_{(0,1)}(\xi_t)f(\xi_t)]$ where ξ is the Wright-Fisher diffusion. Since the law of the Wright-Fisher diffusion at any time $t > 0$ (started in an arbitrary initial condition) on $(0,1)$ is absolutely continuous with respect to Lebesgue measure, we see that $E^{\delta_x}[\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle] = 0$ and hence $\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle = 0$ P^{δ_x}-a.s. (Actually, since $\mathcal{Y}$ is a one-dimensional superprocess, one can prove that $\mathcal{Y}_t$, restricted to $(0,1)$, for $t > 0$ is almost surely absolutely continuous with respect to Lebesgue measure.)

On the other hand, if $\langle \ell, f\rangle > 0$, then by formulas (2.5.10) (ii), (2.5.11) (ii), (2.5.13), and (2.5.15) in Theorem 2.46,

$$e^{-\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle} \xrightarrow{\mathrm{P}} 1_{\{W_{(0,1)}=0\}}. \tag{2.7.26}$$

Hence, for general $f \in B_+[0,1]$,

$$e^{-\langle \mathcal{Y}_t, 1_{(0,1)}f\rangle} \xrightarrow{\mathrm{P}} 1_{\{\langle \ell, f\rangle = 0 \text{ or } W_{(0,1)}=0\}}, \tag{2.7.27}$$

where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability. Inserting (2.7.25) and (2.7.27) into (2.7.24) we arrive at the first equality in (2.7.23). Using formula (2.5.12) and checking the eight possibilities for $f(0), f(1), \langle \ell, f\rangle$ to be zero or positive, we find the second equality in (2.7.23).

In particular, setting $\mu = \delta_x$ in (2.7.23) we see that $\mathcal{U}_t f$ converges in a bounded pointwise way to 0 or to one of the functions $p_{0,0}, \ldots, p_{1,1}$ from (2.5.17), where $p_{0,0} = 0$ if $\alpha \leq 1$ and $p_{0,0} > 0$ on $(0,1)$ otherwise. It follows from Lemma 2.59 that the convergence in (2.5.16) is in fact uniform.

The fact that $p_{l,r}(0) = l$ and $p_{l,r}(1) = r$ will follow from Proposition 2.48. The statements about smoothness of fixed points will be proved in Section 2.7.5 below. ∎

Proof of Proposition 2.48 By Proposition 2.47, for the functions $p_{0,0}, \ldots, p_{1,1}$ from (2.5.17),

$$\left.\begin{array}{l} p_{0,0}(x) = \lim_{t\to\infty} \mathcal{U}_t 1_{(0,1)}(x), \\ p_{1,0}(x) = \lim_{t\to\infty} \mathcal{U}_t 1_{\{0\}}(x) = \lim_{t\to\infty} \mathcal{U}_t 1_{[0,1)}(x), \\ p_{0,1}(x) = \lim_{t\to\infty} \mathcal{U}_t 1_{\{1\}}(x) = \lim_{t\to\infty} \mathcal{U}_t 1_{(0,1]}(x), \\ p_{1,1}(x) = \lim_{t\to\infty} \mathcal{U}_t 1 \end{array}\right\} \quad (x \in [0,1]). \tag{2.7.28}$$

Since by formula (2.5.4), for each Borel measurable $B \subset [0,1]$, $P^{\delta_x}[Y_t(B) > 0] = U_t 1_B = \mathcal{U}_t 1_B(x)$ $(t \geq 0,\ x \in [0,1])$, we can rewrite the expressions in the right-hand side of (2.7.28) as in (2.5.18). ∎

2.7.5 Smoothness of fixed points

In order to finish the proof of Proposition 2.47 we need to show that the functions $p_{0,0}, \ldots, p_{1,1}$ occurring there are twice continuously differentiable on $[0,1]$. We begin with the following.

Lemma 2.67 (Smoothness of fixed points) *If $p \in B_+[0,1]$ is a fixed point under $\mathcal{U}(\overline{A}, \alpha, \alpha)$, then $p \in \mathcal{D}(\overline{A})$ and $\overline{A}p + \alpha\, p(1-p) = 0$.*

Proof For any $t \geq 0$, Lemma 2.59 implies that $p = \mathcal{U}_t p \in \mathcal{C}_+[0,1]$. Moreover, since $u_t := p$ $(t \geq 0)$ is a mild solution of (2.5.6) (recall (2.6.31)),

$$p = S_t p + \int_0^t S_s\big(\alpha p(1-p)\big)\mathrm{d}s \qquad (t \geq 0). \tag{2.7.29}$$

Hence

$$\overline{A}p := \lim_{t\to 0} t^{-1}(S_t p - p) = -\lim_{t\to 0} t^{-1}\int_0^t S_s\big(\alpha\, p(1-p)\big)\mathrm{d}s = -\alpha\, p(1-p), \tag{2.7.30}$$

where the limit exists in $\mathcal{C}[0,1]$. ∎

In this one-dimensional situation, the domain of $\overline{A}$ is known explicitly. One has, see [EK86, Theorem 8.1.1]

$$\mathcal{D}(\overline{A}) = \Big\{ f \in \mathcal{C}[0,1] \cap \mathcal{C}^{(2)}(0,1) : \lim_{x\to r} \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2} f(x) = 0 \ \ (r = 0,1) \Big\}. \tag{2.7.31}$$

Here $\mathcal{C}[0,1] \cap \mathcal{C}^{(2)}(0,1)$ denotes the class of continuous real functions on $[0,1]$ that are twice continuously differentiable on $(0,1)$.

Proof of the smoothness of fixed points It suffices to show that $p_{0,0}$ and $p_{0,1}$ are twice continuously differentiable on $[0,1]$ and solve (2.5.8). The statement for $p_{1,0}$ then follows by symmetry, while for the constant functions 0 and $p_{1,1} = 1$ (see Proposition 2.48), the claim is obvious. Since $p_{0,0}, p_{0,1}$ are fixed points under $\mathcal{U}(\overline{A}, \alpha, \alpha)$, it follows from Lemma 2.67 and formula (2.7.31) that $p_{0,0}, p_{0,1}$ are continuous on $[0,1]$, twice continuously differentiable on $(0,1)$, and solve equation (2.5.8) on $(0,1)$. We are done if we can show that their first and second derivatives can be extended to continuous functions on $[0,1]$. (If f is twice continuously differentiable on $(0,1)$ and the limits $\lim_{x\to r} \frac{\partial}{\partial x} f(x)$ and $\lim_{x\to r} \frac{\partial^2}{\partial x^2} f(x)$ exists $(r = 0,1)$, then these limits coincide with the one-sided derivatives on the boundary. This follows, for example, from Corollary 6.3 in the appendix of [EK86].)

Proposition 2.48 shows that $p_{0,0}, p_{0,1} \leq 1$ and therefore, since they solve (2.5.8) on $(0,1)$, $p_{0,0}$ and $p_{0,1}$ are concave. Proposition 2.48 also shows that $p_{0,0}(0) = p_{0,0}(1) = 0$ and $p_{0,1}(0) = 0$, $p_{0,1}(1) = 1$. (See Figure 2.4 as an illustration.) Since $p_{0,0}$ is concave, $\frac{\partial}{\partial x} p_{0,0}(x)$ increases to a limit in $(-\infty, \infty]$ as $x \downarrow 0$. Lemma 2.61 implies that this limit is finite, and therefore $\frac{\partial}{\partial x} p_{0,0}(x)$ is continuous at $x = 0$. Since $p_{0,0}$ solves (2.5.8) on $(0,1)$,

$$\lim_{x\to 0} \tfrac{\partial^2}{\partial x^2} p_{0,0}(x) = -\lim_{x\to 0} \frac{2\alpha p_{0,0}(x)(1 - p_{0,0}(x))}{x(1-x)} = -2\alpha \tfrac{\partial}{\partial x} p_{0,0}(x)\big|_{x=0}, \tag{2.7.32}$$

which proves that $\frac{\partial^2}{\partial x^2} p_{0,0}(x)$ is continuous at $x = 0$. The same argument proves that $\frac{\partial}{\partial x} p_{0,0}(x)$ and $\frac{\partial^2}{\partial x^2} p_{0,0}(x)$ are continuous at $x = 1$, and that $\frac{\partial}{\partial x} p_{0,1}(x)$ and $\frac{\partial^2}{\partial x^2} p_{0,1}(x)$ are continuous at $x = 0$. Since $p_{0,1}$ is concave, $\frac{\partial}{\partial x} p_{0,1}(x)$ decreases to a limit in $[-\infty, \infty)$ as $x \uparrow 1$. Since $p_{0,1}(1) = 1$ and $p_{0,1} \leq 1$, $\frac{\partial}{\partial x} p_{0,1}(x)\big|_{x=1} \geq 0$. Since $p_{0,1}$ solves (2.5.8) on $(0,1)$ and $\frac{\partial}{\partial x}[p_{0,1}(x)(1 - p_{0,1}(x))]\big|_{x=1} = -\frac{\partial}{\partial x} p_{0,1}(x)\big|_{x=1}$,

$$\lim_{x\uparrow 1} \tfrac{\partial^2}{\partial x^2} p_{0,1}(x) = -\lim_{x\uparrow 1} \frac{2\alpha p_{0,1}(x)(1 - p_{0,1}(x))}{x(1-x)} = -2\alpha \tfrac{\partial}{\partial x} p_{0,1}(x)\big|_{x=1}, \tag{2.7.33}$$

which proves that $\frac{\partial}{\partial x}p_{0,1}(x)$ and $\frac{\partial^2}{\partial x^2}p_{0,1}(x)$ are continuous at $x=1$. ∎

2.8 The renormalization branching process: embedded particle systems

In this section we use embedded particle systems to prove Proposition 2.22. An essential ingredient in the proofs is Proposition 2.82 (a), which will be proved in the Section 2.9.

2.8.1 Weighting and Poissonization

Proof of Proposition 2.20 Obviously $q^h_k \in \mathcal{C}_+(E^h)$ for each $k=1,\ldots,n$. Since $h\in\mathcal{C}_+(E)$ and h is bounded, it is easy to see that the map $\mu\mapsto h\mu$ from $\mathcal{M}(E)$ into $\mathcal{M}(E^h)$ is continuous, and therefore the cluster mechanisms defined in (2.2.32) are continuous. Since

$$\mathcal{U}^h_k f(x) = \frac{q_k(x)}{h(x)}E\big[1-e^{-\langle h\mathcal{Z}_x,f\rangle}\big] = \frac{\mathcal{U}_k(hf)(x)}{h(x)} \qquad (x\in E^h,\ f\in B_+(E^h)), \tag{2.8.1}$$

formula (2.2.33) holds on E^h. To see that (2.2.33) holds on $E\backslash E^h$, note that by assumption $\mathcal{U}_k h \leq Kh$ for some $K<\infty$, so if $x\in E\backslash E^h$, then $\mathcal{U}_k h(x)=0$. By monotonicity also $\mathcal{U}_k(hf)(x)=0$, while $h\mathcal{U}^h_k f(x)=0$ by definition. Since $\sup_{x\in E^h}\mathcal{U}^h_k 1(x) = \sup_{x\in E^h}\frac{\mathcal{U}_k h(x)}{h(x)} \leq K<\infty$, the log-Laplace operators $\mathcal{U}^h_k$ satisfy (2.2.14). If $\mathcal{X}$ is started in an initial state $\mathcal{X}_0$, then the Poisson-cluster branching process $\mathcal{X}^h$ with log-Laplace operators $\mathcal{U}^h_1,\ldots,\mathcal{U}^h_n$ started in $\mathcal{X}^h_0 = h\mathcal{X}_0$ satisfies

$$\begin{aligned} E\big[e^{-\langle h\mathcal{X}_k,f\rangle}\big] &= E\big[e^{-\langle \mathcal{X}_0,\mathcal{U}_1\circ\cdots\circ\mathcal{U}_k(hf)\rangle}\big]\\ &= E\big[e^{-\langle \mathcal{X}_0,h\mathcal{U}^h_1\circ\cdots\circ\mathcal{U}^h_k(f)\rangle}\big] = E\big[e^{-\langle \mathcal{X}^h_k,f\rangle}\big] \qquad (f\in B_+(E^h)), \end{aligned} \tag{2.8.2}$$

which proves (2.2.34). ∎

Proof of Proposition 2.21 We start by noting that by (2.2.13),

$$\mathcal{U}_k f(x) = q(x)E\big[1-e^{-\langle \mathcal{Z}^k_x,f\rangle}\big] = q_k(x)P[\mathrm{Pois}(f\mathcal{Z}^k_x)\neq 0] \qquad (x\in E,\ f\in B_+(E)). \tag{2.8.3}$$

Into (2.2.35), we insert

$$\begin{aligned} &P\big[\mathrm{Pois}(h\mathcal{Z}^k_x)\in\cdot\,\big]\\ &\quad = P\big[\mathrm{Pois}(h\mathcal{Z}^k_x)\in\cdot\,\big|\,\mathrm{Pois}(h\mathcal{Z}^k_x)\neq 0\big]P[\mathrm{Pois}(h\mathcal{Z}^k_x)\neq 0] + \delta_0 P[\mathrm{Pois}(h\mathcal{Z}^k_x)=0]. \end{aligned} \tag{2.8.4}$$

Here and in similar formulas below, if in a conditional probability the symbol $\mathrm{Pois}(\,\cdot\,)$ occurs twice with the same argument, then it always refers to the same random variable (and not to independent Poisson point measures with the same intensity, for example). Using moreover (2.8.3) we can rewrite (2.2.35) as

$$Q^h_k(x,\,\cdot\,) = \frac{\mathcal{U}_k h(x)}{h(x)}P\big[\mathrm{Pois}(h\mathcal{Z}^k_x)\in\cdot\,\big|\,\mathrm{Pois}(h\mathcal{Z}^k_x)\neq 0\big] + \frac{h(x)-\mathcal{U}_k h(x)}{h(x)}\delta_0(\,\cdot\,). \tag{2.8.5}$$

In particular, since we are assuming that h is $\mathcal{U}_k$-subharmonic, this shows that $Q^h_k(x,\,\cdot\,)$ is a probability measure. Let X^h be the branching particle system with offspring mechanisms $Q^h_1,\ldots,Q^h_k$. Let $Z^{h,k}_x$ be random variables such that $\mathcal{L}(Z^{h,k}_x)=Q^h_k(x,\,\cdot\,)$. Then, by (2.2.29), (2.2.35), (2.2.31), and (2.8.3),

$$\begin{aligned}U^h_k f(x)&=P[\mathrm{Thin}_f(Z^{h,k}_x)\neq 0]=\frac{q_k(x)}{h(x)}P[\mathrm{Thin}_f(\mathrm{Pois}(h\mathcal{Z}^k_x))\neq 0]\\&=\frac{q_k(x)}{h(x)}P[\mathrm{Pois}(hf\mathcal{Z}^k_x)\neq 0]=\frac{1}{h(x)}\mathcal{U}_k(hf)(x)\qquad(x\in E^h).\end{aligned}\tag{2.8.6}$$

If $x\in E\backslash E^h$, then $\mathcal{U}_k(hf)(x)\leq\mathcal{U}_k(h)(x)\leq h(x)=0=:h\mathcal{U}^h(f)(x)$. This proves (2.2.36). To see that Q^h_k is a *continuous* offspring mechanism, by [Kal76, Theorem 4.2] it suffices to show that $x\mapsto\int Q^h_k(x,\mathrm{d}\nu)e^{-\langle\nu,g\rangle}$ is continuous for all bounded $g\in\mathcal{C}_+(E^h)$. Indeed, setting $f:=1-e^{-g}$, one has $\int Q^h_k(x,\mathrm{d}\nu)e^{-\langle\nu,g\rangle}=\int Q^h_k(x,\mathrm{d}\nu)(1-f)^\nu=1-\mathcal{U}^h_k f(x)=1-\mathcal{U}_k(hf)(x)/h(x)$ which is continuous on E^h by the continuity of q_k and $\mathcal{Q}_k$.

To see that also (2.2.37) holds, just note that by (2.2.30), (2.2.36), and (2.2.16),

$$\begin{aligned}&P^{\mathcal{L}(\mathrm{Pois}(h\mu))}[\mathrm{Thin}_f(X^h_n)=0]=P[\mathrm{Thin}_{U^h_1\circ\cdots\circ U^h_n f}(\mathrm{Pois}(h\mu))=0]\\&\quad=P[\mathrm{Pois}((hU^h_1\circ\cdots\circ U^h_n f)\mu)=0]=P[\mathrm{Pois}((\mathcal{U}_1\circ\cdots\circ\mathcal{U}_n(hf))\mu)=0]\\&\quad=P^\mu[\mathrm{Pois}(hf\mathcal{X}_n)=0]=P^\mu[\mathrm{Thin}_f(\mathrm{Pois}(h\mathcal{X}_n))=0].\end{aligned}\tag{2.8.7}$$

Here $P^{\mathcal{L}(\mathrm{Pois}(h\mu))}$ denotes the law of the process started with initiallaw $\mathcal{L}(\mathrm{Pois}(h\mu))$. Since this formula holds for all $f\in B_{[0,1]}(E^h)$, formula (2.2.37) follows. ∎

Remark 2.68 (Boundedness of h) Propositions 2.20 and 2.21 generalize to the case that h is unbounded, except that in this case the cluster mechanism in (2.2.32) and the offspring mechanism in (2.2.35) need in general not be continuous. Here, in order for (2.2.33) and (2.2.36) to be well-defined, one needs to extend the definition of $\mathcal{U}_k f$ to unbounded functions f, which can always be done unambiguously (see Lemma 2.53). ◊

2.8.2 Sub- and superharmonic functions

This section contains a number of pivotal calculations involving the log-Laplace operators $\mathcal{U}_\gamma$ from (2.2.20). In particular, we will prove that the functions $h_{1,1}$, $h_{0,0}$, and $h_{0,1}$ from Lemmas 2.23, 2.24, and 2.25, respectively, are $\mathcal{U}_\gamma$-superharmonic.

We start with an observation that holds for general log-Laplace operators.

Lemma 2.69 (Constant multiples) *Let $\mathcal{U}$ be a log-Laplace operator of the form (2.2.13) satisfying (2.2.14) and let $f\in B_+(E)$. Then $\mathcal{U}(rf)\leq r\mathcal{U}f$ for all $r\geq 1$, and $\mathcal{U}(rf)\geq r\mathcal{U}f$ for all $0\leq r\leq 1$. In particular, if f is $\mathcal{U}$-superharmonic then rf is $\mathcal{U}$-superharmonic for each $r\geq 1$, and if f is $\mathcal{U}$-subharmonic then rf is $\mathcal{U}$-superharmonic for each $0\leq r\leq 1$.*

Proof If $\mathcal{X}$ is a branching process and $\mathcal{U}$ is the log-Laplace operator of the transition law from $\mathcal{X}_0$ to $\mathcal{X}_1$ then, using Jensen's inequality, for all $r\geq 1$,

$$e^{-\langle\mu,\mathcal{U}(rf)\rangle}=E^\mu\big[e^{-\langle\mathcal{X}_1,rf\rangle}\big]=E^\mu\big[\big(e^{-\langle\mathcal{X}_1,f\rangle}\big)^r\big]\geq\big(E^\mu\big[e^{-\langle\mathcal{X}_1,f\rangle}\big]\big)^r=e^{-\langle\mu,r\mathcal{U}f\rangle}.\tag{2.8.8}$$

Since this holds for all $\mu \in \mathcal{M}(E)$, it follows that $\mathcal{U}(rf) \le r\mathcal{U}f$. The proof of the statements for $0 \le r \le 1$ is the same but with the inequality signs reversed. ∎

We next turn our attention to the functions $h_{1,1}$ and $h_{0,0}$.

Lemma 2.70 (The catalyzing function $h_{1,1}$) *One has*

$$\mathcal{U}_\gamma(rh_{1,1})(x) = \frac{1+\gamma}{\frac{1}{r}+\gamma} \qquad (\gamma, r > 0,\ x \in [0,1]). \tag{2.8.9}$$

In particular, $h_{1,1}$ is $\mathcal{U}_\gamma$-harmonic for each $\gamma > 0$.

Proof Recall (2.2.18)–(2.2.20). Let $\sigma_{1/r}$ be an exponentially distributed random variable with mean $1/r$, independent of τ_γ. Then

$$\mathcal{U}_\gamma(rh_{1,1})(x) = (\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} r\mathrm{d}t}\big] = (\tfrac{1}{\gamma}+1)P[\sigma_{1/r} < \tau_\gamma] = (\tfrac{1}{\gamma}+1)\frac{\gamma}{\frac{1}{r}+\gamma}, \tag{2.8.10}$$

which yields (2.8.9). ∎

Lemma 2.71 (The catalyzing function $h_{0,0}$) *One has $\mathcal{U}_\gamma(rh_{0,0}) \le rh_{0,0}$ for each $\gamma, r > 0$.*

Proof Let Γ^γ_x be the invariant law from Corollary 2.30. Then, for any $\gamma > 0$ and $f \in B_+[0,1]$,

$$\begin{aligned}\mathcal{U}_\gamma f(x) &= (\tfrac{1}{\gamma}+1)E\big[1 - e^{-\langle \mathcal{Z}^\gamma_x, f\rangle}\big] \le (\tfrac{1}{\gamma}+1)E[\langle \mathcal{Z}^\gamma_x, f\rangle] \\ &= (\tfrac{1}{\gamma}+1)E\Big[\int_0^{\tau_\gamma} f(\mathbf{y}^\gamma_x(-t/2))\,\mathrm{d}t\Big] = (1+\gamma)\langle \Gamma^\gamma_x, f\rangle \qquad (x \in [0,1]),\end{aligned} \tag{2.8.11}$$

where we have used that τ_γ is independent of $\mathbf{y}^\gamma_x$ and has mean γ. In particular, setting $f = rh_{0,0}$ and using (2.3.25) we find that $\mathcal{U}_\gamma(rh_{0,0}) \le rh_{0,0}$. ∎

The aim of the remainder of this section is to derive various bounds on $\mathcal{U}_\gamma f$ for $f \in \mathcal{H}_{0,1}$. We start with a formula for $\mathcal{U}_\gamma f$ that holds for general $[0,1]$-valued functions f.

Lemma 2.72 (Action of $\mathcal{U}_\gamma$ on $[0,1]$-valued functions) *Let $\mathbf{y}^\gamma_x$ be the stationary solution to (2.2.17) and let $\tau_{\gamma/2}$ be an independent exponentially distributed random variable with mean $\gamma/2$. Let $(\beta_i)_{i\ge1}$ be independent exponentially distributed random variables with mean $\frac{1}{2}$, independent of $\mathbf{y}^\gamma_x$ and $\tau_{\gamma/2}$, and let $\sigma_k := \sum_{i=1}^k \beta_i$ $(k \ge 0)$. Then*

$$1 - \mathcal{U}_\gamma f(x) = E\Big[\prod_{k\ge0:\,\sigma_k<\tau_\gamma} \big(1 - f(\mathbf{y}^\gamma_x(-\sigma_k))\big)\Big] \qquad (\gamma > 0,\ f \in B_{[0,1]}[0,1],\ x \in [0,1]). \tag{2.8.12}$$

Proof By Lemma 2.70, the constant function $h_{1,1}(x) := 1$ satisfies $\mathcal{U}_\gamma h_{1,1} = h_{1,1}$ for all $\gamma > 0$. Therefore, by Proposition 2.21, Poissonizing the Poisson-cluster branching process $\mathcal{X}$ with the density $h_{1,1}$ yields a branching particle system $X^{h_{1,1}} = (X^{h_{1,1}}_{-n}, \ldots, X^{h_{1,1}}_0)$ with generating operators $U^{h_{1,1}}_{\gamma_{n-1}}, \ldots, U^{h_{1,1}}_{\gamma_0}$, where

$$U^{h_{1,1}}_\gamma f = \mathcal{U}_\gamma f \qquad (f \in B_{[0,1]}[0,1],\ \gamma > 0). \tag{2.8.13}$$

By (2.2.29) and (2.8.5),

$$U_\gamma^{h_{1,1}} f(x) = 1 - E\big[(1-f)^{\mathrm{Pois}(\mathcal{Z}_x^\gamma)} \,\big|\, \mathrm{Pois}(\mathcal{Z}_x^\gamma) \neq 0\big] \quad (f \in B_{[0,1]}[0,1],\ x \in [0,1],\ \gamma > 0). \tag{2.8.14}$$

Therefore, (2.8.12) will follow provided that

$$P\big[\mathrm{Pois}(\mathcal{Z}_x^\gamma) \in \cdot \,\big|\, \mathrm{Pois}(\mathcal{Z}_x^\gamma) \neq 0\big] = \mathcal{L}\Big(\sum_{k \geq 0:\ \sigma_k < \tau_{\gamma/2}} \delta_{\mathbf{y}_x^\gamma(-\sigma_k)} \Big). \tag{2.8.15}$$

Indeed, it is not hard to see that

$$\mathrm{Pois}(\mathcal{Z}_x^\gamma) \stackrel{\mathcal{D}}{=} \sum_{k > 0:\ \sigma_k < \tau_{\gamma/2}} \delta_{\mathbf{y}_x^\gamma(-\sigma_k)}. \tag{2.8.16}$$

This follows from the facts that $\mathcal{Z}_x^\gamma = 2\int_0^{\tau_{\gamma/2}} \delta_{\mathbf{y}_x^\gamma(-s)} \mathrm{d}s$ and

$$\sum_{k > 0:\ \sigma_k < \tau_{\gamma/2}} \delta_{-\sigma_k} \stackrel{\mathcal{D}}{=} \mathrm{Pois}(2\,1_{(-\tau_{\gamma/2},0]}). \tag{2.8.17}$$

Conditioning $\mathrm{Pois}(2\,1_{(-\tau_{\gamma/2},0]})$ on being nonzero means conditioning on $\tau_{\gamma/2} > \sigma_1$. Since $\tau_{\gamma/2} - \sigma_1$, conditioned on being nonnegative, is exponentially distributed with mean $\gamma/2$, using the stationarity of $\mathbf{y}_x^\gamma$, we arrive at (2.8.15). ∎

The next lemma generalizes the duality (2.3.22) to mixed moments of the Wright-Fisher diffusion $\mathbf{y}$ at multiple times. We can interpret the left-hand side of (2.8.18) as the probability that $m_1, \ldots, m_n$ organisms sampled from the population at times $t_1, \ldots, t_n$ are all of the genetic type I.

Lemma 2.73 (Sampling at multiple times) *Fix $0 \leq t_1 < \cdots < t_n = t$ and nonnegative integers $m_1, \ldots, m_n$. Let $\mathbf{y}$ be the diffusion in (2.3.20). Then*

$$E^y\Big[\prod_{k=1}^n \mathbf{y}_{t_k}^{m_k}\Big] = E\big[y^{\phi_t} x^{\psi_t}\big], \tag{2.8.18}$$

where $(\phi_s, \psi_s)_{s \in [0,t]}$ is a Markov process in $\mathbb{N}^2$ started in $(\phi_0, \psi_0) = (m_n, 0)$, that jumps deterministically as

$$(\phi_s, \psi_s) \to (\phi_s + m_k, \psi_s) \quad \text{at time} \quad t - t_k \quad (k < n), \tag{2.8.19}$$

and between these deterministic times jumps with rates as in (2.3.21).

Proof Induction, with repeated application of (2.3.22). ∎

For any $m \geq 1$, we put

$$h_m(x) := 1 - (1-x)^m \qquad (x \in [0,1]). \tag{2.8.20}$$

The next lemma shows that we have particular good control on the action of $\mathcal{U}_\gamma$ on the functions h_m.

Lemma 2.74 (Action of $\mathcal{U}_\gamma$ on the functions h_m) *Let $m \geq 1$ and let τ_γ be an exponentially distributed random variable with mean γ. Conditional on τ_γ, let $(\phi'_t, \psi'_t)_{t\geq 0}$ be a Markov process in $\mathbb{N}^2$, started in $(\phi'_0, \psi'_0) = (m, 0)$ that jumps at time t as:*

$$\begin{array}{ll} (\phi'_t, \psi'_t) \to (\phi'_t - 1, \psi'_t) & \text{with rate } \phi'_t(\phi'_t - 1), \\ (\phi'_t, \psi'_t) \to (\phi'_t - 1, \psi'_t + 1) & \text{with rate } \frac{1}{\gamma}\phi'_t, \\ (\phi'_t, \psi'_t) \to (\phi'_t + m, \psi'_t) & \text{with rate } 1_{\{\tau_{\gamma/2} < t\}}. \end{array} \tag{2.8.21}$$

Then the limit $\lim_{t\to\infty} \psi'_t =: \psi'_\infty$ exists a.s., and

$$\mathcal{U}_\gamma h_m(x) = E^{(m,0)}\big[1 - (1-x)^{\psi'_\infty}\big] \qquad (m \geq 1,\ x \in [0,1]). \tag{2.8.22}$$

Proof Let $\mathbf{y}^\gamma_x$, $\tau_{\gamma/2}$, and $(\sigma_k)_{k\geq 0}$ be as in Lemma 2.72. Then, by (2.8.12),

$$\mathcal{U}_\gamma h_m(x) = 1 - E\Big[\prod_{k\geq 0:\ \sigma_k < \tau_{\gamma/2}} \big(1 - \mathbf{y}^\gamma_x(-\sigma_k)\big)^m\Big]. \tag{2.8.23}$$

Let $(\phi', \psi') = (\phi'_t, \psi'_t)_{t\geq 0}$ be a $\mathbb{N}^2$-valued process started in $(\phi'_0, \psi'_0) = (m, 0)$ such that conditioned on τ_γ and $(\sigma_k)_{k\geq 0}$, (ϕ', ψ') is a Markov process that jumps deterministically as

$$(\phi'_t, \psi'_t) \to (\phi'_t + m, \psi'_s) \quad \text{at time} \quad \sigma_k \quad (k \geq 1:\ \sigma_k < \tau_{\gamma/2}) \tag{2.8.24}$$

and between these times jumps with rates as in (2.3.21). Then $(\phi'_t, \psi'_t) \to (0, \psi'_\infty)$ as $t \to \infty$ a.s. for some $\mathbb{N}$-valued random variable ψ'_∞, and (2.8.22) follows from Lemma 2.73, using the symmetry $y \leftrightarrow 1 - y$. Since $\sigma_{k+1} - \sigma_k$ are independent exponentially distributed random variables with mean one, (ϕ', ψ') is the Markov process with jump rates as in (2.8.21). ∎

The next result is a simple application of Lemma 2.74.

Lemma 2.75 (The catalyzing function h_1) *The function $h_1(x) := x$ $(x \in [0,1])$ is $\mathcal{U}_\gamma$-subharmonic for each $\gamma > 0$.*

Proof Since $\psi'_\infty \geq 1$ a.s., one has $1 - (1-x)^{\psi'_\infty} \geq x$ a.s. $(x \in [0,1])$ in (2.8.22). In particular, setting $m = 1$ yields $\mathcal{U}_\gamma h_1 \geq h_1$. ∎

We now set out to prove that h_7, which is the function $h_{0,1}$ from Lemma 2.25, is $\mathcal{U}_\gamma$-superharmonic. In order to do so, we will derive upper bounds on the expectation of ψ'_∞. We derive two estimates: one that is good for small γ and one that is good for large γ.

In order to avoid tedious formal arguments, it will be convenient to recall the interpretation of the process (ϕ', ψ') and Lemma 2.73. Recall from the discussion following (2.3.22) that $(\mathbf{y}^\gamma_x(t))_{t\in\mathbb{R}}$ describes the equilibrium frequency of genetic type I as a function of time in a population that is in genetic exchange with an infinite reservoir. From this population we sample at times $-\sigma_k$ $(k \geq 0,\ \sigma_k < \tau_{\gamma/2})$ each time m individuals, and ask for the probability that they are not all of the genetic type II. In order to find this probability, we follow the ancestors of the sampled individuals back in time. Then ϕ'_t and ψ'_t are the number of ancestors that lived at time $-t$ in the population and the reservoir, respectively, and $E[1 - (1-x)^{\psi'_\infty}]$ is the probability that at least one ancestor is of type I.

Lemma 2.76 (Bound for small γ) *For each $\gamma \in (0,\infty)$ and $m \geq 1$,*

$$\frac{1}{m}E^{(m,0)}[\psi'_\infty] \leq \frac{1}{m}\sum_{i=0}^{m-1}\frac{1+\gamma}{1+i\gamma} =: \chi_m(\gamma). \tag{2.8.25}$$

The function χ_m is concave and satisfies $\chi_m(0) = 1$ for each $m \geq 1$.

Proof Note that

$$E\big[\big|\{k \geq 0 :\ \sigma_k < \tau_{\gamma/2}\}\big|\big] = 1+\gamma. \tag{2.8.26}$$

We can estimate (ϕ',ψ') from above by a process where ancestors from individuals sampled at different times cannot coalesce. Therefore,

$$E^{(m,0)}[\psi'_\infty] \leq (1+\gamma)E^{(m,0)}[\psi_\infty], \tag{2.8.27}$$

where (ϕ,ψ) is the Markov process in (2.3.21). Note that if (ϕ,ψ) is in the state $(m+1,0)$, then the next jump is to $(m,1)$ with probability

$$\frac{\frac{1}{\gamma}(m+1)}{\frac{1}{\gamma}(m+1)+m(m+1)} = \frac{1}{1+m\gamma} \tag{2.8.28}$$

and to $(m,0)$ with one minus this probability. Therefore,

$$\begin{aligned} E^{(m+1,0)}[\psi_\infty] &= \frac{1}{1+m\gamma}E^{(m,1)}[\psi_\infty] + \Big(1-\frac{1}{1+m\gamma}\Big)E^{(m,0)}[\psi_\infty] \\ &= \frac{1}{1+m\gamma}\Big(E^{(m,0)}[\psi_\infty]+1\Big) + \Big(1-\frac{1}{1+m\gamma}\Big)E^{(m,0)}[\psi_\infty] \\ &= E^{(m,0)}[\psi_\infty] + \frac{1}{1+m\gamma}. \end{aligned} \tag{2.8.29}$$

By induction, it follows that

$$E^{(m,0)}[\psi_\infty] = \sum_{i=0}^{m-1}\frac{1}{1+i\gamma}. \tag{2.8.30}$$

Inserting this into (2.8.27) we arrive at (2.8.25). Finally, since

$$\frac{\partial^2}{\partial\gamma^2}\frac{1+\gamma}{1+i\gamma} = \frac{2i(i-1)}{(1+i\gamma)^3} \geq 0 \qquad (i \geq 0,\ \gamma \geq 0), \tag{2.8.31}$$

the function χ_m is convex. ∎

Lemma 2.77 (Bound for large γ) *For each $\gamma \in (0,\infty)$ and $m \geq 1$,*

$$E^{(m,0)}[\psi'_\infty] \leq (\tfrac{1}{\gamma}+1)\sum_{k=1}^{m}\frac{1}{k} + \frac{3}{2}. \tag{2.8.32}$$

Proof We start by observing that $\frac{\partial}{\partial t}E[\psi'_t] = \frac{1}{\gamma}E[\phi'_t]$, and therefore

$$E[\psi'_\infty] = \tfrac{1}{\gamma}\int_0^\infty E[\phi'_t]\mathrm{d}t. \tag{2.8.33}$$

Unlike in the proof of the last lemma, this time we cannot fully ignore the coalescence of ancestors sampled at different times. In order to deal with this we use a trick: at time zero we introduce an extra ancestor that can only jump to the reservoir when $t \geq \tau_\gamma$ and there are no other ancestors left in the population. We further assume that all other ancestors do not jump to the reservoir on their own. Let ξ_t be one as long as this extra ancestor is in the population and zero otherwise, and let ϕ''_t be the number of other ancestors in the population according to these new rules. Then we have at a Markov process (ξ, ϕ'') started in $(\xi_0, \phi''_0) = (1, m)$ that jumps as:

$$\begin{array}{lll} (\xi_t, \phi''_t) \to (\xi_t, \phi''_t - 1) & \text{with rate } (\phi''_t + 1)\phi''_t, \\ (\xi_t, \phi''_t) \to (\xi_t, \phi''_t + m) & \text{with rate } 1_{\{\tau_{\gamma/2} < t\}}, \\ (\xi_t, \phi''_t) \to (\xi_t - 1, \phi''_t) & \text{with rate } \frac{1}{\gamma}1_{\{\tau_{\gamma/2} \geq t\}}1_{\{\phi''_t = 0\}}. \end{array} \tag{2.8.34}$$

It is not hard to show that (ξ, ϕ'') and ϕ' can be coupled such that $\xi_t + \phi''_t \geq \phi'_t$ for all $t \geq 0$. We now simplify even further and ignore all coalescence between ancestors belonging to the process ϕ'' that are introduced at different times. Let $\phi^{(k)}_t$ be the number of ancestors in the population that were introduced at the time σ_k $(k \geq 0)$. Thus, for $t < \sigma_k$ one has $\phi^{(k)}_t = 0$, for $t = \sigma_k$ one has $\phi^{(k)}_t = m$, while for $t > \sigma_k$, the process $\phi^{(k)}_t$ jumps from n to $n-1$ with rate $(n+1)n$. Then it is not hard to see that, for an appropriate coupling, $\phi''_t \leq \sum_{k\geq 0:\sigma_k < \tau_{\gamma/2}} \phi^{(k)}_t$ for all $t \geq 0$. We let ξ' be a process such that $\xi'_0 = 1$ and ξ'_t jumps to zero with rate

$$\frac{1}{\gamma}1_{\{\tau_{\gamma/2} \geq t\}} \prod_{k\geq 0:\sigma_k < \tau_{\gamma/2}} 1_{\{\phi^{(k)}_t = 0\}}. \tag{2.8.35}$$

Then for an appropriate coupling $\xi'_t \geq \xi_t$ $(t \geq 0)$. Thus, we can estimate

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \leq \int_0^\infty E[\xi'_t]\mathrm{d}t + \int_0^\infty E\Big[\sum_{k\geq 0:\sigma_k < \tau_{\gamma/2}} \phi^{(k)}_t\Big]\mathrm{d}t. \tag{2.8.36}$$

Set $\rho := \inf\{t \geq \tau_{\gamma/2} : \phi^{(k)}_t = 0 \ \forall k \geq 0 \text{ with } \sigma_k < \tau_{\gamma/2}\}$ and $\pi := \inf\{t \geq 0 : \xi'_t = 0\}$. Then

$$\int_0^\infty E[\xi'_t]\mathrm{d}t = E[\tau_{\gamma/2}] + E[\rho - \tau_{\gamma/2}] + E[\pi - \rho] = \frac{3}{2}\gamma + E[\rho - \tau_{\gamma/2}]. \tag{2.8.37}$$

Since

$$\begin{aligned} E[\rho - \tau_{\gamma/2}] &\leq \int_0^\infty E\Big[1_{\{\sum_{k\geq 0:\sigma_k < \tau_{\gamma/2}} \phi^{(k)}_t \neq 0\}}\Big]\mathrm{d}t \\ &\leq \int_0^\infty E\Big[\sum_{k\geq 0:\sigma_k < \tau_{\gamma/2}} 1_{\{\phi^{(k)}_t \neq 0\}}\Big]\mathrm{d}t, \end{aligned} \tag{2.8.38}$$

using moreover (2.8.36) and (2.8.37), we can estimate

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \le \frac{3}{2}\gamma + \int_0^\infty E\Big[\sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} (\phi_t^{(k)} + 1_{\{\phi_t^{(k)}\neq 0\}})\Big]\mathrm{d}t. \tag{2.8.39}$$

Since $E\big[\big|\{k \ge 0:\ \sigma_k < \tau_{\gamma/2}\}\big|\big] = 1+\gamma$, we obtain

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \le \frac{3}{2}\gamma + (1+\gamma)\int_0^\infty E[\phi_t^{(0)} + 1_{\{\phi_t^{(0)}\neq 0\}}]\mathrm{d}t. \tag{2.8.40}$$

Since $\phi_t^{(0)}$ jumps from n to $n-1$ with rate $(n+1)n$, the expected total time that $\phi_t^{(0)} = n$ equals $1/((n+1)n)$, and therefore

$$\int_0^\infty E[\phi_t^{(0)} + 1_{\{\phi_t^{(0)}\neq 0\}}]\mathrm{d}t = \sum_{n=1}^m \frac{1}{(n+1)n}(n + 1_{\{n\neq 0\}}) = \sum_{n=1}^m \frac{1}{n}. \tag{2.8.41}$$

Inserting this into (2.8.40), using (2.8.33), we arrive at (2.8.32). ∎

Lemma 2.78 (The catalyzing function $h_{0,1}$) *One has $\mathcal{U}_\gamma(h_{0,1}) \le h_{0,1}$ for each $\gamma > 0$. Moreover, for each $r > 1$ and $\gamma > 0$,*

$$\sup_{x\in(0,1]} \frac{\mathcal{U}_\gamma(rh_{0,1})(x)}{rh_{0,1}(x)} < 1. \tag{2.8.42}$$

Proof Recall that $h_{0,1}(x) = h_7(x) = 1-(1-x)^7$ $(x \in [0,1])$. We will show that

$$E^{(7,0)}[\psi'_\infty] < 7 \tag{2.8.43}$$

for each $\gamma \in (0,\infty)$. The function $\chi_m(\gamma)$ from Lemma 2.76 satisfies

$$\chi_m(1) = \frac{1}{m}\sum_{n=1}^m \frac{2}{n} < 1 \qquad (m \ge 5). \tag{2.8.44}$$

Since $\chi_m(\gamma)$ is concave in γ and satisfies $\chi_m(0) = 1$, it follows that $\chi_m(\gamma) < 1$ for all $0 < \gamma \le 1$ and $m \ge 5$. By Lemma 2.77, for all $\gamma \ge 1$,

$$E^{(m,0)}[\psi'_\infty] \le 2\sum_{k=1}^m \frac{1}{k} + \frac{3}{2} < m \qquad (m \ge 7). \tag{2.8.45}$$

Therefore, if $m \ge 7$, then $m' := E^{(m,0)}[\psi'_\infty] < m$. It follows by (2.8.22) and Jensen's inequality applied to the concave function $z \mapsto 1-(1-x)^z$ that

$$\mathcal{U}_\gamma h_m(x) \le 1-(1-x)^{E^{(m,0)}[\psi'_\infty]} = 1-(1-x)^{m'} \le h_m(x) \qquad (x \in [0,1],\ \gamma > 0). \tag{2.8.46}$$

This shows that h_m is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. By Lemma 2.69, for each $r > 1$,

$$\frac{\mathcal{U}_\gamma(rh_m)(x)}{rh_m(x)} \le \frac{r\mathcal{U}_\gamma(h_m)(x)}{rh_m(x)} \le \frac{1-(1-x)^{m'}}{1-(1-x)^m} \qquad (x \in (0,1]). \tag{2.8.47}$$

By Lemma 2.70 and the monotonicity of $\mathcal{U}_\gamma$,

$$\frac{\mathcal{U}_\gamma(rh_m)(x)}{rh_m(x)} \leq \frac{\mathcal{U}_\gamma(r)(x)}{rh_m(x)} \leq \frac{1+\gamma}{1+r\gamma}\frac{1}{(1-(1-x)^m)} \qquad (x \in (0,1]). \tag{2.8.48}$$

Since the right-hand side of (2.8.47) is smaller than 1 for $x \in (0,1)$ and tends to $m'/m < 1$ as $x \to 0$, since the right-hand side of (2.8.48) is smaller than 1 for x in an open neighborhood of 1, and since both bounds are continuous, (2.8.42) follows. ∎

2.8.3 Extinction versus unbounded growth

In this section we show that Lemmas 2.23–2.25 are equivalent to Proposition 2.26. (This follows from the equivalence of conditions (i) and (ii) in Lemma 2.79 below.) We moreover prove Lemmas 2.23 and 2.25 and prepare for the proof of Lemma 2.24. We start with some general facts about log-Laplace operators and branching processes.

For the next lemma, let E be a separable, locally compact, metrizable space. For $n \geq 0$, let $q_n \in \mathcal{C}_+(E)$ be continuous weight functions, let $\mathcal{Q}_n$ be continuous cluster mechanisms on E, and assume that the associated log-Laplace operators $\mathcal{U}_n$ defined in (2.2.13) satisfy (2.2.14). Assume that $0 \neq h \in \mathcal{C}_+(E)$ is bounded and $\mathcal{U}_n$-superharmonic for all n, let $E^h := \{x \in E : h(x) > 0\}$, and define generating operators $U^h_n : B_{[0,1]}(E^h) \to B_{[0,1]}(E)$ as in (2.2.36). For each $n \geq 0$, let $(\mathcal{X}^{(n)}_0, \mathcal{X}^{(n)}_1)$ be a one-step Poisson cluster branching process with log-Laplace operator $\mathcal{U}_n$, and let $(X^{(n),h}_0, X^{(n),h}_1)$ be the one-step branching particle system with generating operator U^h_n. (In a typical application of this lemma, the operators $\mathcal{U}_n$ will be iterates of other log-Laplace operators, and $\mathcal{X}^{(n)}_0, \mathcal{X}^{(n)}_1$ will be the initial and final state, respectively, of a Poisson cluster branching process with many time steps.)

Lemma 2.79 (Extinction versus unbounded growth) *Assume that $\rho \in \mathcal{C}_{[0,1]}(E^h)$ and put*

$$p(x) := \begin{cases} h(x)\rho(x) & \text{if } x \in E^h, \\ 0 & \text{if } x \in E \backslash E^h. \end{cases} \tag{2.8.49}$$

Then the following statements are equivalent:

$$\text{(i)} \quad P^{\delta_x}\big[|X^{(n),h}_1| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \rho(x)\delta_\infty + (1-\rho(x))\delta_0 \quad \textit{locally uniformly for } x \in E^h,$$

$$\text{(ii)} \quad P^{\delta_x}\big[\langle \mathcal{X}^{(n)}_1, h\rangle \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} e^{-p(x)}\delta_0 + \big(1-e^{-p(x)}\big)\delta_\infty \quad \textit{locally uniformly for } x \in E,$$

$$\text{(iii)} \quad \mathcal{U}_n(\lambda h)(x) \underset{n\to\infty}{\longrightarrow} p(x) \quad \textit{locally uniformly for } x \in E \quad \forall \lambda > 0,$$

$$\text{(iv)} \quad \exists 0 < \lambda_1 < \lambda_2 < \infty: \quad \mathcal{U}_n(\lambda_i h)(x) \underset{n\to\infty}{\longrightarrow} p(x) \quad \textit{locally uniformly for } x \in E \quad (i = 1,2).$$

Proof of Lemma 2.79 It is not hard to see that (i) is equivalent to

$$P^{\delta_x}[\mathrm{Thin}_\lambda(X_1^{(n),h}) \neq 0] \underset{n\to\infty}{\longrightarrow} \rho(x) \tag{2.8.50}$$

locally uniformly for $x \in E^h$, for all $0 < \lambda \leq 1$. It follows from (2.2.30) and (2.2.36) that $h(x)P^{\delta_x}[\mathrm{Thin}_\lambda(X_1^{(n),h}) \neq 0] = hU^h(\lambda)(x) = \mathcal{U}(\lambda h)(x)$ $(x \in E)$, so (i) is equivalent to

$$\text{(i)}' \quad \mathcal{U}_n(\lambda h)(x) \underset{n\to\infty}{\longrightarrow} p(x)$$
$$\text{locally uniformly for } x \in E \quad \forall 0 < \lambda \leq 1.$$

By (2.2.15), condition (ii) implies that

$$e^{-\mathcal{U}_n(\lambda h)(x)} = E^{\delta_x}\big[e^{-\lambda\langle \mathcal{X}_1, h\rangle}\big] \underset{n\to\infty}{\longrightarrow} e^{-p(x)} \tag{2.8.51}$$

locally uniformly for $x \in E$ for all $\lambda > 0$, and therefore (ii) implies (iii). Obviously (iii)$\Rightarrow$ (i)$'$ $\Rightarrow$(iv) so we are done if we show that (iv)$\Rightarrow$(ii). Indeed, (iv) implies that

$$E^{\delta_x}\big[e^{-\lambda_1\langle \mathcal{X}_1^{(n)}, h\rangle} - e^{-\lambda_2\langle \mathcal{X}_1^{(n)}, h\rangle}\big] \underset{n\to\infty}{\longrightarrow} 0 \tag{2.8.52}$$

locally uniformly for $x \in E$, which shows that

$$P^{\delta_x}\big[c < \langle \mathcal{X}_1^{(n)}, h\rangle < C\big] \underset{n\to\infty}{\longrightarrow} 0 \tag{2.8.53}$$

for all $0 < c < C < \infty$. Using (iv) once more we arrive at (ii). ∎

Our next lemma gives sufficient conditions for the n-th iterates of a single log-Laplace operator $\mathcal{U}$ to satisfy the equivalent conditions of Lemma 2.79. Let E (again) be separable, locally compact, and metrizable. Let $q \in \mathcal{C}_+(E)$ be a weight function, $\mathcal{Q}$ a continuous cluster mechanism on E, and assume that the associated log-Laplace operator $\mathcal{U}$ defined in (2.2.13) satisfies (2.2.14). Let $\mathcal{X} = (\mathcal{X}_0, \mathcal{X}_1, \ldots)$ be the Poisson-cluster branching process with log-Laplace operator $\mathcal{U}$ in each step, let $0 \neq h \in \mathcal{C}_+(E)$ be bounded and $\mathcal{U}$-superharmonic, and let $X^h = (X_0^h, X_1^h, \ldots)$ denote the branching particle system on E^h obtained from $\mathcal{X}$ by Poissonization with a $\mathcal{U}$-superharmonic function h, in the sense of Proposition 2.21.

Lemma 2.80 (Sufficient condition for extinction versus unbounded growth) *Assume that*

$$\sup_{x\in E^h} \frac{\mathcal{U}h(x)}{h(x)} < 1. \tag{2.8.54}$$

Then the process X^h started in any initial law $\mathcal{L}(X_0^h) \in \mathcal{M}_1(E^h)$ satisfies

$$\lim_{k\to\infty} |X_k^h| = \infty \quad \textit{or} \quad \exists k \geq 0 \text{ s.t. } X_k^h = 0 \qquad \textit{a.s.} \tag{2.8.55}$$

Moreover, if the function $\rho : E^h \to [0,1]$ defined by

$$\rho(x) := P^{\delta_x}[X_n^h \neq 0 \quad \forall n \geq 0] \qquad (x \in E^h) \tag{2.8.56}$$

satisfies $\inf_{x\in E^h} \rho(x) > 0$, then ρ is continuous.

Proof of Lemma 2.80 Let $\mathcal{A}$ denote the tail event $\mathcal{A} = \{X^h_n \neq 0 \ \forall n \geq 0\}$ and let $(\mathcal{F}_k)_{k\geq 0}$ be the filtration generated by X^h. Then, by the Markov property and continuity of the conditional expectation with respect to increasing limits of σ-fields (see Complement 10(b) from [Loe63, Section 29] or [Loe78, Section 32])

$$P[X^h_n \neq 0 \ \forall n \geq 0 | X_k] = P(\mathcal{A}|\mathcal{F}_k) \underset{k\to\infty}{\longrightarrow} 1_{\mathcal{A}} \qquad \text{a.s.} \tag{2.8.57}$$

In particular, this implies that a.s. on the event $\mathcal{A}$ one must have $P[X^h_{k+1} = 0|X^h_k] \to 0$ a.s. By (2.2.30) and (2.2.36), $P^{\delta_x}[X^h_1 \neq 0] = U^h 1(x) = (\mathcal{U}h(x))/h(x)$, which is uniformly bounded away from one by (2.8.54). Therefore, $P[X^h_{k+1} = 0|X^h_k] \to 0$ a.s. on $\mathcal{A}$ is only possible if the number of particles tends to infinity.

The continuity of ρ can be proved by a straightforward adaptation of the proof of [FS04, Proposition 5 (d)] to the present setting with discrete time and noncompact space E. An essential ingredient in the proof, apart from (2.8.54), is the fact that the map $\nu \mapsto P^\nu[X^h_n \in \cdot\,]$ from $\mathcal{N}(E)$ to $\mathcal{M}_1(\mathcal{N}(E))$ is continuous, which follows from the continuity of Q^h. ∎

We now turn our attention more specifically to the renormalization branching process $\mathcal{X}$. In the remainder of this section, $(\gamma_k)_{k\geq 0}$ is a sequence of positive constants such that $\sum_n \gamma_n = \infty$ and $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$, and $\mathcal{X} = (\mathcal{X}_{-n}, \ldots, \mathcal{X}_0)$ is the Poisson cluster branching process on $[0,1]$ defined in Section 2.2.4. We put $\mathcal{U}^{(n)} := \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}$. If $0 \neq h \in \mathcal{C}[0,1]$ is $\mathcal{U}_{\gamma_k}$-superharmonic for all $k \geq 0$, then $\mathcal{X}^h$ and X^h denote the branching process and the branching particle system on $\{x \in [0,1] : h(x) > 0\}$ obtained from $\mathcal{X}$ by weighting and Poissonizing with h in the sense of Propositions 2.20 and 2.21, respectively.

Proof of Lemma 2.23 By induction, it follows from Lemma 2.70 that

$$\mathcal{U}^{(n)}(\lambda h_{1,1}) = \frac{\prod_{k=0}^{n-1}(1+\gamma_k)}{\prod_{k=0}^{n-1}(1+\gamma_k) - 1 + \frac{1}{\lambda}} \qquad (\lambda > 0). \tag{2.8.58}$$

It is not hard to see (compare the footnote at (2.1.42)) that

$$\prod_{k=0}^{\infty}(1+\gamma_k) = \infty \quad \text{if and only if} \quad \sum_{k=0}^{\infty}\gamma_k = \infty. \tag{2.8.59}$$

Therefore, since we are assuming that $\sum_n \gamma_n = \infty$,

$$\mathcal{U}^{(n)}(\lambda h_{1,1}) \underset{n\to\infty}{\longrightarrow} h_{1,1}, \tag{2.8.60}$$

uniformly on $[0,1]$ for all $\lambda > 0$. The result now follows from Lemma 2.79 (with $h = h_{1,1}$ and $\rho(x) = 1$ $(x \in [0,1])$). ∎

Remark 2.81 (Conditions on $(\gamma_n)_{n\geq 0}$) Our proof of Lemma 2.23 does not use that $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$. On the other hand, the proof shows that $\sum_n \gamma_n = \infty$ is a necessary condition for (2.2.40). ◊

We do not know if the assumption that $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$ is needed in Lemma 2.24. We guess that it can be dropped, but it will greatly simplify proofs to have it around.

We will show that in order to prove Lemmas 2.24 and 2.25, it suffices to prove their analogues for embedded particle systems in the time-homogeneous processes $\mathcal{Y}^{\gamma^*}$ ($\gamma^* \in [0,\infty)$). More precisely, we will derive Lemmas 2.24 and 2.25 from the following two results. Below, $(\mathcal{U}^0_t)_{t\geq 0}$ is the log-Laplace semigroup of the super-Wright-Fisher diffusion $\mathcal{Y}^0$, defined in (2.2.26). The functions $p^*_{0,1,\gamma^*}$ ($\gamma^* \in [0,\infty)$) are defined in (2.2.45).

Proposition 2.82 (Time-homogeneous embedded particle system with $h_{0,0}$)
(a) *For any $\gamma^* > 0$, one has $(\mathcal{U}_{\gamma^*})^n h_{0,0} \underset{n\to\infty}{\longrightarrow} 0$ uniformly on $[0,1]$.*

(b) *One has $\mathcal{U}^0_t h_{0,0} \underset{t\to\infty}{\longrightarrow} 0$ uniformly on $[0,1]$.*

Proposition 2.83 (Time-homogeneous embedded particle system with $h_{0,1}$)
(a) *For any $\gamma^* > 0$, one has $(\mathcal{U}_{\gamma^*})^n(\lambda h_{0,1}) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}$ uniformly on $[0,1]$, for all $\lambda > 0$.*

(b) *One has $\mathcal{U}^0_t(\lambda h_{0,1}) \underset{t\to\infty}{\longrightarrow} p^*_{0,1,0}$ uniformly on $[0,1]$, for all $\lambda > 0$.*

Propositions 2.82 (b) and 2.83 (b) follow from Proposition 2.47. Proposition 2.82 (a) will be proved in Section 2.9.2.

Proof of Proposition 2.83 (a) By formula (2.8.42) from Lemma 2.78, for each $r > 1$ the function $rh_{0,1}$ satisfies condition (2.8.54) from Lemma 2.80. Set $\rho(x) := P^{\delta_x}[Y_n^{\gamma^*,rh_{0,1}} \neq 0\ \forall n]$. Then, by (2.2.30) and (2.2.36),

$$\begin{aligned}\rho(x) &= \lim_{n\to\infty} P^{\delta_x}[Y_n^{\gamma^*,rh_{0,1}} \neq 0] = \lim_{n\to\infty} (U_{\gamma^*}^{rh_{0,1}})^n 1(x)\\ &= \lim_{n\to\infty} \frac{(\mathcal{U}_{\gamma^*})^n(rh_{0,1})(x)}{rh_{0,1}(x)} \geq \frac{h_1(x)}{rh_{0,1}(x)} \qquad (x \in (0,1]),\end{aligned} \tag{2.8.61}$$

where $h_1(x) = x$ ($x \in [0,1]$) is the $\mathcal{U}_{\gamma^*}$-subharmonic function from Lemma 2.75. It follows that $\inf_{x\in(0,1]} \rho(x) > 0$ and therefore, by Lemma 2.80, ρ is continuous in x.

By Lemma 2.80, we see that the Poissonized particle system $X^{rh_{0,1}}$ exhibits extinction versus unbounded growth in the sense of Lemma 2.79, which implies the statement in Proposition 2.83 (a). ∎

We now show that Propositions 2.82 and 2.83 imply Lemmas 2.24 and 2.25, respectively.

Proof of Lemma 2.24 We start with the proof that the embedded particle system $X^{h_{0,0}}$ is critical. For any $f \in B_+[0,1]$ and $k \geq 1$, we have, by Poissonization (Proposition 2.21) and the definition of $\mathcal{X}$,

$$\begin{aligned}&h_{0,0}(x)E^{-k,\delta_x}[\langle X^{h_{0,0}}_{-k+1}, f\rangle] = E^{-k,\mathcal{L}(\mathrm{Pois}(h_{0,0}\delta_x))}[\langle X^{h_{0,0}}_{-k+1}, f\rangle] = E^{-k,\delta_x}[\langle \mathrm{Pois}(h_{0,0}\mathcal{X}_{-k+1}), f\rangle]\\ &\quad = E^{-k,\delta_x}[\langle \mathcal{X}_{-k+1}, h_{0,0}f\rangle] = (\tfrac{1}{\gamma}+1)E[\langle \mathcal{Z}^\gamma_x, h_{0,0}f\rangle] = (\tfrac{1}{\gamma}+1)\langle \Gamma^{\gamma_{k-1}}_x, h_{0,0}f\rangle,\end{aligned} \tag{2.8.62}$$

where Γ^γ_x is the invariant law of $\mathbf{y}^\gamma_x$ from Corollary 2.30. In particular, setting $f = 1$ gives $h_{0,0}(x)E^{-k,\delta_x}[|X^{h_{0,0}}_{-k+1}|] = h_{0,0}(x)$ by (2.3.25).

To prove (2.2.41), by Lemma 2.79 it suffices to show that

$$\mathcal{U}^{(n)}(\lambda h_{0,0}) \underset{n\to\infty}{\longrightarrow} 0 \tag{2.8.63}$$

uniformly on $[0,1]$ for all $0<\lambda\leq 1$. We first treat the case $\gamma^*>0$. Then, by Theorem 2.19 (a), for each fixed $l\geq 1$ and $f\in\mathcal{C}_+[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{n-l}}f \underset{n\to\infty}{\longrightarrow} (\mathcal{U}_{\gamma^*})^l f \tag{2.8.64}$$

uniformly on $[0,1]$. Therefore, by a diagonal argument, we can find $l(n)\to\infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)}h_{0,0}-\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{n-l(n)}}h_{0,0}\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{2.8.65}$$

Using the fact that the function $h_{0,0}$ is $\mathcal{U}_\gamma$-superharmonic for each $\gamma>0$ and the monotonicity of the operators $\mathcal{U}_\gamma$, we derive from Proposition 2.82 (a) that

$$\mathcal{U}^{(n)}(\lambda h_{0,0})\leq\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{n-l(n)}}h_{0,0} \underset{n\to\infty}{\longrightarrow} 0 \tag{2.8.66}$$

uniformly on $[0,1]$ for all $0<\lambda\leq 1$. This proves (2.8.63) in the case $\gamma^*>0$.

The proof in the case $\gamma^*=0$ is similar. In this case, by Theorem 2.19 (b), for each fixed $t>0$ and $f\in\mathcal{C}_+[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{k_n(t)}}f(x_n) \underset{n\to\infty}{\longrightarrow} \mathcal{U}^0_t f(x) \qquad \forall x_n\to x\in[0,1], \tag{2.8.67}$$

which shows that $\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{k_n(t)}}f$ converges to $\mathcal{U}^0_t f$ uniformly on $[0,1]$. By a diagonal argument, we can find $t(n)\to\infty$ such that

$$\big\|\mathcal{U}^0_t(h_{0,0})-\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{k_n(t(n))}}(h_{0,0})\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0, \tag{2.8.68}$$

and the proof proceeds in the same way as before. ∎

Proof of Lemma 2.25 By Lemma 2.79 and the monotonicity of the operators $\mathcal{U}_\gamma$ it suffices to show that

$$\begin{array}{rl}\text{(i)} & \limsup\limits_{n\to\infty}\mathcal{U}^{(n)}(h_{0,1})\leq p^*_{0,1,\gamma^*},\\ \text{(ii)} & \liminf\limits_{n\to\infty}\mathcal{U}^{(n)}(\tfrac{1}{2}h_{0,1})\geq p^*_{0,1,\gamma^*},\end{array} \tag{2.8.69}$$

uniformly on $[0,1]$. We first consider the case $\gamma^*>0$. By (2.8.64) and a diagonal argument, we can find $l(n)\to\infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)}h_{0,1}-\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{n-l(n)}}h_{0,1}\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{2.8.70}$$

Therefore, by Proposition 2.83 (a), the fact that $h_{0,1}$ is $\mathcal{U}_{\gamma_k}$-superharmonic for each $k\geq 0$, and the monotonicity of the operators $\mathcal{U}_\gamma$, we find that

$$\mathcal{U}^{(n)}h_{0,1}\leq\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_{n-l(n)}}h_{0,1} \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}, \tag{2.8.71}$$

uniformly on $[0,1]$. This proves (2.8.69) (i). To prove also (2.8.69) (ii) we use the $\mathcal{U}_\gamma$-*sub*harmonic (for each $\gamma > 0$) function h_1 from Lemma 2.75. By Lemma 2.69 also $\frac{1}{2}h_1$ is $\mathcal{U}_\gamma$-subharmonic. By bounding $\frac{1}{2}h_1$ from above and below with multiples of $h_{0,1}$ it is easy to derive from Proposition 2.83 (a) that

$$(\mathcal{U}_{\gamma^*})^n(\tfrac{1}{2}h_1) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*} \tag{2.8.72}$$

uniformly on $[0,1]$. Arguing as before, we can find $l(n) \to \infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)}(\tfrac{1}{2}h_1) - \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}}(\tfrac{1}{2}h_1)\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{2.8.73}$$

Therefore, by (2.8.72) and the facts that $\frac{1}{2}h_1$ is $\mathcal{U}_{\gamma_k}$-subharmonic for each $k \geq 0$ and $\frac{1}{2}h_1 \leq \frac{1}{2}h_{0,1}$,

$$\mathcal{U}^{(n)}(\tfrac{1}{2}h_{0,1}) \geq \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}}(\tfrac{1}{2}h_1) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}, \tag{2.8.74}$$

uniformly on $[0,1]$, which proves (2.8.69) (ii). The proof of (2.8.69) in case $\gamma^* = 0$ is completely analogous. ∎

2.9 The renormalization branching process: extinction on the interior

2.9.1 Basic facts

In this section we prove Proposition 2.82 (a). To simplify notation, throughout this section h denotes the function $h_{0,0}$. We fix $0 < \gamma^* < \infty$, we let $Y^h := Y^{\gamma^*,h}$ denote the branching particle system on $(0,1)$ obtained from $\mathcal{Y}^{\gamma^*} = (\mathcal{Y}^{\gamma^*}_0, \mathcal{Y}^{\gamma^*}_1, \ldots)$ by Poissonization with h in the sense of Proposition 2.21, and we denote its log-Laplace operator by $U^h_{\gamma^*}$. We will prove that

$$\rho(x) := P^{\delta_x}\big[Y^h_n \neq 0\ \forall n \geq 0\big] = 0 \qquad (x \in (0,1)). \tag{2.9.1}$$

Since for each n fixed, $x \mapsto \rho_n(x) := P^{\delta_x}[Y^h_n \neq 0]$ is a continuous function that decreases to $\rho(x)$, (2.9.1) implies that $\rho_n(x) \to 0$ locally uniformly on $(0,1)$, which, by an obvious analogon of Lemma 2.79, yields Proposition 2.82 (a).

As a first step, we prove:

Lemma 2.84 (Continuous survival probability) *One has either $\rho(x) = 0$ for all $x \in (0,1)$ or there exists a continuous function $\tilde\rho : (0,1) \to [0,1]$ such that $\rho(x) \geq \tilde\rho(x) > 0$ for all $x \in (0,1)$.*

Proof Put $p(x) := h(x)\rho(x)$. We will show that either $p = 0$ on $(0,1)$ or there exists a continuous function $\tilde p : (0,1) \to (0,1]$ such that $p \geq \tilde p$ on $(0,1)$. Indeed,

$$\begin{aligned} p(x) &= h(x)P^{\delta_x}\big[Y^h_n \neq 0\ \forall n \geq 0\big] = \lim_{n\to\infty} h(x)P^{\delta_x}\big[Y^h_n \neq 0\big] \\ &= h(x)\lim_{n\to\infty}(U^h_{\gamma^*})^n 1(x) = \lim_{n\to\infty}(\mathcal{U}_{\gamma^*})^n h(x) \qquad (x \in (0,1)), \end{aligned} \tag{2.9.2}$$

where we have used (2.2.30) and (2.2.36) in the last two steps. Using the continuity of $\mathcal{U}_{\gamma^*}$ with respect to decreasing sequences, it follows that

$$\mathcal{U}_{\gamma^*} p = p. \tag{2.9.3}$$

We claim that for any $f \in B_{[0,1]}[0,1]$, one has the bounds

$$\langle \Gamma^\gamma_x, f \rangle \leq \mathcal{U}_\gamma f(x) \leq (1+\gamma)\langle \Gamma^\gamma_x, f \rangle \qquad (\gamma > 0,\ x \in [0,1]). \tag{2.9.4}$$

Indeed, by Lemma 2.72, $\mathcal{U}_\gamma f(x) \geq 1 - E[(1 - f(\mathbf{y}^\gamma_x(0)))] = \langle \Gamma^\gamma_x, f \rangle$, while the upper bound in (2.9.4) follows from (2.8.11).

By Remark 2.31, $(0,1) \ni x \mapsto \langle \Gamma^\gamma_x, f \rangle$ is continuous for all $f \in B_{[0,1]}[0,1]$. Moreover, $\langle \Gamma^\gamma_x, f \rangle = 0$ for some $x \in (0,1)$ if and only if $f = 0$ almost everywhere with respect to Lebesgue measure.

Applying these facts to $f = p$ and $\gamma = \gamma^*$, using (2.9.3), we see that there are two possibilities. Either $p = 0$ a.s. with respect to Lebesgue measure, and in this case $p = 0$ by the upper bound in (2.9.4), or p is not almost everywhere zero with respect to Lebesgue measure, and in this case the function $x \mapsto \tilde{p}(x) := \langle \Gamma^\gamma_x, f \rangle$ is continuous, positive on $(0,1)$, and estimates p from below by the lower bound in (2.9.4). ∎

2.9.2 A representation for the Campbell law

(Local) extinction properties of critical branching processes are usually studied using Palm laws. Our proof of formula (2.9.1) is no exception, except that we will use the closely related Campbell laws. Loosely speaking, Palm laws describe a population that is size-biased at a given position, plus 'typical' particle sampled from that position, while Campbell laws describe a population that is size-biased as a whole, plus a 'typical' particle sampled from a random position.

Let $\mathcal{P}$ be a probability law on $\mathcal{N}(0,1)$ with $\int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)|\nu| = 1$. Then the *size-biased law* $\mathcal{P}_{\rm size}$ associated with $\mathcal{P}$ is the probability law on $\mathcal{N}(0,1)$ defined by

$$\mathcal{P}_{\rm size}(\,\cdot\,) := \int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)\, |\nu| 1_{\{\nu\,\in\,\cdot\,\}}. \tag{2.9.5}$$

The *Campbell law* associated with $\mathcal{P}$ is the probability law on $(0,1) \times \mathcal{N}(0,1)$ defined by

$$\mathcal{P}_{\rm Camp}(A \times B) := \int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)\, \nu(A) 1_{\{\nu\,\in\,B\}} \tag{2.9.6}$$

for all Borel-measurable $A \subset (0,1)$ and $B \subset \mathcal{N}(0,1)$. If (v, V) is a $(0,1) \times \mathcal{N}(0,1)$-valued random variable with law $\mathcal{P}_{\rm Camp}$, then $\mathcal{L}(V) = \mathcal{P}_{\rm size}$, and v is the position of a 'typical' particle chosen from V.

Let

$$\mathcal{P}^{x,n}(\,\cdot\,) := P^{\delta_x}\left[Y^h_n \in \cdot\,\right] \tag{2.9.7}$$

denote the law of Y^h at time n, started at time 0 with one particle at position $x \in (0,1)$. Note that by criticality, $\int_{\mathcal{N}(0,1)} \mathcal{P}^{x,n}(\mathrm{d}\nu)|\nu| = 1$. Using again criticality, it is easy to see that in order to prove the extinction formula (2.9.1), it suffices to show that

$$\lim_{n\to\infty} \mathcal{P}^{x,n}_{\mathrm{size}}\big(\{1,\ldots,N\}\big) = 0 \qquad (x \in (0,1),\ N \geq 1). \tag{2.9.8}$$

In order to prove (2.9.8), we will write down an expression for $\mathcal{P}^{x,n}_{\mathrm{Camp}}$. Let Q^h denote the offspring mechanism of Y^h, and, for fixed $x \in (0,1)$, let $Q^h_{\mathrm{Camp}}(x,\,\cdot\,)$ denote the Campbell law associated with $Q^h(x,\,\cdot\,)$. The next proposition is a time-inhomogeneous version of Kallenberg's famous backward tree technique; see [Lie81, Satz 8.2].

Proposition 2.85 (Representation of Campbell law) *Let $(\mathbf{v}_k, V_k)_{k\geq 0}$ be the Markov process in $(0,1)\times\mathcal{N}(0,1)$ with transition laws*

$$P\big[(\mathbf{v}_{k+1}, V_{k+1}) \in \cdot\,\big|\,(\mathbf{v}_k, V_k) = (x,\nu)\big] = Q^h_{\mathrm{Camp}}(x,\,\cdot\,) \qquad ((x,\nu) \in (0,1)\times\mathcal{N}(0,1)), \tag{2.9.9}$$

started in $(\mathbf{v}_0, V_0) = (\delta_x, 0)$. Let $(Y^{h,(k)})^{k\geq 1}$ be branching particle systems with offspring mechanism Q^h, conditionally independent given $(\mathbf{v}_k, V_k)_{k\geq 0}$, started in $Y^{h,(k)}_0 = V_k - \delta_{\mathbf{v}_k}$. Then

$$\mathcal{P}^{x,n}_{\mathrm{Camp}} = \mathcal{L}\Big(\mathbf{v}_n, \delta_{\mathbf{v}_n} + \sum_{k=1}^{n} Y^{h,(k)}_{n-k}\Big). \tag{2.9.10}$$

Formula (2.9.10) says that the Campbell law at time n arises in such a way, that an 'immortal' particle at positions $\mathbf{v}_0,\ldots,\mathbf{v}_n$ sheds off offspring $V_1 - \delta_{\mathbf{v}_1},\ldots,V_n - \delta_{\mathbf{v}_n}$, distributed according to the size-biased law with one 'typical' particle taken out, and this offspring then evolve under the usual forward dynamics till time n. Note that the position of the immortal particle $(\mathbf{v}_k)_{k\geq 0}$ is an autonomous Markov chain.

We need a bit of explicit control on Q^h_{Camp}.

Lemma 2.86 (Campbell law) *One has*

$$Q^h_{\mathrm{Camp}}(x, A\times B) = \frac{\frac{1}{\gamma^*}+1}{h(x)} \int P[\mathrm{Pois}(h\mathcal{Z}^{\gamma^*}_x) \in \mathrm{d}\chi]\chi(A)1_{\{\chi\in A\}}, \tag{2.9.11}$$

where the random measures $\mathcal{Z}^{\gamma^*}_x$ are defined in (2.2.18).

Proof By the definition of the Campbell law (2.9.6), and (2.2.35),

$$\begin{aligned} Q^h_{\mathrm{Camp}}(x, A\times B) &= \int Q^h(x,\mathrm{d}\chi)\chi(A)1_{\{\chi\in B\}} \\ &= \frac{\frac{1}{\gamma^*}+1}{h(x)} \int P[\mathrm{Pois}(h\mathcal{Z}^{\gamma^*}_x) \in \mathrm{d}\chi]\chi(A)1_{\{\chi\in B\}} + \Big(1 - \frac{\frac{1}{\gamma^*}+1}{h(x)}\Big)\cdot 0. \end{aligned} \tag{2.9.12}$$

∎

Recall that by (2.2.18),

$$\mathcal{Z}^{\gamma^*}_x := \int_0^{\tau_{\gamma^*}} \delta_{\mathbf{y}^{\gamma^*}_x(-t/2)}\mathrm{d}t, \tag{2.9.13}$$

where $(\mathbf{y}_x^{\gamma^*}(t))_{t\in\mathbb{R}}$ is a stationary solution to the SDE (2.2.17) with $\gamma = \gamma^*$. By Lemma 2.86, the transition law of the Markov chain $(\mathbf{v}_k)_{k\geq 0}$ from Proposition 2.85 is given by

$$P[\mathbf{v}_{k+1} \in \mathrm{d}y | \mathbf{v}_k = x] = \frac{\frac{1}{\gamma^*}+1}{h(x)} E[\mathrm{Pois}(h\mathcal{Z}_x^{\gamma^*})(\mathrm{d}y)] = \frac{1+\gamma^*}{h(x)} h(y)\Gamma_x^{\gamma^*}(\mathrm{d}y), \tag{2.9.14}$$

where $\Gamma_x^{\gamma^*}$ is the invariant law of $\mathbf{y}_x^{\gamma^*}$ from Corollary 2.30. In the next section we will prove the following lemma.

Lemma 2.87 (Immortal particle stays in interior) *The Markov chain $(\mathbf{v}_k)_{k\geq 0}$ started in any $\mathbf{v}_0 = x \in (0,1)$ satisfies*

$$(\mathbf{v}_k)_{k\geq 0} \textit{ has a cluster point in } (0,1) \quad \textit{a.s.} \tag{2.9.15}$$

We now show that Lemma 2.87, together with our previous results, implies Proposition 2.82 (a).

Proof of Proposition 2.82 (a) We need to prove (2.9.1). By our previous analysis, it suffices to prove (2.9.8) under the assumption that $\rho \neq 0$. By Proposition 2.85,

$$\mathcal{P}^{x,n}_{\mathrm{size}} = \mathcal{L}\Big(\delta_{\mathbf{v}_n} + \sum_{k=1}^{n} Y^{h,(k)}_{n-k}\Big). \tag{2.9.16}$$

Conditioned on $(\mathbf{v}_k, V_k)_{k\geq 0}$, the $(Y^{h,(k)}_{n-k})_{k=1,\ldots,n}$ are independent random variables with

$$P\big[Y^{h,(k)}_{n-k} \neq 0\big] \geq P\big[Y^{h,(k)}_{m} \neq 0 \ \forall m \geq 0\big] = P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0]. \tag{2.9.17}$$

Therefore, (2.9.8) will follow by Borel-Cantelli provided that we can show that

$$\sum_{k=1}^{\infty} P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1}] = \infty \quad \text{a.s.} \tag{2.9.18}$$

Define $f(x) := P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1} = x]$ $(x \in (0,1))$. We need to show that $\sum_{k=1}^{\infty} f(x) = \infty$ a.s. Using Lemma 2.84 and Lemma 2.86 we can estimate

$$f(x) \geq P[\mathrm{Thin}_{\tilde\rho}(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1} = x] = \int_{\mathcal{N}(0,1)} Q^h_{\mathrm{Camp}}(x, \mathrm{d}y, \mathrm{d}\nu)\{1 - (1-\tilde\rho)^{\nu - \delta_y}\} > 0 \tag{2.9.19}$$

for all $x \in (0,1)$. Since $\mathcal{Q}_{\gamma^*}$, defined in (2.2.19), is a continuous cluster mechanism, also $Q^h_{\mathrm{Camp}}(x,\cdot)$ is continuous as a function of x, hence the bound in (2.9.19) is locally uniform on $(0,1)$, hence Lemma 2.87 implies that there is an $\varepsilon > 0$ such that

$$P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1}] \geq \varepsilon \tag{2.9.20}$$

at infinitely many times $k-1$, which in turn implies (2.9.18). ∎

2.9.3 The immortal particle

Proof of Lemma 2.87 Let $K(x,\mathrm{d}y)$ denote the transition kernel (on $(0,1)$) of the Markov chain $(\mathbf{v}_k)_{k\geq 0}$, i.e., by (2.9.14),

$$K(x,\mathrm{d}y) = (1+\gamma^*)\frac{y(1-y)}{x(1-x)}\Gamma_x^{\gamma^*}(\mathrm{d}y). \tag{2.9.21}$$

It follows from (2.3.24) that

$$\int K(x,\mathrm{d}y)y(1-y) = \frac{x(1-x)+\gamma^*(1+\gamma^*)}{(1+2\gamma^*)(1+3\gamma^*)}. \tag{2.9.22}$$

Set

$$g(x) := \int K(x,\mathrm{d}y)y(1-y) - x(1-x) \qquad (x\in(0,1)). \tag{2.9.23}$$

Then

$$M_n := \mathbf{v}_n(1-\mathbf{v}_n) - \sum_{k=0}^{n-1} g(\mathbf{v}_k) \qquad (n\geq 0) \tag{2.9.24}$$

defines a martingale $(M_n)_{n\geq 0}$. Since $g>0$ in an open neighborhood of $\{0,1\}$,

$$P[(\mathbf{v}_k)_{k\geq 0} \text{ has no cluster point in } (0,1)] \leq P[\lim_{n\to\infty} M_n = -\infty] = 0, \tag{2.9.25}$$

where in the last equality we have used that $(M_n)_{n\geq 0}$ is a martingale. ∎

2.10 Proof of the main result

Proof of Theorem 2.17 Part (a) has been proved in Section 2.3.3. It follows from (2.1.42), (2.1.43), (2.2.21), and (2.2.22) that part (b) is equivalent to the following statement. Assuming that

$$\text{(i)} \quad \sum_{n=1}^{\infty}\gamma_n = \infty \qquad \text{and} \qquad \text{(ii)} \quad \gamma_n \underset{n\to\infty}{\longrightarrow} \gamma^* \tag{2.10.1}$$

for some $\gamma^*\in[0,\infty)$, one has, uniformly on $[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_0}(p) \underset{n\to\infty}{\longrightarrow} p^*_{l,r,\gamma^*}, \tag{2.10.2}$$

where p^*_{l,r,γ^*} is the unique solution in $\mathcal{H}_{l,r}$ of

$$\begin{array}{llll} \text{(i)} & \mathcal{U}_{\gamma^*}p^* = p^* & & \text{if } 0<\gamma^*<\infty, \\ \text{(ii)} & \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p^*(x) - p^*(x)(1-p^*(x)) = 0 & (x\in[0,1]) & \text{if } \gamma^*=0. \end{array} \tag{2.10.3}$$

It follows from Proposition 2.22 that the left-hand side of (2.10.2) converges uniformly to a limit p^*_{l,r,γ^*} which is given by (2.2.45). We must show 1° that $p^*_{l,r,\gamma^*}\in\mathcal{H}_{l,r}$ and 2° that p^*_{l,r,γ^*} is the unique solution in this class to (2.10.3). We first treat the case $\gamma^*>0$.

1° Since $p^*_{0,0,\gamma^*} \equiv 0$ and $p^*_{1,1,\gamma^*} \equiv 1$, it is obvious that $p^*_{0,0,\gamma^*} \in \mathcal{H}_{0,0}$ and $p^*_{1,1,\gamma^*} \in \mathcal{H}_{1,1}$. Therefore, by symmetry, it suffices to show that $p^*_{0,1,\gamma^*} \in \mathcal{H}_{0,1}$. By Lemmas 2.75 and 2.78, $x \leq p \leq 1-(1-x)^7$ implies $x \leq \mathcal{U}_{\gamma_k} p \leq 1-(1-x)^7$ for each k. Iterating this relation, using (2.10.2), we find that

$$x \leq p^*_{0,1,\gamma^*}(x) \leq 1-(1-x)^7. \tag{2.10.4}$$

By Proposition 2.37, the left-hand side of (2.10.2) is nondecreasing and concave in x if p is, so taking the limit we find that $p^*_{0,1,\gamma^*}$ is nondecreasing and concave. Combining this with (2.10.4) we conclude that $p^*_{0,1,\gamma^*}$ is Lipschitz continuous. Moreover $p^*_{0,1,\gamma^*}(0) = 0$ and $p^*_{0,1,\gamma^*}(1) = 1$ so $p^*_{0,1,\gamma^*} \in \mathcal{H}_{0,1}$.

2° Taking the limit $n \to \infty$ in $(\mathcal{U}_{\gamma^*})^n p = \mathcal{U}_{\gamma^*}(\mathcal{U}_{\gamma^*})^{n-1} p$, using the continuity of $\mathcal{U}_{\gamma^*}$ (Corollary 2.36) and (2.10.2), we find that $\mathcal{U}_{\gamma^*} p^*_{l,r,\gamma^*} = p^*_{l,r,\gamma^*}$. It follows from (2.10.2) that p^*_{l,r,γ^*} is the only solution in $\mathcal{H}_{l,r}$ to this equation.

For $\gamma^* = 0$, it has been shown in [FS03, Proposition 3] that $p^*_{l,r,0}$ is the unique solution in $\mathcal{H}_{l,r}$ to (2.10.3) (ii). In particular, it has been shown there that $p^*_{0,1,0}$ is twice continuously differentiable on $[0,1]$ (including the boundary). This proves parts (b) and (c) of the theorem.∎

Chapter 3

Branching-coalescing particle systems.

3.1 Introduction and main results

3.1.1 Introduction

In this chapter we study systems of particles subject to a stochastic dynamics with the following description. 1° Each particle moves independently of the others according to a continuous time Markov process on a lattice Λ, which jumps from site i to site j with rate $a(i,j)$. 2° Each particle splits with rate $b \geq 0$ into two new particles, created on the position of the old one. 3° Each pair of particles, present on the same site, coalesces with rate $2c$ (with $c \geq 0$) to one particle. 4° Each particle dies with rate $d \geq 0$. Throughout this chapter, we make the following assumptions.

(i) Λ is a finite or countably infinite set.

(ii) The transition rates $a(i,j)$ are irreducible, i.e., if $\Delta \subset \Lambda$ is neither Λ nor $\emptyset$, then there exist $i \in \Delta$ and $j \in \Lambda\backslash\Delta$ such that $a(i,j) > 0$ or $a(j,i) > 0$.

(iii) $\sup_i \sum_j a(i,j) < \infty$.

(iv) $\sum_j a^\dagger(i,j) = \sum_j a(i,j)$, where $a^\dagger(i,j) := a(j,i)$.

(v) b, c, and d are nonnegative constants.

Here and elsewhere sums and suprema over i,j always run over Λ, unless stated otherwise. Assumption (iv) says that the counting measure is an invariant σ-finite measure for the Markov process with jump rates a. With respect to this invariant measure, the time-reversed process jumps from i to j with rate $a^\dagger(i,j)$.

Let $X_t(i)$ denote the number of particles present at site $i \in \Lambda$ and time $t \geq 0$. Then $X = (X_t)_{t\geq 0}$, with $X_t = (X_t(i))_{i\in\Lambda}$, is a Markov process with formal generator

$$\begin{aligned} Gf(x) := &\sum_{ij} a(i,j)x(i)\{f(x+\delta_j-\delta_i)-f(x)\} + b\sum_i x(i)\{f(x+\delta_i)-f(x)\} \\ &+c\sum_i x(i)(x(i)-1)\{f(x-\delta_i)-f(x)\} + d\sum_i x(i)\{f(x-\delta_i)-f(x)\}, \end{aligned} \tag{3.1.1}$$

where $\delta_i(j) := 1$ if $i = j$ and $\delta_i(j) := 0$ otherwise. The process X can be defined for finite initial states and also for some infinite initial states in an appropriate Liggett-Spitzer space (see Section 3.1.3). We call $(X_t)_{t\geq 0}$ a branching coalescing particle system with underlying motion (Λ, a), branching rate b, coalescence rate c and death rate d, or shortly the (a, b, c, d)-braco-process.

Some typical examples of underlying motions we have in mind are nearest neighbour random walk on $\Lambda = \mathbb{Z}^d$ and on $\Lambda = \mathbb{T}^d$, the homogeneous tree of degree $d+1$. We will not restrict ourselves to symmetric underlying motions (i.e., $a = a^\dagger$) but also allow $a(i,j) = 1_{\{j=i+1\}}$ on $\mathbb{Z}$, for example. The reason why we do not restrict ourselves to graphs, is that we also want to include the case $\Lambda = \Omega_d$, the hierarchical group with freedom d, i.e.,

$$\Omega_d := \{i = (i_0, i_1, \ldots) : i_\alpha \in \{0, \ldots, d-1\}\ \forall \alpha \geq 0,\ i_\alpha \neq 0 \text{ finitely often}\}, \tag{3.1.2}$$

equipped with componentwise addition modulo n. On Ω_d, one typically chooses transition rates $a(i,j)$ that depend only on the hierarchical distance $|i-j| := \min\{\alpha \geq 0 : i_\beta = j_\beta\ \forall \beta \geq \alpha\}$. The hierarchical group has found widespread applications in population biology and is therefore a natural choice for the underlying space.

3.1.2 Motivation

Our motivation for studying branching-coalescing particle systems comes from three directions.

Reaction diffusion models, Schlögl's first model. Branching-coalescing particle systems are known in the physics literature as a reaction diffusion models. More precisely, our model is a special case of Schlögl's first model [Sch72], where in the latter there is an additional rate with which particles are spontaneously created. For $d = 0$, our model is known as the autocatalytic reaction. Reaction diffusion models have been studied intensively by physicists and more recently also by probabilists [DDL90, Mou92, Neu90]. All work that we are aware of is restricted to the case $\Lambda = \mathbb{Z}^d$.

Population dynamics, the contact process. Branching-coalescing particle systems may be thought of as a more or less realistic model for the spread and growth of a population of organisms. Here, the underlying motion models the migration of organisms, births and deaths have their obvious interpretations, while coalescence of particles should be thought of as additional deaths, caused by local overpopulation. In this respect, our model is similar to the contact process. The latter is often referred to as a model for the spread of an infection, but in fact it is a reasonable model for the population dynamics of many organisms, from trees in a forest to killer bees. There are two striking differences between the contact process and branching-coalescing particle systems. First, whereas the total population at one site is subject to a rigid bound in the contact process (namely one), it may reach arbitrarily high values in a branching-coalescing system. However, when the local population is high, the coalescence (which grows quadratically in the number of organisms) dominates the branching (which grows linearly), and in this way the population is reduced. A second difference is that in the contact process, if one site infects its neighbor, the original site is still infected. As opposed to this, even when the death rate is zero, it is possible that a branching coalescing particle system goes to local extinction due to migration only. Thus, we can say that the gain from infection is guaranteed in the contact process, whereas the reward for migration is uncertain in a branching-coalescing particle system.

Resampling with selection and negative mutations. Our third motivation also comes from population dynamics, but from a different perspective. Assume that at each site $i \in \Lambda$ there lives a large, fixed number of organisms, and that each of these organisms carries a gene that comes in two types: a healthy and a defective one. Let us model the evolution of the population as follows. 1° with rate $a(i,j)$, we let an organism at site i migrate to site j. 2° to model the effect of natural selection, we let each organism with rate b choose another organism, living on the same site. If the first organism carries a healthy gene and the second organism a defective gene, then the latter is replaced by an organism with a healthy gene. 3° to model the effect of random mating, we resample each pair of organisms living at the same site with rate $2c$, i.e., we choose one of the two at random and replace it by an organism with the type of the other one. 4° with rate d, we let a healthy gene mutate into a defective gene. In the limit that the number of organisms at each site is large, the frequencies $\mathcal{X}_t(i)$ of healthy organisms at site i and time t are described by the unique pathwise solution to the infinite dimensional stochastic differential equation (SDE) (see [SU86]):

$$\begin{aligned} \mathrm{d}\mathcal{X}_t(i) = & \sum_j a(j,i)(\mathcal{X}_t(j) - \mathcal{X}_t(i))\,\mathrm{d}t + b\mathcal{X}_t(i)(1-\mathcal{X}_t(i))\,\mathrm{d}t - d\mathcal{X}_t(i)\,\mathrm{d}t \\ & + \sqrt{2c\mathcal{X}_t(i)(1-\mathcal{X}_t(i))}\,\mathrm{d}B_t(i) \qquad (t \geq 0,\ i \in \Lambda). \end{aligned} \tag{3.1.3}$$

We call the $[0,1]^\Lambda$-valued process $\mathcal{X} = (\mathcal{X}_t)_{t\geq 0}$ the resampling-selection process with underlying motion (Λ, a), selection rate b, resampling rate c and mutation rate d, or shortly the (a,b,c,d)-resem-process (the letters in 'resem' standing for <u>re</u>sampling, <u>se</u>lection and <u>m</u>utation).

It is known that branching-coalescing particle systems are dual to resampling-selection processes. To be precise, for any $\phi \in [0,1]^\Lambda$ and $x \in \mathbb{N}^\Lambda$, write

$$\phi^x := \prod_i \phi(i)^{x(i)}, \tag{3.1.4}$$

where $0^0 := 1$. Let $\mathcal{X}$ be the (a,b,c,d)-resem-process and let $X^\dagger$ be the $(a^\dagger,b,c,d)$-braco-process. Then (see Theorem 3.1 (a) below)

$$E^\phi[(1-\mathcal{X}_t)^x] = E^x[(1-\phi)^{X^\dagger_t}]. \tag{3.1.5}$$

Formula (3.1.5) has the following interpretation: $E^\phi[(1-\mathcal{X}_t)^x]$ is the probability that x organisms, sampled from the population at time t, all have defective genes. If we want to calculate this probability, we must follow back in time those organisms that could possibly be healthy ancestors of these x organisms. In this way we end up with a system of branching coalescing $a^\dagger$-random walks, which die when a mutation occurs, coalesce when two potential ancestors descend from the same ancestor, and branch when a selection event takes place. If we end up with at least one healthy potential ancestor at time zero, then we know that not all the x particles have defective genes.

Resampling-selection processes of the form (3.1.3) are also known as *stepping stone models* (with selection and one type of mutation). These were studied by Shiga and Uchiyama in [SU86], a paper similar in spirit to ours. The duality (3.1.5) is a special case of Lemma 2.1 [SU86]. Moment duals for genetic diffusions in a more general but non-spatial context go back to [Shi81]. The idea of incorporating selection in resampling models by introducing branching into the usual coalescent dual seems to have been independently reinvented in [KN97]. They were probably the first to interpret the duality (3.1.5) in terms of potential ancestors. For some recent versions of this duality, see also [DK99, DG99, BES04]. A SDE that is dual to branching-*annihilating* random walks occurs in [BEM03, Lemma 2.1]. A SPDE version of (3.1.3) (with $d = 0$) has been derived as the rescaled limit of long-range biased voter models in [MT95, Theorem 2].

Note that for $c = 0$, the process $\mathcal{X}$ is deterministic. In this case, the semigroup $(U_t)_{t\geq 0}$ defined by $U_t\phi := \mathcal{X}_t$ $(t \geq 0)$, where $\mathcal{X}$ is the deterministic solution of (3.1.3) with initial state $\mathcal{X}_0 = \phi \in [0,1]^\Lambda$, is called the generating semigroup of the branching particle system $X^\dagger$. (For this terminology, see for example [FS04].) Thus, the duality relation (3.1.5) says that, loosely speaking, branching-coalescing particle systems have a random generating semigroup. The SDE (3.1.3) will be our main tool for studying branching-coalescing particle systems.

3.1.3 Preliminaries

In this section we introduce the notation and definitions that we will use throughout the chapter.

(Inner product and norm notation) For $\phi, \psi \in [-\infty, \infty]^\Lambda$, we write

$$\langle \phi, \psi \rangle := \sum_i \phi(i)\psi(i) \qquad \text{and} \qquad |\phi| := \sum_i |\phi(i)|, \tag{3.1.6}$$

whenever the infinite sums are defined.

(Poisson measures) If ϕ is a $[0,\infty)^\Lambda$-valued random variable, then by definition a Poisson measure with random intensity ϕ is an $\mathbb{N}^\Lambda$-valued random variable $\mathrm{Pois}(\phi)$ whose law is uniquely determined by

$$E[(1-\psi)^{\mathrm{Pois}(\phi)}] = E[e^{-\langle \phi, \psi \rangle}] \qquad (\psi \in [0,1]^\Lambda). \tag{3.1.7}$$

In particular, when ϕ is nonrandom, then the components $(\mathrm{Pois}(\phi)(i))_{i\in\Lambda}$ are independent Poisson distributed random variables with intensity $\phi(i)$.

(Thinned point measures) If x and ϕ are random variables taking values in $\mathbb{N}^\Lambda$ and $[0,1]^\Lambda$, respectively, then by definition a ϕ-thinning of x is an $\mathbb{N}^\Lambda$-valued random variable $\mathrm{Thin}_\phi(x)$ whose law is uniquely determined by

$$E[(1-\psi)^{\mathrm{Thin}_\phi(x)}] = E[(1-\phi\psi)^x] \qquad (\psi \in [0,1]^\Lambda). \tag{3.1.8}$$

In particular, when x and ϕ are nonrandom, and $x = \sum_{n=1}^{m} \delta_{i_n}$, then a ϕ-thinning of x can be constructed as $\mathrm{Thin}_\phi(x) := \sum_{n=1}^{m} \chi_n \delta_{i_n}$ where the χ_n are independent $\{0,1\}$-valued random variables with $P[\chi_n = 1] = \phi(i_n)$.

If ϕ and x are both random, then it will always be understood that they are independent. Thus, $\mathcal{L}(\mathrm{Thin}_\phi(x))$ depends on the laws $\mathcal{L}(\phi)$ and $\mathcal{L}(x)$ alone, and it is only the map $(\mathcal{L}(\phi), \mathcal{L}(x)) \mapsto \mathcal{L}(\mathrm{Thin}_\phi(x))$ that is of interest to us. We have chosen the present notation in terms of random variables instead of their laws to keep things simple if ϕ and x are nonrandom.

We leave it to the reader to check the elementary relations

$$\mathrm{Thin}_\psi(\mathrm{Thin}_\phi(x)) \stackrel{\mathcal{D}}{=} \mathrm{Thin}_{\psi\phi}(x) \quad \text{and} \quad \mathrm{Thin}_\psi(\mathrm{Pois}(\phi)) \stackrel{\mathcal{D}}{=} \mathrm{Pois}(\psi\phi), \tag{3.1.9}$$

where $\stackrel{\mathcal{D}}{=}$ denote equality in distribution.

(Weak convergence) We let $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ denote the one-point compactification of $\mathbb{N}$, and equip $\overline{\mathbb{N}}^\Lambda$ with the product topology. We say that probability measures ν_n on $\overline{\mathbb{N}}^\Lambda$ converge weakly to a limit ν, denoted as $\nu_n \Rightarrow \nu$, when $\int \nu_n(\mathrm{d}x) f(x) \to \int \nu(\mathrm{d}x) f(x)$ for every $f \in \mathcal{C}(\overline{\mathbb{N}}^\Lambda)$, the space of continuous real functions on $\overline{\mathbb{N}}^\Lambda$. One has $\nu_n \Rightarrow \nu$ if and only if $\nu_n(\{x : x(i) = y(i)\ \forall i \in \Delta\}) \to \nu(\{x : x(i) = y(i)\ \forall i \in \Delta\})$ for all finite $\Delta \subset \Lambda$ and $y \in \mathbb{N}^\Delta$.

We equip the space $[0,1]^\Lambda$ with the product topology, and we say that probability measures μ_n on $[0,1]^\Lambda$ converge weakly to a limit μ, denoted as $\mu_n \Rightarrow \mu$, when $\int \mu_n(\mathrm{d}\phi) f(\phi) \to \int \mu(\mathrm{d}\phi) f(\phi)$ for every $f \in \mathcal{C}([0,1]^\Lambda)$.

(Monotone convergence) If ν_1, ν_2 are probability measures on $\overline{\mathbb{N}}^\Lambda$, then we say that ν_1 and ν_2 are stochastically ordered, denoted as $\nu_1 \leq \nu_2$, if $\overline{\mathbb{N}}^\Lambda$-valued random variables Y_1, Y_2 with laws $\mathcal{L}(Y_i) = \nu_i$ $(i = 1,2)$ can be coupled such that $Y_1 \leq Y_2$. We say that a sequence of probability measures ν_n on $\mathbb{N}^\Lambda$ decreases (increases) stochastically to a limit ν, denoted as $\nu_n \downarrow \nu$ $(\nu_n \uparrow \nu)$, if random variables Y_n, Y with laws $\mathcal{L}(Y_n) = \nu_n$ and $\mathcal{L}(Y) = \nu$ can be coupled such that $Y_n \downarrow Y$ $(Y_n \uparrow Y)$. It is not hard to see that $\nu_n \downarrow \nu$ $(\nu_n \uparrow \nu)$ implies $\nu_n \Rightarrow \nu$. Stochastic ordering and monotone convergence of probability measures on $[0,1]^\Lambda$ are defined in the same way.

(Finite systems) We denote the set of finite particle configurations by $\mathcal{N}(\Lambda) := \{x \in \mathbb{N}^\Lambda : |x| < \infty\}$ and let

$$\mathcal{S}(\mathcal{N}(\Lambda)) := \{f : \mathcal{N}(\Lambda) \to \mathbb{R} : |f(x)| \leq K|x|^k + M \text{ for some } K, M, k \geq 0\} \tag{3.1.10}$$

denote the space of real functions on $\mathcal{N}(\Lambda)$ satisfying a polynomial growth condition. For finite initial conditions, the (a,b,c,d)-braco-process X is well-defined as a Markov process in $\mathcal{N}(\Lambda)$ (in particular, X does not explode), $f(X_t)$ is absolutely integrable for each $f \in \mathcal{S}(\mathcal{N}(\Lambda))$ and $t \geq 0$, and the semigroup

$$S_t f(x) := E^x[f(X_t)] \qquad (t \geq 0,\ x \in \mathcal{N}(\Lambda),\ f \in \mathcal{S}(\mathcal{N}(\Lambda))) \tag{3.1.11}$$

maps $\mathcal{S}(\mathcal{N}(\Lambda))$ into itself (see Proposition 3.8 below).

(Liggett-Spitzer space) Set $a_{\mathrm{s}}(i,j) := a(i,j) + a^\dagger(i,j)$. It follows from our assumptions on a that there exist (strictly) positive constants $(\gamma_i)_{i \in \Lambda}$ such that

$$\sum_i \gamma_i < \infty \quad \text{and} \quad \sum_j a_{\mathrm{s}}(i,j)\gamma_j \leq K\gamma_i \quad (i \in \Lambda) \tag{3.1.12}$$

for some $K < \infty$. We fix such $(\gamma_i)_{i\in\Lambda}$ throughout the chapter and define the Liggett-Spitzer space (after [LS81])

$$\mathcal{E}_\gamma(\Lambda) := \{x \in \mathbb{N}^\Lambda : \|x\|_\gamma < \infty\}, \tag{3.1.13}$$

where for $x \in \mathbb{Z}^\Lambda$ we put

$$\|x\|_\gamma := \sum_i \gamma_i |x(i)|. \tag{3.1.14}$$

We let $\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ denote the class of Lipschitz functions on $\mathcal{E}_\gamma(\Lambda)$, i.e., $f : \mathcal{E}_\gamma(\Lambda) \to \mathbb{R}$ such that $|f(x) - f(y)| \leq L\|x - y\|_\gamma$ for some $L < \infty$.

(Infinite systems) It is known ([Che87], see also Proposition 3.11 below) that for each $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ and $t \geq 0$, the function $S_t f$ defined in (3.1.11) can be extended to a unique Lipschitz function on $\mathcal{E}_\gamma(\Lambda)$, also denoted by $S_t f$. Moreover, there exists a time-homogeneous Markov process X in $\mathcal{E}_\gamma(\Lambda)$ (also called (a, b, c, d)-braco-process) with transition laws given by

$$E^x[f(X_t)] = S_t f(x) \qquad (f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda)),\ x \in \mathcal{E}_\gamma(\Lambda),\ t \geq 0). \tag{3.1.15}$$

We will show (in Proposition 3.11 below) that X has a modification with cadlag sample paths, a fact that may seem obvious but to our knowledge has not been proved before.

(Survival and extinction) We say that the (a, b, c, d)-braco-process survives if

$$P^x[X_t \neq 0\ \forall t \geq 0] > 0 \quad \text{for some} \quad x \in \mathcal{N}(\Lambda). \tag{3.1.16}$$

If X does not survive we say that X dies out. Note that the process with death rate $d = 0$ survives, since the number of particles can no longer decrease once only one particle is left. If Λ is finite then the (a, b, c, d)-braco-process survives if and only if $d = 0$, but for infinite Λ survival often holds also for some $d > 0$. For $\Lambda = \mathbb{Z}^d$ and b sufficiently large survival has been proved in [SU86, Theorem 3.1]. We plan to study sufficient conditions for survival in more detail in a forthcoming paper.

(Nontrivial measures) We say that a probability measure ν on $\overline{\mathbb{N}}^\Lambda$ is nontrivial if $\nu(\{0\}) = 0$, where $0 \in \overline{\mathbb{N}}^\Lambda$ denotes the zero configuration. Likewise, we say that a probability measure μ on $[0,1]^\Lambda$ is nontrivial if $\mu(\{0\}) = 0$.

(Homogeneous lattices) By definition, an *automorphism* of (Λ, a) is a bijection $g : \Lambda \to \Lambda$ such that $a(gi, gj) = a(i, j)$ for all $i, j \in \Lambda$. We denote the group of all automorphisms of (Λ, a) by $\mathrm{Aut}(\Lambda, a)$. We say that a subgroup $G \subset \mathrm{Aut}(\Lambda, a)$ is *transitive* if for each $i, j \in \Lambda$ there exists a $g \in G$ such that $gi = j$. We say that (Λ, a) is *homogeneous* if $\mathrm{Aut}(\Lambda, a)$ is transitive. We define shift operators $T_g : \mathbb{N}^\Lambda \to \mathbb{N}^\Lambda$ by

$$T_g x(j) := x(g^{-1}j) \qquad (i \in \Lambda,\ x \in \mathbb{N}^\Lambda,\ g \in \mathrm{Aut}(\Lambda, a)). \tag{3.1.17}$$

If G is a subgroup of $\mathrm{Aut}(\Lambda, a)$, then we say that a probability measure ν on $\mathbb{N}^\Lambda$ is G-*homogeneous* if $\nu \circ T_g^{-1} = \nu$ for all $g \in G$. For example, if $\Lambda = \mathbb{Z}^d$ and $a(i, j) = 1_{\{|i-j|=1\}}$ (nearest-neighbor random walk), then the group G of translations $i \mapsto i + j$ $(j \in \Lambda)$ form a transitive subgroup of $\mathrm{Aut}(\Lambda, a)$ and the G-homogeneous probability measures are the translation invariant probability measures. Shift operators and G-homogeneous measures on $[0,1]^\Lambda$ are defined analogously.

3.1.4 Main results

Our first result is a tool that we exploit substantially towards the main result. Part (a) is known [SU86, Lemma 2.1], but we are not aware of parts (b) and (c) occuring anywhere in the literature.

Theorem 3.1 (Dualities and Poissonization) *Let X and $\mathcal{X}$ be the (a,b,c,d)-braco-process and the (a,b,c,d)-resem-process, respectively, and let $\mathcal{X}^\dagger$ denote the $(a^\dagger,b,c,d)$-resem-process. Then the following holds:*

(a) (Duality)

$$P^x[\mathrm{Thin}_\phi(X_t)=0]=P^\phi[\mathrm{Thin}_{\mathcal{X}_t^\dagger}(x)=0]\qquad(t\geq 0,\ \phi\in[0,1]^\Lambda,\ x\in\mathcal{E}_\gamma(\Lambda)).\tag{3.1.18}$$

(b) (Self-duality) *Assume $c>0$, then*

$$P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t\psi)=0]=P^\psi[\mathrm{Pois}(\tfrac{b}{c}\phi\mathcal{X}_t^\dagger)=0]\qquad(t\geq 0,\ \phi,\psi\in[0,1]^\Lambda).\tag{3.1.19}$$

(c) (Poissonization) *Assume $c>0$, then*

$$P^{\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\phi))}[X_t\in\cdot\,]=P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t)\in\cdot\,]\qquad(t\geq 0,\ \phi\in[0,1]^\Lambda),\tag{3.1.20}$$

i.e., if X is started in the initial law $\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\phi))$ and $\mathcal{X}$ is started in ϕ, then X_t and $\mathrm{Pois}(\frac{b}{c}\mathcal{X}_t)$ are equal in law.

Note that $P[\mathrm{Thin}_\phi(x)=0]=(1-\phi)^x$. Therefore, Theorem 3.1 (a) is just a reformulation of the duality relation (3.1.5). Theorem 3.1 (b) says that resampling-selection processes are in addition dual with respect to each other. In particular, if the underlying motion is symmetric, i.e., $a=a^\dagger$, then this is a self-duality. Since $P[\mathrm{Pois}(\phi)=0]=e^{-|\phi|}$, formula (3.1.19) can be rewritten as

$$E^\phi\big[e^{-\frac{b}{c}\langle\mathcal{X}_t,\psi\rangle}\big]=E^\psi\big[e^{-\frac{b}{c}\langle\phi,\mathcal{X}_t^\dagger\rangle}\big]\qquad(t\geq 0,\ \phi,\psi\in[0,1]^\Lambda).\tag{3.1.21}$$

We note that by [Kal83, Lemma 15.5.1], for $b>0$, the distribution of $\mathcal{X}_t$ is determined uniquely by all $E[e^{-\frac{b}{c}\langle\mathcal{X}_t,\psi\rangle}]$ with $\psi\in[0,1]^\Lambda$. To convince the reader that the notation in (3.1.18) and (3.1.19), which may feel a little uneasy in the beginning, is convenient, we give here the proof of the Poissonization formula (3.1.20).

Proof of Theorem 3.1 (c) By (3.1.9) and the duality relations (3.1.18) and (3.1.19),

$$\begin{aligned}P^{\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\phi))}[\mathrm{Thin}_\psi(X_t)=0]&=P^\psi[\mathrm{Thin}_{\mathcal{X}_t^\dagger}(\mathrm{Pois}(\tfrac{b}{c}\phi))=0]\\&=P^\psi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t^\dagger\phi)=0]=P^\phi[\mathrm{Pois}(\tfrac{b}{c}\psi\mathcal{X}_t)=0]=P^\phi[\mathrm{Thin}_\psi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t))=0].\end{aligned}\tag{3.1.22}$$

Since this is true for all $\psi\in[0,1]^\Lambda$, the random variables X_t and $\mathrm{Pois}(\frac{b}{c}\mathcal{X}_t)$ are equal in distribution. ∎

Our next result shows that it is possible to start the (a,b,c,d)-braco-process with infinitely many particles at each site. This result (except for parts (b) and (f)) has been proved for branching-coalescing particle systems with more general branching and coalescing mechanisms on $\mathbb{Z}^d$ in [DDL90]. Their methods are not restricted to the case $\Lambda=\mathbb{Z}^d$, but we give an independent proof using duality, which has the additional appeal of yielding the explicit bound in part (b).

Theorem 3.2 (The maximal branching-coalescing process) *Assume that $c > 0$. Then there exists an $\mathcal{E}_\gamma(\Lambda)$-valued process $X^{(\infty)} = (X_t^{(\infty)})_{t>0}$ with the following properties:*

(a) *For each $\varepsilon > 0$, $(X_t^{(\infty)})_{t\geq\varepsilon}$ is the (a, b, c, d)-braco-process starting in $X_\varepsilon^{(\infty)}$.*

(b) *Set $r := b - d + c$. Then*

$$E[X_t^{(\infty)}(i)] \leq \begin{cases} \frac{r}{c(1-e^{-rt})} & \text{if } r \neq 0, \\ \frac{1}{ct} & \text{if } r = 0 \end{cases} \qquad (i \in \Lambda,\ t > 0). \tag{3.1.23}$$

(c) *If $X^{(n)}$ are (a, b, c, d)-braco-processes starting in initial states $x^{(n)} \in \mathcal{E}_\gamma(\Lambda)$ such that*

$$x^{(n)}(i) \uparrow \infty \quad \text{as } n \uparrow \infty \qquad (i \in \Lambda), \tag{3.1.24}$$

then

$$\mathcal{L}(X_t^{(n)}) \uparrow \mathcal{L}(X_t^{(\infty)}) \quad \text{as } n \uparrow \infty \qquad (t > 0). \tag{3.1.25}$$

(d) *There exists an invariant measure $\overline{\nu}$ of the (a, b, c, d)-braco-process such that*

$$\mathcal{L}(X_t^{(\infty)}) \downarrow \overline{\nu} \quad \text{as } t \uparrow \infty. \tag{3.1.26}$$

(e) *If ν is another invariant measure for the (a, b, c, d)-braco-process, then $\nu \leq \overline{\nu}$.*

(f) *The measure $\overline{\nu}$ is uniquely characterised by*

$$\int \overline{\nu}(\mathrm{d}x)(1-\phi)^x = P^\phi[\exists t \geq 0 \text{ such that } \mathcal{X}_t^\dagger = 0] \qquad (\phi \in [0,1]^\Lambda), \tag{3.1.27}$$

where $\mathcal{X}^\dagger$ denotes the $(a^\dagger, b, c, d)$-resem-process.

We call $X^{(\infty)}$ the maximal (a, b, c, d)-braco process and we call $\overline{\nu}$ the upper invariant measure. To see why Theorem 3.2 (f) holds, note that by Theorem 3.1 (a) and Theorem 3.2 (c),

$$P[\mathrm{Thin}_\phi(X_t^{(\infty)}) = 0] = \lim_{n\uparrow\infty} P^\phi[\mathrm{Thin}_{\mathcal{X}^\dagger}(x^{(n)}) = 0] = P^\phi[\mathcal{X}_t^\dagger = 0] \qquad (\phi \in [0,1]^\Lambda,\ t > 0). \tag{3.1.28}$$

Now 0 is an absorbing state for the (a, b, c, d)-resem-process, and therefore $P^\phi[\mathcal{X}_t^\dagger = 0] = P^\phi[\exists s \leq t \text{ such that } \mathcal{X}_s^\dagger = 0]$. Therefore, taking the limit $t \uparrow \infty$ in (3.1.28) we arrive at (3.1.27).

The (a, b, c, d)-resem process has an upper invariant measure too. Of our next theorem, parts (a)–(c) are simple, but part (d) lies somewhat deeper.

Theorem 3.3 (The maximal resampling-selection process) *Let $\mathcal{X}^1$ denote the (a, b, c, d)-resem-process started in $\mathcal{X}_0^1(i) = 1$ $(i \in \Lambda)$. Then the following holds.*

(a) *There exists an invariant measure $\overline{\mu}$ of the (a, b, c, d)-resem process such that*

$$\mathcal{L}(\mathcal{X}_t^1) \downarrow \overline{\mu} \quad \text{as } t \uparrow \infty. \tag{3.1.29}$$

(b) *If μ is another invariant measure, then $\mu \leq \overline{\mu}$.*

(c) *Let $X^\dagger$ denote the $(a^\dagger, b, c, d)$-braco-process. Then*

$$\int \overline{\mu}(\mathrm{d}\phi)(1-\phi)^x = P^x[\exists t \geq 0 \text{ such that } X^\dagger_t = 0] \qquad (x \in \mathcal{N}(\Lambda)), \tag{3.1.30}$$

and the measure $\overline{\mu}$ is nontrivial if and only if the $(a^\dagger, b, c, d)$-braco-process survives.

(d) *Assume that $c > 0$ and that Λ is infinite. If $\mathcal{Y}$ is a random variable such that $\overline{\mu} = \mathcal{L}(\mathcal{Y})$, then the upper invariant measure of the (a, b, c, d)-braco-process is given by $\overline{\nu} = \mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{Y}))$. If $\overline{\mu}$ is nontrivial then so is $\overline{\nu}$.*

Note that $\int \overline{\mu}(\mathrm{d}\phi)(1-\phi)^x$ is the probability that x individuals, sampled from a population with resampling and selection in the equilibrium measure $\overline{\mu}$, all have defective genes.

The following is our main result.

Theorem 3.4 (Convergence to the upper invariant measure) *Assume that (Λ, a) is infinite and homogeneous, G is a transitive subgroup of $\mathrm{Aut}(\Lambda, a)$, and $c > 0$.*

(a) *Let X be the (a, b, c, d)-braco process started in a G-homogeneous nontrivial initial law $\mathcal{L}(X_0)$. Then $\mathcal{L}(X_t) \Rightarrow \overline{\nu}$ as $t \to \infty$, where $\overline{\nu}$ is the upper invariant measure.*

(b) *Let $\mathcal{X}$ be the (a, b, c, d)-resem process started in a G-homogeneous nontrivial initial law $\mathcal{L}(\mathcal{X}_0)$. Then $\mathcal{L}(\mathcal{X}_t) \Rightarrow \overline{\mu}$ as $t \to \infty$, where $\overline{\mu}$ is the upper invariant measure.*

Shiga and Uchiyama [SU86, Theorems 1.3 and 1.4] proved Theorem 3.4 (b) under the additional assumptions that $\Lambda = \mathbb{Z}^d$ and that a satisfies a first moment condition in case the death rate d is zero. As we will show below Theorem 3.4 (b) can be derived from Theorem 3.4 (a) by Poissonization, but not vice versa.

3.1.5 Methods

A key ingredient in the proofs of Theorem 3.3 (d) and Theorem 3.4 is the following property of resampling-selection processes, which is of some interest on its own.

Lemma 3.5 (Extinction versus unbounded growth) *Assume that $c > 0$. Let $\mathcal{X}$ be the (a, b, c, d)-resem-process starting in an initial state $\phi \in [0,1]^\Lambda$ with $|\phi| < \infty$. Then $e^{-\frac{b}{c}|\mathcal{X}_t|}$ is a submartingale, and a martingale if $d = 0$. If moreover Λ is infinite, then*

$$\mathcal{X}_t = 0 \quad \textit{for some } t \geq 0 \quad \textit{or} \quad \lim_{t\to\infty} |\mathcal{X}_t| = \infty \quad \text{a.s.} \tag{3.1.31}$$

Note that by Theorem 3.1 (b),

$$E^\phi\big[e^{-\frac{b}{c}\langle \mathcal{X}_t, 1\rangle}\big] = E^1\big[e^{-\frac{b}{c}\langle \phi, \mathcal{X}^\dagger_t\rangle}\big] \geq e^{-\frac{b}{c}\langle \phi, 1\rangle} \qquad (\phi \in [0,1]^\Lambda), \tag{3.1.32}$$

with equality if $d = 0$, since 1 is a stationary state for the $(a^\dagger, b, c, 0)$-resem-process. This shows that $e^{-\frac{b}{c}|\mathcal{X}_t|}$ is a submartingale, and a martingale if $d = 0$. By submartingale convergence, $|\mathcal{X}_t|$ converges a.s. to a limit in $[0, \infty]$. All the hard work of Lemma 3.5 consists of proving that this limit is a.s. either 0 or ∞, and that $\mathcal{X}$ gets extinct in finite time if the limit is zero.

Once Lemma 3.5 is established the proof of Theorem 3.3 (d) is simple.

Proof of Theorem 3.3 (d) Let $\mathcal{Y}$ be a random variable such that $\overline{\mu} = \mathcal{L}(\mathcal{Y})$ and let Y be a random variable such that $\overline{\nu} = \mathcal{L}(Y)$. By (3.1.9), Theorem 3.1 (b), and Theorem 3.2 (f)

$$\begin{aligned} P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{Y})) = 0] &= \lim_{t\to\infty} P^1[\mathrm{Pois}(\tfrac{b}{c}\phi\mathcal{X}_t) = 0] = \lim_{t\to\infty} P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}^\dagger_t) = 0] \\ &\overset{!}{=} P^\phi[\exists t \geq 0 \text{ such that } \mathcal{X}^\dagger_t = 0] = P[\mathrm{Thin}_\phi(Y) = 0], \end{aligned} \tag{3.1.33}$$

where we have used Lemma 3.5 in the equality marked with '!'. Since (3.1.33) holds for all $\phi \in [0,1]^\Lambda$, the random variables $\mathrm{Pois}(\frac{b}{c}\mathcal{Y})$ and Y are equal in distribution. By Lemma 3.5, $|\mathcal{Y}| \in \{0,\infty\}$ a.s. and therefore if $\overline{\mu}$ is nontrivial then $\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{Y}))$ is nontrivial. ∎

In view of Theorem 3.3 (d), it is natural to ask if for infinite lattices, every invariant law of the (a,b,c,d)-braco-process is the Poissonization of an invariant law of the (a,b,c,d)-resem-process. We do not know the answer to this question.

In order to give a very short proof of Theorem 3.4, we need one more lemma.

Lemma 3.6 (Systems with particles everywhere) *Assume that (Λ,a) is infinite and homogeneous and that G is a transitive subgroup of $\mathrm{Aut}(\Lambda,a)$. Let X be the (a,b,c,d)-braco process started in a G-homogeneous nontrivial initial law $\mathcal{L}(X_0)$. Then, for any $t>0$*

$$\lim_{n\to\infty} P[\mathrm{Thin}_{\phi_n}(X_t) = 0] = 0, \tag{3.1.34}$$

for all $\phi_n \in [0,1]^\Lambda$ satisfying $|\phi_n| \to \infty$.

Proof of Theorem 3.4 (a) Let $\mathcal{X}^\dagger$ denote the $(a^\dagger,b,c,d)$-resem-process started in ϕ. By Theorem 3.1 (a), Lemmas 3.5 and 3.6, and Theorem 3.2 (f),

$$\begin{aligned} \lim_{t\to\infty} P[\mathrm{Thin}_\phi(X_t) = 0] &= \lim_{t\to\infty} P[\mathrm{Thin}_{\mathcal{X}^\dagger_{t-1}}(X_1) = 0] \\ &= P[\exists t \geq 0 \text{ such that } \mathcal{X}^\dagger_t = 0] = \int \overline{\nu}(\mathrm{d}x)\,(1-\phi)^x. \end{aligned} \tag{3.1.35}$$

Since this holds for all $\phi \in [0,1]^\Lambda$, it follows that $\mathcal{L}(X_t) \Rightarrow \overline{\nu}$. ∎

Proof of Theorem 3.4 (b) Let X_∞ and $\mathcal{X}_\infty$ be random variables with laws $\overline{\nu}$ and $\overline{\mu}$, respectively. Let $\mathcal{X}$ be the (a,b,c,d)-resem-process started in a G-homogeneous nontrivial initial law $\mathcal{L}(\mathcal{X}_0)$. Let X be the (a,b,c,d)-braco-process started in $\mathcal{L}(X_0) := \mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{X}_0))$. Then by Theorem 3.4 (a), $\mathcal{L}(X_t) \Rightarrow \mathcal{L}(X_\infty)$ as $t\to\infty$. Therefore, by Poissonization (Theorem 3.1 (c)) and by Theorem 3.3 (d), $\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{X}_t)) \Rightarrow \mathcal{L}(X_\infty) = \mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{X}_\infty))$. It follows that

$$\begin{aligned} P\big[e^{-\frac{b}{c}\langle \mathcal{X}_t,\phi\rangle}\big] &= P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t)) = 0] \\ &\Longrightarrow P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_\infty)) = 0] = P\big[e^{-\frac{b}{c}\langle \mathcal{X}_\infty,\phi\rangle}\big] \qquad \text{as } t\to\infty. \end{aligned} \tag{3.1.36}$$

Since this holds for all $\phi \in [0,1]^\Lambda$, we conclude that $\mathcal{L}(\mathcal{X}_t) \Rightarrow \mathcal{L}(\mathcal{X}_\infty)$. ∎

Note that there is no easy way to convert the last argument: if $\mathcal{L}(X_0)$ is homogeneous and nontrivial then we cannot in general find a random variable $\mathcal{X}_0$ such that $\mathcal{L}(X_0) = \mathcal{L}(\mathrm{Pois}(\frac{b}{c}\mathcal{X}_0))$.

For example, this is the case if $X_0(i) \leq 1$ for each $i \in \Lambda$ a.s. Therefore, Theorem 3.4 (a) is stronger than Theorem 3.4 (b).

Summarizing, all the hard work for getting Theorem 3.4 is in proving Lemmas 3.5 and 3.6, as well as the more basic Theorems 3.1 and 3.2. The heart of the proof of Theorem 3.2 is the bound in part (b). We derive this bound using a 'duality' relation with a nonnegative error term, between the (a,b,c,d)-braco-process and a super random walk (Proposition 3.23). We call this relation a subduality. Theorem 3.2 (b) yields a lower bound on the finite time extinction probabilities of the (a,b,c,d)-resem-process started with small initial mass (Lemma 3.24, in particular formula (3.6.1)), which plays a key role in the proof of Lemma 3.5.

Our methods are similar to those of Shiga and Uchiyama [SU86]. Since they prove a version of our Theorem 3.4 (b), while our main focus is on proving the stronger Theorem 3.4 (a), the roles of X and $\mathcal{X}$ are interchanged in their work. Their Lemma 3.2 and Theorem 4.2 are analogues for the (a,b,c,d)-braco-process X of our Lemma 3.5. The proof of the latter is considerably more involved, however. This is because of the fact that we do not want to use spatial homogeneity and we have to prove that $|\mathcal{X}_t| \to 0$ implies $\mathcal{X}_t = 0$ for some $t \geq 0$, which is obvious for the (a,b,c,d)-braco-process X. On the other hand, we can use the submartingale property of $e^{-\frac{b}{c}|\mathcal{X}_t|}$, a very useful fact that has no analogue for the particle system. Lemma 2.5 in [SU86] is the analogue for the (a,b,c,d)-resem-process $\mathcal{X}$ of our Lemma 3.6. By adapting elements of their proof to our situation, we were able to simplify and considerably shorten our original proof of Lemma 3.6.

Our original proof of Lemma 3.6 assumed that Λ has a group structure, and used an L^2 spatial ergodic theorem for general countable groups that need not be amenable.

3.1.6 Discussion

Generalizing our model, let X be a process in a Liggett-Spitzer subspace of $\mathbb{N}^\Lambda$, with local jump rates

$$\begin{aligned} x &\mapsto x + \delta_j - \delta_i && \text{with rate } a(i,j) \\ x &\mapsto x + \delta_i && \text{with rate } \textstyle\sum_{n=0}^{k} b_n x^{(n)}, \\ x &\mapsto x - \delta_i && \text{with rate } \textstyle\sum_{n=1}^{k+1} c_n x^{(n)}, \end{aligned} \tag{3.1.37}$$

where $x^{(0)} := 1$ and $x^{(n)} := x(x-1)\cdots(x-n+1)$ $(n \geq 1)$. In particular, the (a,b,c,d)-braco-process corresponds to the case $k = 1$, $b_0 = 0$, $b_1 = b$, $c_1 = d$, and $c_2 = c$. Processes with jump rates as in (3.1.37) are known as reaction-diffusion systems. It has been known for a long time that if the coefficients satisfy

$$a = a^\dagger \quad \text{and} \quad b_n = \lambda c_n \quad \text{for some} \quad \lambda \geq 0, \tag{3.1.38}$$

then $\mathcal{L}(\mathrm{Pois}(\lambda))$ is a reversible equilibrium for the corresponding reaction-diffusion system. Note that the (a,b,c,d)-braco-process satisfies (3.1.38) if and only if $a = a^\dagger$ and $d = 0$.

The ergodic behavior of reaction-diffusion systems on $\Lambda = \mathbb{Z}^d$ satisfying the reversibility condition (3.1.38) was studied by Ding, Durrett and Liggett in [DDL90]. For our model with $a = a^\dagger$ and $d = 0$ on $\mathbb{Z}^d$, they show that all homogeneous invariant measures are convex combinations of δ_0 and $\mathcal{L}(\mathrm{Pois}(\frac{b}{c}))$. Their proof uses the fact that for a large block in $\mathbb{Z}^d$, surface terms are small compared to volume terms, i.e., $\mathbb{Z}^d$ is amenable. Such arguments typically fail on nonamenable lattices such as trees, and therefore it is not immediately obvious if their methods can be generalized to such lattices. Our Theorem 3.4 (a) shows that all homogeneous invariant measures of the (a,b,c,d)-braco-process are convex combinations of δ_0 and $\overline{\nu}$, also in the non-reversible case $d>0$ and for nonamenable lattices. Thus, neither reversibility nor amenability are essential here.

On the other hand, we believe that amenability is essential for more subtle ergodic properties of reaction-diffusion processes. In analogy with the contact process, let us say that a reaction-diffusion process with $b_0=0$ exhibits complete convergence, if

$$P^x[X_t\in\cdot\,]\Rightarrow\rho(x)\overline{\nu}+(1-\rho(x))\delta_0\quad\text{as}\quad t\to\infty\qquad(x\in\mathcal{N}(\Lambda)),\tag{3.1.39}$$

where $\rho(x):=P^x[X_t\neq 0\ \forall t\geq 0]$ denotes the survival probability. It has been shown by Mountford [Mou92] that complete convergence holds for reaction-diffusion systems on $\Lambda=\mathbb{Z}^d$ satisfying the reversibility condition (3.1.38), $b_0=0$, and a first moment condition on a. We conjecture that complete convergence holds more generally if $a=a^\dagger$ and Λ is amenable, but not in general on nonamenable lattices. As a motivation for this conjecture, we note that complete convergence holds for the contact process on $\mathbb{Z}^d$ but not in general on $\mathbb{T}^d$; see Liggett [Lig99].

The self-duality of resampling-selection processes (Theorem 3.1 (b)) is reminiscent of the self-duality of the contact process. It is an interesting question whether our methods can be adapted to the contact process, to show that the upper invariant measure of the contact process on a countable group is the limit started from any homogeneous nontrivial initial law.

Other interesting processes that some of our techniques might be applied to are multitype branching-coalescing particle systems. For example, it seems natural to color the particles in a branching-coalescing particle system in two (or more) colors, with the rule that in coalescence of differently colored particles, the newly created particle chooses the color of one of its parents with equal probabilities (neutral selection) or with a prejudice towards one color (positive selection). More difficult questions refer to what happens when the two colors have different parameters b,c,d or even different underlying motions a.

One also wonders whether the techniques in this chapter can be generalized to reaction-diffusion processes with higher-order branching and coalescence as in (3.1.37). It seems that at least some of these systems have some sort of a resampling-selection dual too, now with 'resampling' and 'selection' events involving three and more particles.

We conclude with an intriguing question. Does survival of the (a,b,c,d)-braco-process X imply survival of the $(a^\dagger,b,c,d)$-braco-process $X^\dagger$? If X survives, then Theorem 3.3 (c) and (d) and Theorem 3.4 (a) show that the upper invariant measure of $X^\dagger$ is nontrivial, which suggests that $X^\dagger$ should survive. Survival of $X^\dagger$ is obvious if (Λ,a) and $(\Lambda,a^\dagger)$ are isomorfic, as is the case if $a=a^\dagger$, or if Λ is an Abelian group, with group action denoted by $+$, and $a(i,j)$ depends only on $j-i$. However, even when (Λ,a) is homogeneous, (Λ,a) and $(\Lambda,a^\dagger)$ need in general not be isomorphic, and in this case we don't know the answer to our question.

3.1.7 Outline

We start in Section 3.2 with a few generalities about martingale problems that will be needed in our proofs. In Section 3.3 we construct (a, b, c, d)-braco-processes and (a, b, c, d)-resem-processes and prove some of their elementary properties, such as comparison, approximation with finite systems, moment estimates and martingale problems. Section 3.4 contains the proof of Theorem 3.1 and of the subduality between branching-coalescing particle systems and super random walks. In Section 3.5 we prove Theorems 3.2 and 3.3. In Section 3.6, finally, we prove Lemma 3.5 and Lemma 3.6, thereby completing the proof of Theorem 3.4.

Acknowledgements We thank Klaus Fleischmann who played a stimulating role during the early stages of this project and answered a question about Laplace functionals, Claudia Neuhauser for answering questions about branching-coalescing processes, Olle Häggström for answering questions on nonamenable groups, and Tokuzo Shiga for answering our questions about his work. We thank the referee for drawing our attention to the reference [SU86]. Part of this work was carried out during the visits of Siva Athreya to the Weierstrass Institute for Applied Analysis and Stochastics, Berlin and to the Friedrich-Alexander University Erlangen-Nuremberg, and of Jan Swart to the Indian Statistical Institute, Delhi. We thank all these places for their kind hospitality.

3.2 Martingale problems

3.2.1 Definitions

If E be a metrizable space, we denote by $M(E), B(E)$ the spaces of real Borel measurable and bounded real Borel measurable functions on E, respectively. If A is a linear operator from a domain $\mathcal{D}(A) \subset M(E)$ into $M(E)$ and X is an E-valued process, then we say that X solves the martingale problem for A if X has cadlag sample paths and for each $f \in \mathcal{D}(A)$,

$$E\big[|f(X_t)|\big] < \infty \quad \text{and} \quad \int_0^t E\big[|Af(X_s)|\big]\mathrm{d}s < \infty \qquad (t \geq 0), \tag{3.2.1}$$

and the process $(M_t)_{t\geq 0}$ defined by

$$M_t := f(X_t) - \int_0^t Af(X_s)\mathrm{d}s \qquad (t \geq 0) \tag{3.2.2}$$

is a martingale with respect to the filtration generated by X.

3.2.2 Duality with error term

For later use in Section 3.4, we formulate a theorem giving sufficient conditions for two martingale problems to be dual to each other up to a possible error term. Although the techniques for proving Theorem 3.7 below are well-known (see, for example, [EK86, Section 4.4]), we don't know a good reference for the theorem as is formulated here.

Theorem 3.7 (Duality with error term) *Assume that E_1, E_2 are metrizable spaces and that for $i = 1, 2$, A_i is a linear operator from a domain $\mathcal{D}(A_i) \subset B(E_i)$ into $M(E_i)$. Assume that $\Psi \in B(E_1 \times E_2)$ satisfies $\Psi(\cdot, x_2) \in \mathcal{D}(A_1)$ and $\Psi(x_1, \cdot) \in \mathcal{D}(A_2)$ for each $x_1 \in E_1$ and $x_2 \in E_2$, and that*

$$\Phi_1(x_1, x_2) := A_1\Psi(\cdot, x_2)(x_1) \quad \text{and} \quad \Phi_2(x_1, x_2) := A_2\Psi(x_1, \cdot)(x_2) \qquad (x_1 \in E_1,\ x_2 \in E_2) \tag{3.2.3}$$

are jointly measurable in x_1 and x_2. Assume that X^1 and X^2 are independent solutions to the martingale problems for A_1 and A_2, respectively, and that

$$\int_0^T \mathrm{d}s \int_0^T \mathrm{d}t\, E\big[|\Phi_i(X^1_s, X^2_t)|\big] < \infty \qquad (T \geq 0,\ i = 1, 2). \tag{3.2.4}$$

Then

$$E[\Psi(X^1_T, X^2_0)] - E[\Psi(X^1_0, X^2_T)] = \int_0^T \mathrm{d}t\, E[R(X^1_t, X^2_{T-t})] \qquad (T \geq 0), \tag{3.2.5}$$

where $R(x_1, x_2) := \Phi_1(x_1, x_2) - \Phi_2(x_1, x_2) \quad (x_1 \in E_1,\ x_2 \in E_2)$.

Proof Put

$$F(s,t) := E[\Psi(X^1_s, X^2_t)] \qquad (s, t \geq 0). \tag{3.2.6}$$

Then, for each $T > 0$,

$$\begin{aligned}\int_0^T \mathrm{d}t\, \big\{F(t,0) - F(0,t)\big\} &= \int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t) - F(T-t,t) + F(t,0)\big\} \\ &= \int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t)\big\} - \int_0^T \mathrm{d}t\, \big\{F(t, T-t) - F(t,0)\big\},\end{aligned} \tag{3.2.7}$$

where we have subsituted $t \mapsto T - t$ in the term $-F(T-t,t)$. Since X^1 solves the martingale problem for A_1,

$$E\big[\Psi(X^1_{T-t}, x_2)\big] - E\big[\Psi(X^1_0, x_2)\big] = \int_0^{T-t} \mathrm{d}s\, E\big[\Phi_1(X^1_s, x_2)\big] \qquad (x_2 \in E_2), \tag{3.2.8}$$

and therefore, integrating the x_2-variable with respect to the law of X^2_t, using the independence of X^1 and X^2 and (3.2.4), we find that

$$\begin{aligned}\int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t)\big\} &= \int_0^T \mathrm{d}t\, \big\{E\big[\Psi(X^1_{T-t}, X^2_t)\big] - E\big[\Psi(X^1_0, X^2_t)\big]\big\} \\ = \int_0^T \mathrm{d}t \int_0^{T-t} \mathrm{d}s\, E\big[\Phi_1(X^1_s, X^2_t)\big] &= \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_1(X^1_{t-s}, X^2_s)\big].\end{aligned} \tag{3.2.9}$$

Treating the second term in the right-hand side of (3.2.7) in the same way, we find that

$$\int_0^T \mathrm{d}t\, \big\{F(t,0) - F(0,t)\big\} = \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_1(X^1_{t-s}, X^2_s)\big] - \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_2(X^1_{t-s}, X^2_s)\big]. \tag{3.2.10}$$

Differentiating with respect to T we arrive at (3.2.5). ∎

3.3 Construction and comparison

3.3.1 Finite branching-coalescing particle systems

For finite initial conditions, the (a, b, c, d)-braco-process X can be constructed explicitly using exponentially distributed random variables. The only thing one needs to check is that X does not explode. This is part of the next proposition. Recall the definitions of $\mathcal{N}(\Lambda)$ and $\mathcal{S}(\mathcal{N}(\Lambda))$ from (3.1.10) and of G from (3.1.1).

Proposition 3.8 (Finite braco-processes) *Let X be the (a, b, c, d)-braco-process started in a finite state x. Then X does not explode. Moreover, with $z^{\langle k\rangle} := z(z+1)\cdots(z+k-1)$, one has*

$$E^x\big[|X|_t^{\langle k\rangle}\big] \leq |x|^{\langle k\rangle} e^{kbt} \qquad (k = 1, 2, \ldots,\ t \geq 0). \tag{3.3.1}$$

For each $f \in \mathcal{S}(\mathcal{N}(\Lambda))$, one has $Gf \in \mathcal{S}(\mathcal{N}(\Lambda))$ and X solves the martingale problem for the operator G with domain $\mathcal{S}(\mathcal{N}(\Lambda))$.

Proof Introduce stopping times $\tau_N := \inf\{t \geq 0 : |X_t| \geq N\}$. Put $f_t^k(x) := |x|^{\langle k\rangle} e^{-kbt}$. It is easy to see that

$$\{G + \tfrac{\partial}{\partial t}\} f_t^k(x) \leq kb|x|^{\langle k\rangle} e^{-kbt} - kb|x|^{\langle k\rangle} e^{-kbt} = 0. \tag{3.3.2}$$

The stopped process $(X_{t\wedge\tau_N})_{t\geq 0}$ is a jump process in $\{x \in \mathbb{N}^\Lambda : |x| \leq N\}$ with bounded jump rates, and therefore standard theory tells us that the process $(M_t)_{t\geq 0}$ given by

$$M_t := f_{t\wedge\tau_N}^k(X_{t\wedge\tau_N}) - \int_0^{t\wedge\tau_N} \big(\{G + \tfrac{\partial}{\partial s}\} f_s^k\big)(X_s)\,\mathrm{d}s \qquad (t \geq 0) \tag{3.3.3}$$

is a martingale. By (3.3.2), it follows that $E^x\big[|X_{t\wedge\tau_N}|^{\langle k\rangle} e^{-kb(t\wedge\tau_N)}\big] \leq |x|^{\langle k\rangle}$ and therefore

$$E^x\big[|X_{t\wedge\tau_N}|^{\langle k\rangle}\big] \leq |x|^{\langle k\rangle} e^{kbt} \qquad (k = 1, 2, \ldots,\ t \geq 0). \tag{3.3.4}$$

In particular, setting $k = 1$, we see that

$$NP^x[\tau_N \leq t] \leq E^x\big[|X_{t\wedge\tau_N}|\big] \leq |x|e^{bt} \qquad (t \geq 0), \tag{3.3.5}$$

which shows that $\lim_{N\to\infty} P^x[\tau_N \leq t] = 0$ for all $t \geq 0$, i.e., the process does not explode. Taking the limit $N \uparrow \infty$ in (3.3.4), using Fatou, we arrive at (3.3.1).

If $f \in \mathcal{S}(\mathcal{N}(\Lambda))$ then f is bounded on sets of the form $\{x \in \mathbb{N}^\Lambda : |x| \leq N\}$, and therefore Gf is well-defined. By standard theory, the processes $(M_t^N)_{t\geq 0}$ given by

$$M_t^N := f(X_{t\wedge\tau_N}) - \int_0^{t\wedge\tau_N} Gf(X_s)\mathrm{d}s \qquad (t \geq 0) \tag{3.3.6}$$

are martingales. It is easy to see that $f \in \mathcal{S}(\mathcal{N}(\Lambda))$ implies $Gf \in \mathcal{S}(\mathcal{N}(\Lambda))$, and therefore $\int_0^t E[|Gf(X_s)|\mathrm{d}s < \infty$ for all $t \geq 0$ by (3.3.1). Using (3.3.4), one can now check that for fixed $t \geq 0$, the random variables $\{M_t^N\}_{N\geq 1}$ are uniformly integrable. Taking the pointwise limit in (3.3.6), one can now check that X solves the martingale problem for G with domain $\mathcal{S}(\mathcal{N}(\Lambda))$. ∎

3.3.2 Monotonicity and subadditivity

In this section we present two simple comparison results for finite branching-coalescing particle systems.

Lemma 3.9 (Comparison of branching-coalescing particle systems) *Let X and $\tilde{X}$ be the (a,b,c,d)-braco-process and the $(a,\tilde{b},\tilde{c},\tilde{d})$-braco-process started in finite initial states x and $\tilde{x}$, respectively. Assume that*

$$x \leq \tilde{x}, \quad b \leq \tilde{b}, \quad c \geq \tilde{c}, \quad d \geq \tilde{d}. \tag{3.3.7}$$

Then X and $\tilde{X}$ can be coupled in such a way that

$$X_t \leq \tilde{X}_t \qquad (t \geq 0). \tag{3.3.8}$$

Proof We will construct a bivariate process (B, W), say of black and white particles, such that $X = B$ are the black particles and $\tilde{X} = B + W$ are the black and white particles together. To this aim, we let the particles evolve in such a way that black and white particles branch with rates b and $\tilde{b}$, respectively, and additionally black particles give birth to white particles with rate $\tilde{b} - b$. Moreover, all pairs of particles coalesce with rate $2\tilde{c}$, where the new particle is black if at least one of its parents is black, and additionally each pair of black particles is with rate $2c - 2\tilde{c}$ replaced by a pair consisting of one black and one white particle. Finally, all particles die with rate $\tilde{d}$, and additionally, black particles change into white particles with rate $d - \tilde{d}$. It is easy to see that with these rules, X and $\tilde{X}$ are the (a,b,c,d)-braco-process and the $(a,\tilde{b},\tilde{c},\tilde{d})$-braco-process, respectively. ∎

The next lemma has been proved for $\Lambda = \mathbb{Z}^d$ in [SU86, Lemma 2.2]. It can be proved (with particles in three colors) in a similar way as the previous lemma.

Lemma 3.10 (Subadditivity) *Let X, Y, Z be (a,b,c,d)-braco-processes started in finite initial states x, y, and $x + y$, respectively. Then X, Y, Z may be coupled in such a way that X and Y are independent and*

$$Z_t \leq X_t + Y_t \qquad (t \geq 0). \tag{3.3.9}$$

3.3.3 Infinite branching-coalescing particle systems

In this section we carry out the construction of branching-coalescing particle systems for infinite initial conditions. We will also derive two results on the approximation of infinite systems with finite systems, that are needed later on. Except for the statement about sample paths, the next proposition has been proved in [Che87], but we give a proof here for the sake of completeness.

Proposition 3.11 (Construction of branching-coalescing particle systems) *For each $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ and $t \geq 0$, the function $S_t f$ defined in (3.1.11) can be extended to a unique Lipschitz function on $\mathcal{E}_\gamma(\Lambda)$, also denoted by $S_t f$. There exists a unique (in distribution) time-homogeneous Markov process with cadlag sample paths in the space $\mathcal{E}_\gamma(\Lambda)$ equipped with the norm $\|\cdot\|_\gamma$, such that*

$$E^x[f(X_t)] = S_t f(x) \qquad (f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda)),\ x \in \mathcal{E}_\gamma(\Lambda),\ t \geq 0). \tag{3.3.10}$$

We start with the following lemma.

Lemma 3.12 (Action of the semigroup on Lipschitz functions) *If $f : \mathcal{N}(\Lambda) \to \mathbb{R}$ is Lipschitz continuous in the norm $\|\cdot\|_\gamma$ from (3.1.14), with Lipschitz constant L, and K is the constant from (3.1.12), then*

$$|S_t f(x) - S_t f(y)| \leq L e^{(K+b-d)t} \|x-y\|_\gamma \qquad (x, y \in \mathcal{N}(\Lambda),\ t \geq 0). \tag{3.3.11}$$

Proof It follows from Propostion 3.8 that $\frac{\partial}{\partial t} E[f(X_t)] = E[Gf(X_t)]$ for all $f \in \mathcal{S}(\mathcal{N}(\Lambda))$, $t \geq 0$. Applying this to the function $f(x) := \|x\|_\gamma$ we see that

$$\begin{aligned} \tfrac{\partial}{\partial t} E^x[\|X_t\|_\gamma] = & \sum_{ij} a(i,j)(\gamma_j - \gamma_i) E[X_t(i)] + (b-d) E^x[\|X_t\|_\gamma] \\ & - c \sum_i \gamma_i E[X_t(i)(X_t(i)-1)] \ \leq \ (K+b-d) E[\|X\|_\gamma], \end{aligned} \tag{3.3.12}$$

and therefore

$$E^x[\|X_t\|_\gamma] \leq e^{(K+b-d)t} \|x\|_\gamma \qquad (x \in \mathcal{N}(\Lambda)). \tag{3.3.13}$$

Let X^x denote the (a,b,c,d)-braco-process started in x. By Lemma 3.9, we can couple X^x, X^y, $X^{x\wedge y}$, and $X^{x\vee y}$ such that $X^{x\wedge y}_t \leq X^x_t, X^y_t \leq X^{x\vee y}_t$ for all $t \geq 0$. It follows that

$$E[\|X^x_t - X^y_t\|_\gamma] \leq E[\|X^{x\vee y}_t - X^{x\wedge y}_t\|_\gamma]. \tag{3.3.14}$$

By Lemma 3.10, we can couple $X^{x\wedge y}$ and $X^{x\vee y}$ to the process $X^{|x-y|}$ such that $X^{x\vee y}_t \leq X^{x\wedge y}_t + X^{|x-y|}_t$ for all $t \geq 0$. Therefore, by (3.3.14) and (3.3.13),

$$E[\|X^x_t - X^y_t\|_\gamma] \leq E[\|X^{|x-y|}_t\|_\gamma] \leq \|x-y\|_\gamma e^{(K+b-d)t}, \tag{3.3.15}$$

which implies that

$$|S_t f(x) - S_t f(y)| \leq E[|f(X^x_t) - f(X^y_t)|] \leq L E[\|X^x_t - X^y_t\|_\gamma] \leq L\|x-y\|_\gamma e^{(K+b-d)t}, \tag{3.3.16}$$

as required. ∎

Since Lipschitz functions on $\mathcal{N}(\Lambda)$ have a unique Lipschitz extension to $\mathcal{E}_\gamma(\Lambda)$, Lemma 3.12 implies that $S_t f$ can be uniquely extended to a function in $\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ for each $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$.

Lemma 3.13 (Construction of the process for fixed times) *Let $X^{(n)}$ be (a,b,c,d)-braco-processes started in initial states $x^{(n)} \in \mathcal{N}(\Lambda)$ such that $x^{(n)} \uparrow x$ for some $x \in \mathcal{E}_\gamma(\Lambda)$. Then the $X^{(n)}$ may be coupled such that $X^{(n)}_t \uparrow X_t$ $(t \geq 0)$ for some $\overline{\mathbb{N}}^\Lambda$-valued process $X = (X_t)_{t\geq 0}$. The process X satisfies $X_t \in \mathcal{E}_\gamma(\Lambda)$ a.s. $\forall t \geq 0$ and X is a Markov process with semigroup $(S_t)_{t\geq 0}$.*

Proof It follows from Lemma 3.9 that the $X^{(n)}$ can be coupled such that $X^{(n)}_t \leq X^{(n+1)}_t$ $(t \geq 0)$, and therefore $X^{(n)}_t \uparrow X_t$ $(t \geq 0)$ for some $\overline{\mathbb{N}}^\Lambda$-valued random variables X_t. By (3.3.15),

$$E\big[\|X_t - X^{(n)}_t\|_\gamma\big] = \lim_{m\uparrow\infty} E\big[\|X^{(m)}_t - X^{(n)}_t\|_\gamma\big] \leq \|x - x^{(n)}\|_\gamma e^{(K+b-d)t}. \tag{3.3.17}$$

This shows in particular that $E[\|X_t\|_\gamma] < \infty$ and therefore $X_t \in \mathcal{E}_\gamma(\Lambda)$ a.s. $\forall t \geq 0$. If $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ has Lipschitz constant L, then by (3.3.17),

$$\begin{aligned} |E[f(X_t)] - E[f(X_t^{(n)})]| &\leq E[|f(X_t) - f(X_t^{(n)})|] \\ &\leq LE[\|X_t - X_t^{(n)}\|_\gamma] \leq L\|x - x^{(n)}\|_\gamma e^{(K+b-d)t}, \end{aligned} \tag{3.3.18}$$

and therefore

$$E[f(X_t)] = \lim_{n\uparrow\infty} E[f(X_t^{(n)})] = \lim_{n\uparrow\infty} S_t f(x^{(n)}) = S_t f(x). \tag{3.3.19}$$

This proves that for each $x \in \mathcal{E}_\gamma(\Lambda)$ and $t \geq 0$ there exists a probability measure $P_t(x, \cdot)$ on $\mathcal{E}_\gamma(\Lambda)$ such that $\int P_t(x, \mathrm{d}y) f(y) = S_t f(x)$ for all $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$. We need to show that X is the Markov process with transition probabilities $P_t(x, \mathrm{d}y)$. Let $\mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$ denote the class of bounded Lipschitz functions on $\mathcal{E}_\gamma(\Lambda)$. Then $\mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$ is closed under multiplication and S_t maps $\mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$ into itself. Therefore, for all $0 \leq t_0 < \cdots < t_k$ and $f_1, \ldots, f_k \in \mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$, one has

$$E\big[f_1(X_{t_1}^{(n)}) \cdots f_k(X_{t_k}^{(n)})\big] = S_{t_1} f_1 S_{t_2 - t_1} f_2 \cdots S_{t_k - t_{k-1}} f_k(x^{(n)}). \tag{3.3.20}$$

It follows from (3.3.17) that

$$\big|E\big[f_1(X_{t_1}) \cdots f_k(X_{t_k})\big] - E\big[f_1(X_{t_1}^{(n)}) \cdots f_k(X_{t_k}^{(n)})\big]\big| \leq \|x - x^{(n)}\|_\gamma \sum_{i=1}^k L_i e^{(K+b-d)t_k} \prod_{j\neq i} \|f_j\|_\infty, \tag{3.3.21}$$

where L_i is the Lipschitz constant of f_i. Taking the limit $n \uparrow \infty$ in (3.3.20), using (3.3.21), we see that

$$E\big[f_1(X_{t_1}) \cdots f_k(X_{t_k})\big] = S_{t_1} f_1 S_{t_2 - t_1} f_2 \cdots S_{t_k - t_{k-1}} f_k(x), \tag{3.3.22}$$

i.e., X is the Markov process with semigroup $(S_t)_{t\geq 0}$. ∎

Proof of Proposition 3.11 We need to show that the process X from Lemma 3.13 satisfies $X_t \in \mathcal{E}_\gamma(\Lambda)$ $\forall t \geq 0$ a.s. (and not just for fixed times) and that $(X_t)_{t\geq 0}$ has cadlag sample paths with respect to the norm $\|\cdot\|_\gamma$. It suffices to prove these facts on the time interval $[0, 1]$. We will do this by constructing an $\mathcal{E}_\gamma(\Lambda)$-valued process Z such that Z makes only upward jumps, and the number of upward jumps of Z dominates the number of upward jumps of X.

Couple the process $X^{(n)}$ from Lemma 3.13 to a process $Y^{(n)}$ such that the joint process $(X^{(n)}, Y^{(n)})$ is the Markov process in $\mathcal{N}(\Lambda) \times \mathcal{N}(\Lambda)$ with generator

$$\begin{aligned} &G_{X,Y} f(x, y) := \\ &\sum_{ij} a(i,j) x(i) \{f(x + \delta_j - \delta_i, y + \delta_i) - f(x, y)\} + \sum_{ij} a(i,j) y(i) \{f(x, y + \delta_j) - f(x, y)\} \\ &+ b \sum_i x(i) \{f(x + \delta_i, y) - f(x, y)\} + b \sum_i y(i) \{f(x, y + \delta_i) - f(x, y)\} \\ &+ c \sum_i x(i)(x(i) - 1) \{f(x - \delta_i, y + \delta_i) - f(x, y)\} + d \sum_i x(i) \{f(x - \delta_i, y + \delta_i) - f(x, y)\}. \end{aligned} \tag{3.3.23}$$

and initial state $(X_0^{(n)}, Y_0^{(n)}) = (x^{(n)}, 0)$. Indeed, it is not hard to see that the first component of the process with generator $G_{X,Y}$ is the (a,b,c,d)-braco-process, and that $Z^{(n)} := X^{(n)} + Y^{(n)}$ is the Markov process in $\mathcal{N}(\Lambda)$ with generator

$$G_Z f(z) := \sum_{ij} a(i,j) z(i) \{f(z+\delta_j) - f(z)\} + b \sum_i z(i) \{f(z+\delta_i) - f(z)\} \tag{3.3.24}$$

and initial state $Z_0^{(n)} = x^{(n)}$. In analogy with (3.3.13) it is easy to check that

$$E^z[\|Z_t^{(n)}\|_\gamma] \leq \|x^{(n)}\|_\gamma e^{(K+b)t} \qquad (z \in \mathcal{N}(\Lambda),\ t \geq 0). \tag{3.3.25}$$

$Z^{(n)}$ makes only upward jumps and $Z^{(n)}(i)$ makes at least as many upward jumps as $X^{(n)}(i)$. Since $X^{(n)}(i)$ cannot become negative, it follows that

$$|\{t \in [0,1] : X_{t-}^{(n)}(i) \neq X_t^{(n)}(i)\}| \leq x^{(n)}(i) + 2Z_1^{(n)}(i). \tag{3.3.26}$$

Summing with respect to the γ_i, taking expectations, using (3.3.25), we see that

$$\sum_i \gamma_i\, E\big[|\{t \in [0,1] : X_{t-}^{(n)}(i) \neq X_t^{(n)}(i)\}|\big] \leq \|x^{(n)}\|_\gamma (1 + 2e^{K+b}). \tag{3.3.27}$$

Let Z be the increasing limit of the processes $Z^{(n)}$. It follows from (3.3.25) that $Z_1 \in \mathcal{E}_\gamma(\Lambda)$ a.s. Now

$$X_t, X_{t-} \leq Z_t \leq Z_1 \qquad \forall t \in [0,1] \quad \text{a.s.}, \tag{3.3.28}$$

and therefore $X_t, X_{t-} \in \mathcal{E}_\gamma(\Lambda)$ $\forall t \in [0,1]$ a.s. Since a.s. all jumps occur at different times,

$$|\{t \in [0,1] : X_{t-}^{(n)}(i) \neq X_t^{(n)}(i)\}| \uparrow |\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\}| \quad \text{as } n \uparrow \infty. \tag{3.3.29}$$

Thus, taking the limit $n \uparrow \infty$ in (3.3.27) we see that

$$\sum_i \gamma_i\, E\big[|\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\}|\big] \leq \|x\|_\gamma (1 + 2e^{K+b}). \tag{3.3.30}$$

This proves that X has a.s. componentwise cadlag sample paths. If $1 \geq t_n \downarrow t$, then $X_{t_n} \to X_t$ pointwise and $|X_{t_n} - X_t| \leq 2Z_1$, and therefore, by dominated convergence,

$$\|X_{t_n} - X_t\|_\gamma = \sum_i \gamma_i |X_{t_n}(i) - X_t(i)| \to 0. \tag{3.3.31}$$

The same argument shows that $X_{t_n} \to X_{t-}$ for $t_n \uparrow t \leq 1$, i.e., X has cadlag sample paths with respect to the norm $\|\cdot\|_\gamma$. ∎

The proof of Proposition 3.11 yields a useful corollary.

Corollary 3.14 (Locally finite number of jumps) *The (a,b,c,d)-braco-process X satisfies*

$$\sum_i \gamma_i\, E^x\big[|\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\}|\big] \leq \|x\|_\gamma (1 + 2e^{K+b}). \tag{3.3.32}$$

We can now prove two approximation lemmas.

Lemma 3.15 (Convergence of finite dimensional distributions) *Let X^{x_n}, X^x be the (a,b,c,d)-braco-process started in initial states $x_n, x \in \mathcal{E}_\gamma(\Lambda)$, respectively, such that*

$$\lim_{n\to\infty} \|x_n - x\|_\gamma = 0. \tag{3.3.33}$$

Then, for all $0 \le t_1 < \cdots < t_k$, one has

$$(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_k}) \Rightarrow (X_{t_1}, \ldots, X_{t_k}) \qquad \text{as } n \to \infty. \tag{3.3.34}$$

Proof Use (3.3.22) for x_n and then let $n \to \infty$. ∎

Lemma 3.16 (Monotonicities for infinite systems) *Lemmas 3.9 and 3.10 also hold for infinite initial states. If X^x, X^{x_n} are (a,b,c,d)-braco-process started in initial states $x, x_n \in \mathcal{E}_\gamma(\Lambda)$, such that $x_n \uparrow x$, then X^x, X^{x_n} may be coupled such that*

$$X^{x_n}_t(i) \uparrow X^x_t(i) \quad \text{as } n \uparrow \infty \quad \forall i \in \Lambda,\ t \ge 0 \quad \text{a.s.} \tag{3.3.35}$$

Proof The proof of Proposition 3.11 shows that (3.3.35) holds if the x_n are finite. To generalize Lemma 3.9 to infinite initial states $x, \tilde{x}$, it therefore suffices to note that if $x \le \tilde{x}$, then there exist finite $x_n \le \tilde{x}_n$ such that $x_n \uparrow x$ and $\tilde{x}_n \uparrow \tilde{x}$, and then take the limit $n \uparrow \infty$ in (3.3.8) using (3.3.35). Lemma 3.10 can be generalized to infinite x, y by approximation with finite x_n, y_n in the same way. Finally, to see that (3.3.35) remains valid if the x_n are infinite, note that by Lemma 3.9 (which has now been proved in the infinite case), the processes X^{x_n} can be coupled such that $X^{x_n}_t(i) \le X^{x_{n+1}}_t(i)$ for all $i \in \Lambda$ and $t \ge 0$. Denote the increasing limit of the X^{x_n} by X^x. Lemma 3.15 shows that X^x has the same finite dimensional distributions as the (a,b,c,d)-braco-process started in x and it follows from Corollary 3.14 that X^x has componentwise cadlag sample paths, so X^x is a version of the (a,b,c,d)-braco-process started in x. ∎

3.3.4 Construction and comparison of resampling-selection processes

We equip the space $[0,1]^\Lambda$ with the product topology and let $\mathcal{C}([0,1]^\Lambda)$ denote the space of continuous real functions on $[0,1]^\Lambda$, equipped with the supremum norm. By $\mathcal{C}^2_{\text{fin}}([0,1]^\Lambda)$ we denote the space of $\mathcal{C}^2$ functions on $[0,1]^\Lambda$ depending on finitely many coordinates. By definition, $\mathcal{C}^2_{\text{sum}}([0,1]^\Lambda)$ is the space of continuous functions f on $[0,1]^\Lambda$ such that the partial derivatives $\frac{\partial}{\partial\phi(i)}f(\phi)$ and $\frac{\partial^2}{\partial\phi(i)\partial\phi(j)}f(\phi)$ exist for each $x \in (0,1)^\Lambda$ and such that the functions

$$\phi \mapsto \big(\tfrac{\partial}{\partial\phi(i)}f(\phi)\big)_{i\in\Lambda} \quad \text{and} \quad \phi \mapsto \big(\tfrac{\partial^2}{\partial\phi(i)\partial\phi(j)}f(\phi)\big)_{i,j\in\Lambda} \tag{3.3.36}$$

can be extended to continuous functions from $[0,1]^\Lambda$ into the spaces $\ell^1(\Lambda)$ and $\ell^1(\Lambda^2)$ of absolutely summable sequences on Λ and Λ^2, respectively, equipped with the ℓ^1-norm. Define an operator $\mathcal{G} : \mathcal{C}^2_{\text{sum}}([0,1]^\Lambda) \to \mathcal{C}([0,1]^\Lambda)$ by

$$\begin{aligned}\mathcal{G}f(\phi) :=& \sum_{ij} a(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi) + b\sum_i \phi(i)(1-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi)\\ &+c\sum_i \phi(i)(1-\phi(i))\tfrac{\partial^2}{\partial\phi(i)^2}f(\phi) - d\sum_i \phi(i)\tfrac{\partial}{\partial\phi(i)}f(\phi) \qquad (\phi \in [0,1]^\Lambda).\end{aligned} \tag{3.3.37}$$

One can check that for $f \in \mathcal{C}^2_{\rm sum}([0,1]^\Lambda)$, the infinite sums converge in the supremumnorm and the result does not depend on the summation order [Swa99, Lemma 3.4.4]. If a $[0,1]^\Lambda$-valued process $\mathcal{X}$ solves the martingale problem for $\mathcal{G}$ with domain $\mathcal{C}_{\rm fin}([0,1]^\Lambda)$, then also for the larger domain $\mathcal{C}_{\rm sum}([0,1]^\Lambda)$ (see [Swa99, Lemma 3.4.5]).

Let $\mathcal{C}_{[0,1]^\Lambda}[0,\infty)$ denote the space of continuous functions from $[0,\infty)$ into $[0,1]^\Lambda$, equipped with the topology of uniform convergence on compacta. If $\mathcal{X}^{(n)},\mathcal{X}$ are $\mathcal{C}_{[0,1]^\Lambda}[0,\infty)$-valued random variables, then we say that $\mathcal{X}^{(n)}$ converges in distribution to $\mathcal{X}$, denoted as $\mathcal{X}^{(n)}\Rightarrow\mathcal{X}$, when $\mathcal{L}(\mathcal{X}^{(n)})$ converges weakly to $\mathcal{L}(\mathcal{X})$. Convergence in distribution implies convergence of the finite-dimensional distributions (see [EK86, Theorem 3.7.8]). The fact that a $\mathcal{C}_{[0,1]^\Lambda}[0,\infty)$-valued random variable $\mathcal{X}$ solves the martingale problem for $\mathcal{G}$ is a property of the law of $\mathcal{X}$ only. Standard results from [EK86] yield the following (for the details, see for example Lemma 4.1 in [Swa00]):

Lemma 3.17 (Existence and compactness of solutions to the martingale problem) *For each $\phi\in[0,1]^\Lambda$, there exists a solution $\mathcal{X}$ to the martingale problem for $\mathcal{G}$ with initial state $\mathcal{X}_0=\phi$, and each solution to the martingale problem for $\mathcal{G}$ has continuous sample paths. Moreover, the space $\{\mathcal{L}(\mathcal{X}):\mathcal{X}$ solves the martingale problem for $\mathcal{G}\}$ is compact in the topology of weak convergence.*

If $\mathcal{X}$ solves the SDE (3.1.3), then $\mathcal{X}$ solves the martingale problem for $\mathcal{G}$. Conversely, each solution to the martingale problem for $\mathcal{G}$ is equal in distribution to some (weak) solution of the SDE (3.1.3). Thus, existence of (weak) solutions to (3.1.3) follows from Lemma 3.17. Distribution uniqueness of solutions to (3.1.3) follows from pathwise uniqueness, which is in turn implied by the following comparison result.

Lemma 3.18 (Monotone coupling of linearly interacting diffusions) *Let $I\subset\mathbb{R}$ be a closed interval, let $\sigma:I\to\mathbb{R}$ be Hölder-$\frac{1}{2}$-continuous, and let $b_1,b_2:I\to\mathbb{R}$ be Lipschitz continuous functions such that $b_1\leq b_2$. Let $\mathcal{X}^\alpha$ $(\alpha=1,2)$ be solutions, relative to the same system of Brownian motions, of the SDE*

$$\mathrm{d}\mathcal{X}^\alpha_t(i)=\sum_j a(j,i)(\mathcal{X}^\alpha_t(j)-\mathcal{X}^\alpha_t(i))\mathrm{d}t+b_\alpha(\mathcal{X}^\alpha_t(i))\mathrm{d}t+\sigma(\mathcal{X}^\alpha_t(i))\mathrm{d}B_t(i). \tag{3.3.38}$$

$(i\in\Lambda,\ t\geq 0,\ \alpha=1,2)$. *Then*

$$\mathcal{X}^1_0\leq\mathcal{X}^2_0\quad\textit{implies}\quad\mathcal{X}^1_t\leq\mathcal{X}^2_t\quad\forall t\geq 0\quad\text{a.s.} \tag{3.3.39}$$

Proof (sketch) Set $\Delta_t(i):=\mathcal{X}^1_t(i)-\mathcal{X}^2_t(i)$ and write $x^+:=x\vee 0$. Using an appropriate smoothing of the function $x\mapsto x^+$ in the spirit of [YW71, Theorem 1] and arguing as in the proof of [SS80, Theorem 3.2], one can show that

$$E[\|\Delta^+_t\|_\gamma]\leq(K+L)\int_0^t E[\|\Delta^+_s\|_\gamma]\mathrm{d}s, \tag{3.3.40}$$

where $\|\cdot\|_\gamma$ is the norm from (3.1.14), K is the constant from (3.1.12), and L is the Lipschitz-constant of b_2. The result now follows from Gronwall's inequality. ∎

Corollary 3.19 (Comparison of resampling-selection processes) *Assume that $\mathcal{X}, \tilde{\mathcal{X}}$ are solutions to the SDE (3.1.3), relative to the same collection of Brownian motions, with parameters (a,b,c,d) and $(a,\tilde{b},c,\tilde{d})$ and starting in initial states $\phi, \tilde{\phi}$, respectively. Assume that*

$$\phi \leq \tilde{\phi}, \quad d - b \geq \tilde{d} - \tilde{b}, \quad d \geq \tilde{d}. \tag{3.3.41}$$

Then

$$\mathcal{X}_t \leq \tilde{\mathcal{X}}_t \quad \forall t \geq 0 \quad \text{a.s.} \tag{3.3.42}$$

Proof Immediate from Lemma 3.18 and the fact that by (3.3.41), $bx(1-x) - dx \leq \tilde{b}x(1-x) - \tilde{d}x$ for all $x \in [0,1]$. ∎

Our next lemma shows that resampling-selection processes with finite initial mass have finite mass at all later times. The estimate (3.3.43) is not very good if $b - d < 0$, but it suffices for our purposes.

Lemma 3.20 (Summable resampling-selection processes) *Let $\mathcal{X}$ be the (a,b,c,d)-resem-process started in $x \in [0,1]^\Lambda$ with $|x| < \infty$. Set $r := (b-d) \vee 0$. Then*

$$E^x\big[|\mathcal{X}_t|\big] \leq |x| e^{rt} \qquad (t \geq 0), \tag{3.3.43}$$

and $|\mathcal{X}_t| < \infty$ $\forall t \geq 0$ a.s.

Proof Without loss of generality we may assume that $b \geq d$; otherwise, using Corollary 3.19, we can bound $\mathcal{X}$ from above by a braco-process with a higher b. Set $r := b - d$ and put $\mathcal{Y}_t(i) := \mathcal{X}_t(i) e^{-rt}$. By Itô's formula,

$$\mathrm{d}\mathcal{Y}_t(i) = \sum_j a(j,i)(\mathcal{Y}_t(j) - \mathcal{Y}_t(i))\,\mathrm{d}t - be^{-rt}\mathcal{X}_t(i)^2 \mathrm{d}t + e^{-rt}\sqrt{c\mathcal{X}_t(i)(1-\mathcal{X}_t(i))}\,\mathrm{d}B_t(i). \tag{3.3.44}$$

Set $\tau_N := \inf\{t \geq 0 : |\mathcal{X}_t| \geq N\}$. Integrate (3.3.44) up to $t \wedge \tau_N$ and sum over i. The motion terms yield

$$\begin{aligned} &\int_0^{t\wedge\tau_N} \sum_{ij} a(j,i)(\mathcal{Y}_s(j) - \mathcal{Y}_s(i))\,\mathrm{d}s \\ &\quad = \int_0^{t\wedge\tau_N} \sum_j \Big(\sum_i a(j,i)\Big)\mathcal{Y}_s(j)\,\mathrm{d}s - \int_0^{t\wedge\tau_N} \sum_i \Big(\sum_j a^\dagger(i,j)\Big)\mathcal{Y}_s(i)\,\mathrm{d}s = 0, \end{aligned} \tag{3.3.45}$$

where the infinite sums converge in a bounded pointwise way since $|Y_s| \leq N$ for $s \leq \tau_N$. It follows that

$$|\mathcal{Y}_{t\wedge\tau_N}| = |x| - b\sum_i \int_0^{t\wedge\tau_N} \mathcal{X}_s(i)^2 e^{-rs}\mathrm{d}s + \sum_i \int_0^{t\wedge\tau_N} \sqrt{c\mathcal{X}_s(i)(1-\mathcal{X}_s(i))}\, e^{-rs}\mathrm{d}B_s(i), \tag{3.3.46}$$

provided we can show that the infinite sum of stochastic integrals converges. Indeed, for any finite $\Delta \subset \Lambda$, by the Itô isometry,

$$\begin{aligned} &\sum_{i\in\Delta} E\Big[\Big|\int_0^{t\wedge\tau_N} \sqrt{c\mathcal{X}_s(i)(1-\mathcal{X}_s(i))}\, e^{-rs}\mathrm{d}B_s(i)\Big|^2\Big] \\ &\quad = c\sum_{i\in\Delta} E\Big[\int_0^{t\wedge\tau_N} \mathcal{X}_s(i)(1-\mathcal{X}_s(i))e^{-2rs}\mathrm{d}s\Big] \leq cE\Big[\int_0^{t\wedge\tau_N} |\mathcal{X}_s|\mathrm{d}s\Big] \leq ctN, \end{aligned} \tag{3.3.47}$$

which shows that the stochastic integrals in (3.3.46) are absolutely summable in L^2-norm. It follows from (3.3.46) that

$$E^x[|\mathcal{X}_{t\wedge\tau_N}|]e^{-rt} \leq E^x[|\mathcal{X}_{t\wedge\tau_N}|e^{-r(t\wedge\tau_N)}] = E^x[|\mathcal{Y}_{t\wedge\tau_N}|] \leq |x|. \tag{3.3.48}$$

Now $NP^x[\tau_N \leq t] \leq |x|e^{rt}$ for all $t \geq 0$, which shows that $\tau_N \uparrow \infty$ as $N \uparrow \infty$ a.s. Letting $N \uparrow \infty$ in (3.3.48) we arrive at (3.3.43). ∎

We conclude this section with two results on the continuity of $\mathcal{X}$ in its initial state.

Lemma 3.21 (Convergence in law) *Assume that $\mathcal{X}^{(n)}, \mathcal{X}$ are (a,b,c,d)-resem-processes, started in $x^{(n)}, x \in [0,1]^\Lambda$, respectively. Then $x^{(n)} \to x$ implies $\mathcal{X}^{(n)} \Rightarrow \mathcal{X}$.*

Proof By Lemma 3.17, the laws $\mathcal{L}(\mathcal{X}^{(n)})$ are tight and each cluster point of the $\mathcal{L}(\mathcal{X}^{(n)})$ solves the martingale problem for $\mathcal{G}$ with initial state x. Therefore, by uniqueness of solutions to the martingale problem, $\mathcal{X}^{(n)} \Rightarrow \mathcal{X}$. ∎

Lemma 3.22 (Monotone convergence) *Let $\mathcal{X}^{(n)}, \mathcal{X}$ be (a,b,c,d)-resem-processes started in $x^{(n)}, x \in [0,1]^\Lambda$, respectively, such that*

$$x^{(n)} \uparrow x \qquad \textit{as} \quad n \uparrow \infty. \tag{3.3.49}$$

Then $\mathcal{X}^{(n)}, \mathcal{X}$ may be defined on the same probablity space such that

$$\mathcal{X}^{(n)}_t(i) \uparrow \mathcal{X}_t(i) \qquad \forall i \in \Lambda,\ t \geq 0 \quad \textit{as} \quad n \uparrow \infty \quad \text{a.s.} \tag{3.3.50}$$

Proof Let $\mathcal{X}^{(n)}, \mathcal{X}$ be solutions of the SDE (3.1.3) relative to the same system of Brownian motions. By Corollary 3.19, $\mathcal{X}^{(n)} \leq \mathcal{X}^{(n+1)}$ and $\mathcal{X}^{(n)} \leq \mathcal{X}$ for all n. Write $\Delta^{(n)}_t := \mathcal{X}_t - \mathcal{X}^{(n)}_t$ and set $\tau^{(n)}_\varepsilon := \inf\{t \geq 0 : \Delta^{(n)}_t \geq \varepsilon\}$. A calculation as in the proof of Lemma 3.18 shows that

$$\mathrm{d}\|\Delta^{(n)}_t\|_\gamma \leq (K+b)\|\Delta^{(n)}_t\|\mathrm{d}t \quad + \quad \text{martingale terms}. \tag{3.3.51}$$

It follows that

$$E\big[\|\Delta^{(n)}_{t\wedge\tau^{(n)}_\varepsilon}\|_\gamma\big] \leq \|x - x^{(n)}\|_\gamma e^{(K+b)t}. \tag{3.3.52}$$

Now $\varepsilon P[\tau^{(n)}_\varepsilon \leq t] \leq \|x - x^{(n)}\|_\gamma e^{(K+b)t}$ from which we conclude that $\tau^{(n)}_\varepsilon \uparrow \infty$ as $n \uparrow \infty$ for every $\varepsilon > 0$. ∎

3.4 Dualities

3.4.1 Duality and self-duality

Proof of Theorem 3.1 (a) We first prove the statement for finite x. We apply Theorem 3.7. Our duality function is

$$\Psi(x,\phi) := (1-\phi)^x \qquad (x \in \mathcal{N}(\Lambda),\ \phi \in [0,1]^\Lambda). \tag{3.4.1}$$

We need to check that the right-hand side in (3.2.5) is zero, i.e., that

$$G\Psi(\cdot,\phi)(x)=\mathcal{G}^\dagger\Psi(x,\cdot)(\phi)\qquad(\phi\in[0,1]^\Lambda,\ x\in\mathcal{N}(\Lambda)),\tag{3.4.2}$$

where G be the generator of the (a,b,c,d)-braco-process, defined in (3.1.1), and $\mathcal{G}^\dagger$ is the generator of the $(a^\dagger,b,c,d)$-resem-process, defined in (3.3.37). Note that since x is finite, $\Psi(x,\cdot)\in\mathcal{C}^2_{\rm fin}([0,1]^\Lambda)$. We check that

$$\begin{aligned}
&G\Psi(\cdot,\phi)(x)\\
&=\sum_{ij}a(i,j)x(i)\{(1-\phi(j))-(1-\phi(i))\}(1-\phi)^{x-\delta_i}+b\sum_i x(i)\{(1-\phi(i))-1\}(1-\phi)^x\\
&\quad+c\sum_i x(i)(x(i)-1)\{1-(1-\phi(i))\}(1-\phi)^{x-\delta_i}+d\sum_i x(i)\{1-(1-\phi(i)\}(1-\phi)^{x-\delta_i}\\
&=-\sum_{ij}a^\dagger(j,i)(\phi(j)-\phi(i))x(i)(1-\phi)^{x-\delta_i}-b\sum_i\phi(i)(1-\phi(i))x(i)(1-\phi)^{x-\delta_i}\\
&\quad+c\sum_i\phi(i)(1-\phi(i))x(i)(x(i)-1)(1-\phi)^{x-2\delta_i}+d\sum_i\phi(i)x(i)(1-\phi)^{x-\delta_i}\\
&\quad=\mathcal{G}^\dagger\Psi(x,\cdot)(\phi)\qquad(\phi\in[0,1]^\Lambda,\ x\in\mathcal{N}(\Lambda)).
\end{aligned}\tag{3.4.3}$$

Set

$$\Phi(x,\phi):=G\Psi(\cdot,\phi)(x)=\mathcal{G}^\dagger\Psi(x,\cdot)(\phi)\qquad(\phi\in[0,1]^\Lambda,\ x\in\mathcal{N}(\Lambda)).\tag{3.4.4}$$

It is not hard to see that there exists a constant K such that

$$|\Phi(x,\phi)|\le K\Big(1+|x|^2\Big)\qquad(\phi\in[0,1]^\Lambda,\ x\in\mathcal{N}(\Lambda)).\tag{3.4.5}$$

Therefore, condition (3.2.4) is satisfied by (3.3.1).

To generalize the statement from finite x to general $x\in\mathcal{E}_\gamma(\Lambda)$, we apply Lemma 3.16. Choose finite $x^{(n)}$ such that $x^{(n)}\uparrow x$ and couple the (a,b,c,d)-braco-processes $X^{(n)},X$ with initial conditions $x^{(n)},x$, respectively, such that $X^{(n)}\uparrow X$. Then, for each $t\ge 0$ and $\phi\in[0,1]^\Lambda$,

$$E^\phi[(1-\mathcal{X}_t)^{x^{(n)}}]\downarrow E^\phi[(1-\mathcal{X}_t)^x]\quad\text{as }n\uparrow\infty,\tag{3.4.6}$$

and

$$E[(1-\phi)^{X^{(n)}_t}]\downarrow E[(1-\phi)^{X_t}]\quad\text{as }n\uparrow\infty,\tag{3.4.7}$$

where we used the continuity of the function $x\mapsto(1-\phi)^x$ with respect to increasing sequences.∎

Proof of Theorem 3.1 (b) We first prove the statement under the additional assumption that ϕ and ψ are summable. Recall that by Lemma 3.20, if $\mathcal{X}_0$ is summable then $\mathcal{X}_t$ is summable for all $t\ge 0$ a.s. Let $S:=\{\phi\in[0,1]^\Lambda:|\phi|<\infty\}$ denote the space of summable states. We apply Theorem 3.7. Our duality function is

$$\Psi(\phi,\psi):=e^{-\frac{b}{c}\langle\phi,\psi\rangle}\qquad(\phi,\psi\in S).\tag{3.4.8}$$

Let $\mathcal{G},\mathcal{G}^\dagger$ denote the generators of the (a,b,c,d)-resem-process and the $(a^\dagger,b,c,d)$-resem-process, as in (3.3.37), respectively. We need to show that the right-hand side in (3.2.5) is zero, i.e., that $\mathcal{G}\Psi(\cdot,\psi)(\phi) = \mathcal{G}^\dagger\Psi(\phi,\cdot)(\psi)$. It is not hard to see that $\Psi(\cdot,\psi), \Psi(\phi,\cdot) \in \mathcal{C}_{\rm sum}([0,1]^\Lambda)$ for each $\psi,\phi \in S$. We calculate

$$\begin{aligned}
\mathcal{G}\Psi(\cdot,\psi)(\phi) = &\Big\{\sum_{ij} a(j,i)(\phi(j)-\phi(i))(-\tfrac{b}{c})\psi(i) + b\sum_i \phi(i)(1-\phi(i))(-\tfrac{b}{c})\psi(i) \\
&+ c\sum_i \phi(i)(1-\phi(i))(-\tfrac{b}{c})^2\psi(i)^2 - d\sum_i \phi(i)(-\tfrac{b}{c})\psi(i)\Big\} e^{-\frac{b}{c}\langle\phi,\psi\rangle} \\
= &-\tfrac{b}{c}\Big\{\sum_{ij} a(j,i)\phi(j)\psi(i) - \Big(\sum_j a(j,i)\Big)\sum_i \phi(i)\psi(i) \\
&+ b\sum_i \phi(i)(1-\phi(i))\psi(i)(1-\psi(i)) - d\sum_i \phi(i)\psi(i)\Big\} e^{-\frac{b}{c}\langle\phi,\psi\rangle} \\
= &\,\mathcal{G}^\dagger\Psi(\phi,\cdot)(\psi).
\end{aligned} \tag{3.4.9}$$

It is not hard to see that there exists a constant K such that

$$|\mathcal{G}\Psi(\cdot,\psi)(\phi)| \leq K|\phi|\,|\psi| \qquad (\phi,\psi \in S). \tag{3.4.10}$$

Therefore, condition (3.2.4) is implied by Lemma 3.20, and Theorem 3.7 is applicable. To generalize the result to general $\phi,\psi \in [0,1]^\Lambda$, we apply Lemma 3.22. ∎

3.4.2 Subduality

Fix constants $\beta \in \mathbb{R}$, $\gamma \geq 0$. Let $\mathcal{M}(\Lambda) := \{\phi \in [0,\infty)^\Lambda : |\phi| < \infty\}$ be the space of finite measures on Λ, equipped with the topology of weak convergence, and let $\mathcal{Y}$ be the Markov process in $\mathcal{M}(\Lambda)$ given by the unique pathwise solutions to the SDE

$$\mathrm{d}\mathcal{Y}_t(i) = \sum_j a(j,i)(\mathcal{Y}_t(j) - \mathcal{Y}_t(i))\,\mathrm{d}t + \beta\mathcal{Y}_t(i)\,\mathrm{d}t + \sqrt{2\gamma\mathcal{Y}_t(i)}\,\mathrm{d}B_t(i) \tag{3.4.11}$$

($t \geq 0$, $i \in \Lambda$). Then $\mathcal{Y}$ is the well-known super random walk with underlying motion a, growth parameter β and activity γ. One has [Daw93, Section 4.2]

$$E^\phi\big[e^{-\langle\mathcal{Y}_t,\psi\rangle}\big] = e^{-\langle\phi,\mathcal{U}_t\psi\rangle} \tag{3.4.12}$$

for any $\phi \in \mathcal{M}(\Lambda)$ and bounded nonnegative $\psi : \Lambda \to \mathbb{R}$, where $u_t = \mathcal{U}_t\psi$ solves the semilinear Cauchy problem

$$\tfrac{\partial}{\partial t}u_t(i) = \sum_j a(j,i)(u_t(j) - u_t(i)) + \beta u_t(i) - \gamma u_t(i)^2 \qquad (i \in \Lambda,\ t \geq 0) \tag{3.4.13}$$

with initial condition $u_0 = \psi$. The semigroup $(\mathcal{U}_t)_{t\geq 0}$ acting on bounded nonnegative functions ψ on Λ is called the log-Laplace semigroup of $\mathcal{Y}$.

We will show that (a,b,c,d)-braco-process and the super random walk with underlying motion $a^\dagger$, growth parameter $b-d+c$ and activity c are related by a duality formula with a nonnegative error term. In analogy with words such as subharmonic and submartingale, we call this a subduality relation.

Proposition 3.23 (Subduality with a branching process) *Let X be the (a,b,c,d)-braco-process and let $\mathcal{Y}$ be the super random walk with underlying motion $a^\dagger$, growth parameter $b-d+c$ and activity c. Then*

$$E^x\big[e^{-\langle\phi,X_t\rangle}\big]\geq E^\phi\big[e^{-\langle\mathcal{Y}_t,x\rangle}\big]\qquad(x\in\mathcal{E}_\gamma(\Lambda),\ \phi\in\mathcal{M}(\Lambda)).\tag{3.4.14}$$

Proof We first prove the statement for finite x. We apply Theorem 3.7 to X and $\mathcal{Y}$ considered as processes in $\mathcal{N}(\Lambda)$ and $\mathcal{M}(\Lambda)$, respectively. The process $\mathcal{Y}$ solves the martingale problem for the operator

$$\begin{aligned}\mathcal{H}f(\phi):=&\sum_{ij}a^\dagger(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi)+(b-d+c)\sum_i\phi(i)\tfrac{\partial}{\partial\phi(i)}f(\phi)\\&+c\sum_i\phi(i)\tfrac{\partial^2}{\partial\phi(i)^2}f(\phi)\qquad(\phi\in[0,1]^\Lambda),\end{aligned}\tag{3.4.15}$$

defined for functions ϕ in the space $\mathcal{C}^2_{\rm fin,b}[0,\infty)^\Lambda$ of bounded $\mathcal{C}^2$ functions on $[0,\infty)^\Lambda$ depending on finitely many coordinates. Our duality function is $\Psi(x,\phi):=e^{-\langle\phi,x\rangle}$. We observe that $\Psi(x,\cdot)\in\mathcal{C}^2_{\rm fin,b}[0,\infty)^\Lambda$ for all $x\in\mathcal{N}(\Lambda)$ and calculate

$$\begin{aligned}G\Psi(\cdot,\phi)(x)=\Big\{&\sum_{ij}a(i,j)x(i)\big(e^{\phi(i)-\phi(j)}-1\big)+b\sum_i x(i)\big(e^{-\phi(i)}-1\big)\\&+c\sum_i x(i)(x(i)-1)\big(e^{\phi(i)}-1\big)+d\sum_i x(i)\big(e^{\phi(i)}-1\big)\Big\}e^{-\langle\phi,x\rangle},\end{aligned}\tag{3.4.16}$$

and

$$\begin{aligned}\mathcal{H}\Psi(x,\cdot)(\phi)=\Big\{&\sum_{ij}a^\dagger(j,i)x(i)(\phi(i)-\phi(j))-(b-d+c)x(i)\phi(i)\\&+c\sum_i x(i)^2\phi(i)\Big\}e^{-\langle\phi,x\rangle}\end{aligned}\tag{3.4.17}$$

$(x\in\mathcal{N}(\Lambda),\ \phi\in\mathcal{M}(\Lambda))$. It is not hard to see that there exists a constant K such that

$$|G\Psi(\cdot,\phi)(x)|\leq K|x|^2\quad\text{and}\quad|\mathcal{H}\Psi(x,\cdot)(\phi)|\leq K|x|^2\,|\phi|\qquad(x\in\mathcal{N}(\Lambda),\ \phi\in\mathcal{M}(\Lambda)).\tag{3.4.18}$$

and therefore condition (3.2.4) is implied by (3.3.1) and the elementary estimate $E[|\mathcal{Y}_t|]\leq e^{(b-d+c)t}|\phi|$. One has

$$\begin{aligned}G\Psi(\cdot,\phi)(x)-\mathcal{H}\Psi(x,\cdot)(\phi)=\Big\{&\sum_{ij}a(i,j)x(i)\big(e^{\phi(i)-\phi(j)}-1-(\phi(i)-\phi(j))\big)\\+b\sum_i x(i)\big(e^{-\phi(i)}-1+\phi(i)\big)&+c\sum_i x(i)(x(i)-1)\big(e^{\phi(i)}-1-\phi(i)\big)\\+d\sum_i x(i)\big(e^{\phi(i)}-1-\phi(i)\big)\Big\}&e^{-\langle\phi,x\rangle}\geq 0,\end{aligned}\tag{3.4.19}$$

and therefore, for finite x, (3.4.14) is implied by Theorem 3.7. The general case follows by approximation, using Lemma 3.16. ∎

3.5 The maximal processes

3.5.1 The maximal branching-coalescing process

Using Proposition 3.23 we can now prove Theorem 3.2.

Proof of Theorem 3.2 Choose $x^{(n)} \in \mathcal{E}_\gamma(\Lambda)$ such that $x^{(n)}(i) \uparrow \infty$ for all $i \in \Lambda$. By Lemma 3.16, the (a,b,c,d)-braco processes $X^{(n)}$ started in $x^{(n)}$, respectively, can be coupled such that $X^{(n)}_t \leq X^{(n+1)}_t$ for each $t \geq 0$. Define $X^{(\infty)} = (X^{(\infty)}_t)_{t\geq 0}$ as the $\overline{\mathbb{N}}^\Lambda$-valued process that is the pointwise increasing limit of the $X^{(n)}$. By Proposition 3.23 and (3.4.12),

$$E\big[1 - e^{-\langle \varepsilon\delta_i, X^{(n)}_t\rangle}\big] \leq 1 - e^{-\langle \varepsilon\delta_i, \mathcal{U}_t x^{(n)}\rangle} \qquad (t, \varepsilon \geq 0,\ i \in \Lambda). \tag{3.5.1}$$

where $(\mathcal{U}_t)_{t\geq 0}$ is the log-Laplace semigroup of the super random walk with underlying motion $a^\dagger$, growth parameter $r := b - d + c$ and activity c. It follows that

$$E[X^{(n)}_t(i)] = \lim_{\varepsilon\downarrow 0} \varepsilon^{-1} E\big[1 - e^{-\langle \varepsilon\delta_i, X^{(n)}_t\rangle}\big] \leq \lim_{\varepsilon\downarrow 0} \varepsilon^{-1}\big(1 - e^{-\langle \varepsilon\delta_i, \mathcal{U}_t x^{(n)}\rangle}\big) = \mathcal{U}_t x^{(n)}(i) \tag{3.5.2}$$

$(t \geq 0,\ i \in \Lambda)$. Using the explicit solution of (3.4.13) for constant initial conditions, it is easy to see that $\mathcal{U}_t x^{(n)} \uparrow \mathcal{U}_t\infty$, where

$$\mathcal{U}_t\infty := \begin{cases} \frac{r}{c(1-e^{-rt})} & \text{if } r \neq 0, \\ \frac{1}{ct} & \text{if } r = 0. \end{cases} \tag{3.5.3}$$

(See formula (2.6.9).) Letting $n \uparrow \infty$ in (3.5.2) we arrive at Theorem 3.2 (b). Moreover, we see that

$$E\big[\|X^{(\infty)}_t(i)\|_\gamma\big] \leq \mathcal{U}_t\infty \sum_i \gamma_i < \infty \qquad (t > 0), \tag{3.5.4}$$

and therefore $X^{(\infty)}_t \in \mathcal{E}_\gamma(\Lambda)$ a.s. for each $t > 0$. Part (a) of the theorem now follows from Lemma 3.16. Using Theorem 3.1 (a) and the continuity of the function $x \mapsto (1-\phi)^x$ with respect to increasing sequences, reasoning as in (3.1.28), we see that

$$P[\mathrm{Thin}_\phi(X^{(\infty)}_t) = 0] = P^\phi[\mathcal{X}^\dagger_t = 0] \qquad (\phi \in [0,1]^\Lambda,\ t \geq 0), \tag{3.5.5}$$

where $\mathcal{X}^\dagger$ denotes the $(a^\dagger, b, c, d)$-resem-process. Since formula (3.5.5) determines the distribution of $X^{(\infty)}_t$ uniquely, the law of $X^{(\infty)}_t$ does not depend on the choice of the $x^{(n)} \uparrow \infty$ $(t \geq 0)$. This completes the proof of part (c) of the theorem.

To prove part (d), fix $0 \leq s \leq t$. Choose $y_n \in \mathcal{E}_\gamma(\Lambda)$, $y_n(i) \uparrow \infty$ $\forall i \in \Lambda$ and let $\tilde{X}^{(n)}$ be the (a,b,c,d)-braco-process started in $\tilde{X}^{(n)}_0 := X^{(\infty)}_{t-s} \vee y_n$. Then $\tilde{X}^{(n)}_0 \geq X^{(\infty)}_{t-s}$ and therefore, by Lemma 3.9, $\tilde{X}^{(n)}_s$ and $X^{(\infty)}_t$ may be coupled such that $\tilde{X}^{(n)}_s \geq X^{(\infty)}_t$. By part (c) of the theorem, $\tilde{X}^{(n)}_s$ and $X^{(\infty)}_s$ may be coupled such that $\tilde{X}^{(n)}_s \uparrow X^{(\infty)}_s$ and therefore $X^{(\infty)}_s$ and $X^{(\infty)}_t$ may be coupled such that $X^{(\infty)}_s \geq X^{(\infty)}_t$.

It follows that $\mathcal{L}(X_t^{(\infty)}) \downarrow \overline{\nu}$ for some probability measure $\overline{\nu}$ on $\mathcal{E}_\gamma(\Lambda)$. Set $\rho := \mathcal{L}(X_1^{(\infty)})$ and let $(S_t)_{t\geq 0}$ denote the semigroup of the (a,b,c,d)-braco-process. Recall the definition of $\mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$ above (3.3.20). One has

$$\int \overline{\nu}(\mathrm{d}x) f(x) = \lim_{t\to\infty} \int \rho(\mathrm{d}x) S_t f(x) \tag{3.5.6}$$

for every $f \in \mathcal{C}_{\rm Lip,b}(\mathcal{E}_\gamma(\Lambda))$. Therefore, since S_t maps $\mathcal{C}_{\rm Lip,\,b}(\mathcal{E}_\gamma(\Lambda))$ into itself,

$$\int \overline{\nu}(\mathrm{d}x) S_s f(x) = \lim_{t\to\infty} \int \rho(\mathrm{d}x) S_t S_s f(x) = \int \overline{\nu}(\mathrm{d}x) f(x) \qquad (s \geq 0), \tag{3.5.7}$$

for every $f \in \mathcal{C}_{\rm Lip,\,b}(\mathcal{E}_\gamma(\Lambda))$, which shows that $\overline{\nu}$ is an invariant measure. If ν is another invariant measure, then $\mathcal{L}(X_t^{(\infty)}) \geq \nu$ for all $t \geq 0$. Letting $t \to \infty$, we see that $\overline{\nu} \geq \nu$, proving part (e) of the theorem. Part (f) has already been proved in the introduction. ∎

3.5.2 The maximal resampling-selection process

The proof of Theorem 3.3 (a)–(c) is similar to the proof of Theorem 3.2, but easier. Recall that Theorem 3.3 (d) is proved in Section 3.1.5.

Proof of Theorem 3.3 (a)–(c) Part (a) can be proved in the same way as Theorem 3.2 (d), using Lemma 3.22. The proof of part (b) goes analogue to the proof of Theorem 3.2 (e). To see why (3.1.30) holds, note that for any $\phi \in [0,1]^\Lambda$, by Theorem 3.1 (a),

$$\int \overline{\mu}(\mathrm{d}\phi)(1-\phi)^x = \lim_{t\to\infty} P^1[\mathrm{Thin}_{\mathcal{X}_t}(x) = 0] = \lim_{t\to\infty} P^x[\mathrm{Thin}_1(X_t^\dagger) = 0]. \tag{3.5.8}$$

To complete the proof of part (c) we must show that $\overline{\mu}$ is nontrivial if and only if the $(a^\dagger, b, c, d)$-process survives. Using subadditivity (Lemma 3.10) it is easy to see that the $(a^\dagger, b, c, d)$-process survives if and only if $P^{\delta_i}[X_t^\dagger \neq 0 \ \forall t \geq 0] > 0$ for some $i \in \Lambda$. Formula (3.1.30) implies that $\int \overline{\mu}(\mathrm{d}\phi)\phi(i) = P^{\delta_i}[X_t^\dagger \neq 0 \ \forall t \geq 0]$, which shows that $\overline{\mu} = \delta_0$ if and only if the $(a^\dagger, b, c, d)$-process survives. If $\overline{\mu} \neq \delta_0$ then the measure $\overline{\mu}$ conditioned on $\{\phi : \phi \neq 0\}$ is an invariant measure of the (a,b,c,d)-resem-process that is stochastically larger than $\overline{\mu}$. By part (b), this conditioned measure is $\overline{\mu}$ itself, thus $\overline{\mu}(\{0\}) = 0$, i.e., $\overline{\mu}$ is nontrivial. ∎

3.6 Convergence to the upper invariant measure

3.6.1 Extinction versus unbounded growth

In this section we prove Lemma 3.5. It has already been proved in Section 3.1.5 that $e^{-\frac{b}{c}|\mathcal{X}_t|}$ is a submartingale. Therefore, if $b > 0$, then $|\mathcal{X}_t|$ converges a.s. to a limit in $[0, \infty]$. If $b = 0$ then it is easy to see that $|\mathcal{X}_t|$ is a nonnegative supermartingale and therefore also in this case $|\mathcal{X}_t|$ converges a.s. Thus, all we have to do is to show that $\lim_{t\to\infty} |\mathcal{X}_t|$ takes values in $\{0, \infty\}$ a.s. (Proposition 3.25 below), and that $\mathcal{X}$ gets extinct in finite time if the limit is zero (Lemma 3.24). Throughout this section, $c > 0$ and $\mathcal{X}$ is the (a,b,c,d)-resem-process starting in an initial state $\phi \in [0,1]^\Lambda$ with $|\phi| < \infty$.

Lemma 3.24 (Finite time extinction) *One has $\mathcal{X}_t = 0$ for some $t \geq 0$ a.s. on the event $\lim_{t\to\infty} |\mathcal{X}_t| = 0$.*

Proof Choose $x^{(n)} \in \mathcal{E}_\gamma(\Lambda)$ such that $x^{(n)}(i) \uparrow \infty$ for all $i \in \Lambda$. Let $X^{(n)\dagger}$ denote the $(a^\dagger, b, c, d)$-braco-process started in $x^{(n)}$ and let $X^{(\infty)\dagger}$ denote the maximal $(a^\dagger, b, c, d)$-braco-process. By Theorem 3.1 (a) and Theorem 3.2 (b),

$$\begin{aligned} P^\phi[\mathcal{X}_t \neq 0] &= \lim_{n\uparrow\infty} P^\phi[\mathrm{Thin}_{\mathcal{X}_t}(x^{(n)}) \neq 0] = \lim_{n\uparrow\infty} P[\mathrm{Thin}_\phi(X_t^{(n)\dagger}) \neq 0] \\ &= P[\mathrm{Thin}_\phi(X_t^{(\infty)\dagger}) \neq 0] \leq E\big[|\mathrm{Thin}_\phi(X_t^{(\infty)\dagger})|\big] = \langle \phi, E[X_t^{(\infty)\dagger}]\rangle \leq |\phi| \mathcal{U}_t\infty, \end{aligned} \tag{3.6.1}$$

where $\mathcal{U}_t\infty$ is the function on the right-hand side in (3.1.23). Choose $\varepsilon > 0$ and $t_0 > 0$ such that $\varepsilon\mathcal{U}_{t_0}\infty \leq \frac{1}{2}$. Let $(\mathcal{F}_t)_{t\geq 0}$ denote the filtration generated by $\mathcal{X}$. By (3.6.1),

$$\tfrac{1}{2} 1_{\{|\mathcal{X}_t| \,\leq\, \varepsilon\}} \leq P[\mathcal{X}_{t+t_0} = 0 | \mathcal{F}_t] \leq P[\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0 | \mathcal{F}_t]. \tag{3.6.2}$$

Now

$$1_{\{\lim_{s\to\infty} \mathcal{X}_s \,=\, 0\}} \leq \liminf_{t\to\infty} 1_{\{|\mathcal{X}_t| \,\leq\, \varepsilon\}}, \tag{3.6.3}$$

while

$$P[\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0 | \mathcal{F}_t] \to 1_{\{\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0\}} \quad \text{as } t \to \infty \quad \text{a.s.,} \tag{3.6.4}$$

by convergence of right-continuous martingales and the fact that the left-hand side is right-continuous by a general property of strong Markov processes described in Section 2.6.6 from Chapter 2. Letting $t \to \infty$ in (3.6.2), using (3.6.3) and (3.6.4), we find that $\frac{1}{2} 1_{\{\lim_{s\to\infty} \mathcal{X}_s = 0\}} \leq 1_{\{\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0\}}$ a.s. ∎

To finish this section, we need to prove:

Proposition 3.25 (Convergence to zero or infinity) *Assume that Λ is infinite. Then $\lim_{t\to\infty} |\mathcal{X}_t| \in \{0, \infty\}$ a.s.*

Since the proof of Proposition 3.25 is rather long we break it up into a number of steps. At each step, we will skip the proof if it is obvious but tedious. Our first step is:

Lemma 3.26 (Integrable fluctuations) *One has*

$$\int_0^\infty \sum_i \mathcal{X}_t(i)(1 - \mathcal{X}_t(i))\,\mathrm{d}t < \infty \tag{3.6.5}$$

a.s. on the event $\lim_{t\to\infty} |\mathcal{X}_t| \in [0, \infty)$.

Proof For any $\psi \in [0,\infty)^\Lambda$ with $|\psi| < \infty$ one has $e^{-\langle\cdot,\psi\rangle} \in \mathcal{C}^2_{\mathrm{sum}}([0,1]^\Lambda)$ and (compare (3.4.9))

$$\begin{aligned} \mathcal{G} e^{-\langle\cdot,\,\psi\rangle}(\phi) = \Big\{ &-\sum_i \phi(i) \sum_j a^\dagger(j,i)(\psi(j) - \psi(i)) \\ &+ \sum_i \phi(i)(1 - \phi(i))\big(c\psi(i)^2 - b\psi(i)\big) + d\sum_i \phi(i)\psi(i) \Big\} e^{-\langle\phi,\,\psi\rangle}. \end{aligned} \tag{3.6.6}$$

Since $\mathcal{X}$ solves the martingale problem for $\mathcal{G}$,

$$E\Big[\int_0^t \mathcal{G}e^{-\langle\cdot,\psi\rangle}(\mathcal{X}_s)\mathrm{d}s\Big] = E\big[e^{-\langle\mathcal{X}_t,\psi\rangle}\big] - e^{-\langle\phi,\psi\rangle} \qquad (t \geq 0). \tag{3.6.7}$$

Choose $\lambda > 0$ such that $c\lambda^2 - b\lambda =: \mu > 0$ and $\psi_n \in [0,\infty)^\Lambda$ with $|\psi_n| < \infty$ such that $\psi_n \uparrow \lambda$. Then the bounded pointwise limit of the function $i \mapsto \sum_j a^\dagger(j,i)(\psi_n(j) - \psi_n(i))$ is zero and therefore, taking the limit in (3.6.7), using Lemma 3.20, we find that

$$E\Big[\int_0^t \sum_i \Big\{\mu\mathcal{X}_s(i)(1-\mathcal{X}_s(i)) + \lambda d\mathcal{X}_s(i)\Big\}e^{-\lambda|\mathcal{X}_s|}\mathrm{d}s\Big] = E\big[e^{-\lambda|\mathcal{X}_t|}\big] - e^{-\langle\phi,\psi\rangle}. \tag{3.6.8}$$

Letting $t \uparrow \infty$, using the fact that the right-hand side of (3.6.8) is bounded by one, we see that

$$\int_0^\infty \sum_i \Big\{\mu\mathcal{X}_t(i)(1-\mathcal{X}_t(i)) + \lambda d\mathcal{X}_t(i)\Big\}e^{-\lambda|\mathcal{X}_t|}\,\mathrm{d}t < \infty \quad \text{a.s.}, \tag{3.6.9}$$

which implies (3.6.5). ∎

Lemma 3.27 (Process not started with only zeros and ones) *For every $0 < \varepsilon < \frac{1}{4}$ there exists a $\delta, r > 0$ such that*

$$P^\phi\big[\mathcal{X}_t(i) \in (\varepsilon, 1-\varepsilon)\ \forall t \in [0,r]\big] \geq \delta \qquad (i \in \Lambda,\ \phi \in [0,1]^\Lambda,\ \phi(i) \in (2\varepsilon, 1-2\varepsilon)). \tag{3.6.10}$$

Proof Since $\sup_i \sum_j a(i,j) < \infty$ and all the components of the (a,b,c,d)-resem-process take values in $[0,1]$, the maximal drift that the i-th component $\mathcal{X}_t(i)$ can experience (both in the positive and negative direction) can be uniformly bounded. Now the proof of (3.6.10) is just a standard calculation, which we skip. ∎

Lemma 3.28 (Uniform convergence to zero or one) *Almost surely on the event that $\lim_{t\to\infty}|\mathcal{X}_t| \in [0,\infty)$, there exists a set $\Delta \subset \Lambda$ such that*

$$\lim_{t\to\infty}\inf_{i\in\Delta}\mathcal{X}_t(i) = 1 \quad \textit{and} \quad \lim_{t\to\infty}\sup_{i\in\Lambda\setminus\Delta}\mathcal{X}_t(i) = 0. \tag{3.6.11}$$

Proof Imagine that the statement does not hold. Then, by the continuity of sample paths, with positive probability $\lim_{t\to\infty}|\mathcal{X}_t| \in [0,\infty)$ while there exists $0 < \varepsilon < \frac{1}{4}$ such that for every $T > 0$ there exists $t \geq T$ and $i \in \Lambda$ with $\mathcal{X}_t(i) \in (2\varepsilon, 1-2\varepsilon)$. Using Lemma 3.27 and the strong Markov property, it is then not hard to check that with positive probability $\lim_{t\to\infty}|\mathcal{X}_t| \in [0,\infty)$ while there exist infinitely many disjoint time intervals $[t_k, t_k + r]$ and points $i_k \in \Lambda$ such that $\mathcal{X}_t(i_k) \in (\varepsilon, 1-\varepsilon)$ for all $t \in [t_k, t_k + r]$. This contradicts Lemma 3.26.∎

Lemma 3.29 (Convergence to one on a finite nonempty set) *Almost surely on the event $\lim_{t\to\infty}|\mathcal{X}_t| \in (0,\infty)$, the set Δ from Lemma 3.28 is finite and nonempty.*

Proof It is clear that Δ is finite a.s. on the event $\lim_{t\to\infty}|\mathcal{X}_t| < \infty$. Now imagine that Δ is empty. Then, a.s. on the event $\lim_{t\to\infty}|\mathcal{X}_t| > 0$, there exists a random time T such that $\mathcal{X}_t(i) \leq \frac{1}{2}$ for all $t \geq T$ and $i \in \Lambda$. Since $z(1-z) \geq \frac{1}{2}z$ on $[0, \frac{1}{2}]$, it follows that a.s. on the event $\lim_{t\to\infty}|\mathcal{X}_t| > 0$,

$$\int_T^\infty \sum_i \mathcal{X}_t(i)(1-\mathcal{X}_t(i))\mathrm{d}t \geq \frac{1}{2}\int_T^\infty |\mathcal{X}_t| = \infty. \tag{3.6.12}$$

We arrive at a contradiction with Lemma 3.26. ∎

Proof of Proposition 3.25 Let Δ be the random set from Lemma 3.28. We will show that $\Delta = \Lambda$ a.s. on the event $\lim_{t\to\infty}|\mathcal{X}_t| \in (0,\infty)$. In particular, by Lemma 3.29, if Λ is infinite this implies that the event $\lim_{t\to\infty}|\mathcal{X}_t| \in (0,\infty)$ has zero probability. Assume that with positive probability $\lim_{t\to\infty}|\mathcal{X}_t| \in (0,\infty)$ and $\Delta \neq \Lambda$. By Lemma 3.29, Δ is nonempty, and therefore by irreducibility there exist $i \in \Lambda\backslash\Delta$ and $j \in \Delta$ such that $a(i,j) > 0$ or $a(j,i) > 0$. If $a(i,j) > 0$ then by the fact that the counting measure is an invariant measure for the Markov process with jump rates a and by the finiteness of Δ, there must also be an $i' \in \Lambda\backslash\Delta$ and $j' \in \Delta$ such that $a(j',i') > 0$. Thus, there exist $i,j \in \Lambda$ such that $a(j,i) > 0$ and with positive probability $\lim_{t\to\infty}\mathcal{X}_t(i) = 0$, and $\lim_{t\to\infty}\mathcal{X}_t(j) = 1$. It is not hard to see that this violates the evolution in (3.1.3). (We skip the details.) ∎

3.6.2 Convergence to the upper invariant measure

In this section we complete the proof of Theorem 3.4, started in Section 3.1.5, by proving Lemma 3.6. Throughout this section, (Λ, a) is infinite and homogeneous and G is a transitive subgroup of $\mathrm{Aut}(\Lambda, a)$. We fix a reference point $0 \in \Lambda$. We start with two preparatory lemmas.

Lemma 3.30 (Sparse thinning functions) *Assume that $\phi_n \in [0,1]^\Lambda$, $|\phi_n| \to \infty$. Let $\Delta \subset \Lambda$ be finite with $0 \in \Delta$. Then it is possible to choose constants $\lambda_n \to \infty$, finitely supported probability distributions π_n on Λ, and $\{g_i\}_{i\in\mathrm{supp}(\pi_n)}$ with $g_i \in G$ and $g_i(0) = i$ such that the images $\{g_i(\Delta)\}_{i\in\mathrm{supp}(\pi_n)}$ are disjoint, and such that $\lambda_n\pi_n \leq \phi_n$.*

Proof Choose $(g_i)_{i\in\Lambda}$ with $g_i \in G$ such that $g_i(0) = i$. Let $(\xi^{\mathrm{s}}_t)_{t\geq 0}$ be the random walk on Λ that jumps from i to j with the symmetrized jump rates $a^{\mathrm{s}}(i,j) = a(i,j) + a^\dagger(i,j)$. By irreducibility and symmetry, $P^i[\xi^{\mathrm{s}}_t = j] > 0$ for all $t > 0$, $i,j \in \Lambda$. Put

$$\Gamma^\varepsilon_i := \{j \in \Lambda : P^i[\xi^{\mathrm{s}}_1 = j] \geq \varepsilon\} \qquad (i \in \Lambda). \tag{3.6.13}$$

We can choose $\varepsilon > 0$ small enough such that

$$j \notin \Gamma^\varepsilon_i \quad \text{implies} \quad g_i(\Delta) \cap g_j(\Delta) = \emptyset \quad (i,j \in \Lambda). \tag{3.6.14}$$

To see this, set $\delta := \min_{k\in\Delta} P^0[\xi^{\mathrm{s}}_{\frac{1}{2}} = k]$ and put $\varepsilon := \delta^2$. Imagine that $\exists k \in g_i(\Delta) \cap g_j(\Delta)$. Then $P^i[\xi^{\mathrm{s}}_1 = j] \geq P^i[\xi^{\mathrm{s}}_{\frac{1}{2}} = k]P^k[\xi^{\mathrm{s}}_{\frac{1}{2}} = j] \geq \delta^2 = \varepsilon$ by the symmetry of the random walk and homogeneity, and therefore $j \in \Gamma^\varepsilon_i$. Now choose inductively $i_1, i_2, \ldots \in \Lambda$ such that

$$\phi_n \text{ assumes its maximum over } \Lambda\backslash\bigcup_{l=1}^k \Gamma^\varepsilon_{i_l} \text{ in } i_{k+1}. \tag{3.6.15}$$

Then $g_{i_1}(\Delta), g_{i_2}(\Delta), \ldots$ are disjoint by (3.6.14). Since $K := |\Gamma_i^\varepsilon|$ is finite and does not depend on i,

$$\sum_{l=1}^{\infty} \phi_n(i_l) \geq \frac{|\phi_n|}{K}, \tag{3.6.16}$$

and we can choose k_n such that

$$\lambda_n := \sum_{l=1}^{k_n} \phi_n(i_l) \underset{n\to\infty}{\longrightarrow} \infty. \tag{3.6.17}$$

Setting

$$\pi_n := \frac{1}{\lambda_n} \phi_n 1_{\{i_1,\ldots,i_{k_n}\}} \tag{3.6.18}$$

yields λ_n and π_n with the desired properties. ∎

Let $(\xi_t)_{t\geq 0}$ and $(\xi_t^\dagger)_{t\geq 0}$ denote the random walks on Λ that jump from i to j with rates $a(i,j)$ and $a^\dagger(i,j)$, respectively. Then, for any $\Delta \subset \Lambda$, the sets

$$R\Delta := \{i \in \Lambda : P^i[\xi_t \in \Delta] > 0\} \quad \text{and} \quad R^\dagger\Delta := \{i \in \Lambda : P^i[\xi_t^\dagger \in \Delta] > 0\} \quad (t > 0) \tag{3.6.19}$$

of points from which ξ and $\xi^\dagger$ can enter Δ do not depend on $t > 0$. Indeed

$$R\Delta = \big\{i : \exists n \geq 0,\ i_0, \ldots, i_n \text{ s.t. } i_0 = i,\ i_n \in \Delta,\ a(i_{l-1}, i_l) > 0\ \forall l = 1, \ldots, n\big\} \tag{3.6.20}$$

and similarly for $R^\dagger\Delta$. In our next lemma, for $x \in \mathbb{N}^\Lambda$ and $\Delta \subset \Lambda$ we let $x|_\Delta := (x_i)_{i\in\Delta}$ denote the restriction of x to Δ.

Lemma 3.31 (Points from which 0 can be reached) *If μ is a G-homogeneous and nontrivial probability measure on $\mathbb{N}^\Lambda$, then*

$$\mu\big(\{x : x|_{R\{0\}} = 0\}\big) = 0. \tag{3.6.21}$$

Proof Let Y be a $\mathbb{N}^\Lambda$-valued random variable with law μ. We will show that for any $\Delta \subset \Lambda$,

$$P\big[Y|_{R^\dagger R\Delta} = 0\big] = P\big[Y|_{R\Delta} = 0\big]. \tag{3.6.22}$$

Assume that (3.6.22) does not hold. Then there exists an $i \in R^\dagger R\Delta \backslash R\Delta$ such that with positive probability $Y(i) \neq 0$ and $Y|_{R\Delta} = 0$. Since the random walk $(\xi_t^\dagger)_{t\geq 0}$ cannot escape from $R\Delta$ this implies that for any $t > 0$

$$P^i\big[Y(\xi_0^\dagger) \neq 0,\ Y(\xi_s^\dagger) = 0\ \forall s \geq t\big] > 0, \tag{3.6.23}$$

which contradicts the fact that $(Y(\xi_t^\dagger))_{t\geq 0}$ is stationary. This proves (3.6.22). Continuing this process, we see that

$$P\big[Y|_{R\{0\}} = 0\big] = P\big[Y|_{R^\dagger R\{0\}} = 0\big] = P\big[Y|_{RR^\dagger R\{0\}} = 0\big] = \cdots \tag{3.6.24}$$

By irreducibility, the sets $R\{0\}, R^\dagger R\{0\}, RR^\dagger R\{0\}, \ldots$ increase to Λ, and therefore, since μ is nontrivial,

$$P\big[Y|_{R\{0\}} = 0\big] = P\big[Y|_\Lambda = 0\big] = 0. \tag{3.6.25}$$

∎

Proof of Lemma 3.6 For any finite set $\Delta \subset \Lambda$, let X^Δ denote the (a,b,c,d)-braco-process with immediate killing outside Δ. Thus, $X^\Delta_t(i) := 0$ for all $i \in \Lambda\backslash\Delta$ and $t > 0$ and $(X^\Delta_t(i))_{i\in\Delta,\ t\geq 0}$ is the Markov process in $\mathbb{N}^\Delta$ with generator G^Δ given by (compare (3.1.1))

$$\begin{aligned} G^\Delta f(x) := &\sum_{i,j\in\Delta} a(i,j)x(i)\{f(x+\delta_j-\delta_i)-f(x)\} + \sum_{i\in\Delta, j\in\Lambda\backslash\Delta} a(i,j)x(i)\{f(x-\delta_i)-f(x)\} \\ &+b\sum_{i\in\Delta} x(i)\{f(x+\delta_i)-f(x)\} + c\sum_{i\in\Delta} x(i)(x(i)-1)\{f(x-\delta_i)-f(x)\} \\ &+d\sum_{i\in\Delta} x(i)\{f(x-\delta_i)-f(x)\}. \end{aligned} \tag{3.6.26}$$

It is not hard to see that if $\Delta_1, \ldots, \Delta_n$ are disjoint finite sets, then it is possible to couple the processes X and $X^{\Delta_1}, \ldots, X^{\Delta_n}$ in such a way that

$$X_t \leq \sum_{i=1}^n X_t^{\Delta_i} \qquad (t \geq 0) \tag{3.6.27}$$

and the $(X^{\Delta_i})_{i=1,\ldots,n}$ are independent.

Let X denote the (a,b,c,d)-braco-process and assume that $\phi_n \in [0,1]^\Lambda$ satisfy $|\phi_n| \to \infty$. Fix $t > 0$. Assume that $\Delta \subset \Lambda$ is a finite set such that $0 \in \Delta$ and

$$x|_\Delta \neq 0 \quad \Rightarrow \quad P^x[X^\Delta_t(0) > 0] > 0. \tag{3.6.28}$$

Choose λ_n, π_n, and $\{g_i\}_{i\in\mathrm{supp}(\pi_n)}$ as in Lemma 3.30. Then, for deterministic $x \in \mathcal{E}_\gamma(\Lambda)$, we can estimate

$$\begin{aligned} P^x\big[\mathrm{Thin}_{\phi_n}(X_t) = 0\big] &\leq P^x\big[\mathrm{Thin}_{\lambda_n\pi_n}(X_t) = 0\big] \\ &\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^x\big[\mathrm{Thin}_{\lambda_n\pi_n(i)}(X_t^{g_i(\Delta)}(i)) = 0\big] \\ &\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-\lambda_n\pi_n(i)X^\Delta_t(i)}\big] \\ &\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-X^\Delta_t(i)}\big]^{\lambda_n\pi_n(i)}, \end{aligned} \tag{3.6.29}$$

where the $T_{g_i^{-1}}$ are shift operators as in (3.1.17) and we have used that $P[\mathrm{Thin}_\phi(x) = 0] = E[(1-\phi)^x] = E[e^{\langle\log(1-\phi),x\rangle}] \leq E[e^{-\langle\phi,x\rangle}]$ for any $\phi \in [0,1]^\Lambda$, $x \in \mathbb{N}^\Lambda$.

If $\mathcal{L}(X_0)$ is G-homogeneous, then by (3.6.29) and Hölder's inequality,

$$\begin{aligned} P\big[\mathrm{Thin}_{\phi_n}(X_t)=0\big] &\leq \int P[X_0\in \mathrm{d}x] \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-X_t^\Delta(i)}\big]^{\lambda_n\pi_n(i)} \\ &\leq \prod_{i\in\mathrm{supp}(\pi_n)} \Big(\int P[X_0\in \mathrm{d}x]\,P^{T_{g_i^{-1}}x}\big[e^{-X_t^\Delta(i)}\big]^{\lambda_n}\Big)^{\pi_n(i)} \\ &= \int P[X_0\in \mathrm{d}x]\,P^x\big[e^{-X_t^\Delta(0)}\big]^{\lambda_n}, \end{aligned} \tag{3.6.30}$$

and therefore, by (3.6.28) and the fact that $\lambda_n\to\infty$,

$$\limsup_{n\to\infty} P\big[\mathrm{Thin}_{\phi_n}(X_t)=0\big] \leq P\big[X_0|_\Delta=0\big]. \tag{3.6.31}$$

Put

$$\Delta_k := \bigcup_{n=0}^{k} \big\{i : \exists i_0,\ldots,i_n \text{ s.t. } i_0=i,\ i_n=0,\ a(i_{l-1},i_l)>\tfrac{1}{k}\ \forall l=1,\ldots,n\big\}. \tag{3.6.32}$$

Then the Δ_k satisfy (3.6.28) and $\Delta_k\uparrow R\{0\}$ as $k\uparrow\infty$, where $R\{0\}$ is defined in (3.6.20). Therefore, inserting $\Delta=\Delta_k$ in (3.6.31) and taking the limit $k\uparrow\infty$, using Lemma 3.31, we arrive at (3.1.34). ∎

Chapter 4

The contact process seen from a typical infected site

4.1 Introduction and main results

4.1.1 Contact processes on countable groups

The aim of this chapter is to study contact processes on rather general lattices. In particular, we are interested in the way how a certain property of the lattice, namely subexponential growth, influences the behavior of the process.

To keep things reasonably simple, we assume that the lattice Λ is a countably infinite group with group action $(i,j) \mapsto ij$ and unit element 0, also referred to as the origin. Each site $i \in \Lambda$ can be in one of two states: healthy or infected. Infected sites become healthy with *recovery rate* $\delta \geq 0$. An infected site i infects another site j with *infection rate* $a(i,j) \geq 0$. We assume that the infection rates are invariant with respect to the left action of the group, summable, and statisfy a condition that is a bit stronger than irreducibility:

$$\begin{aligned}
&\text{(i)} \quad a(i,j) = a(ki,kj) \qquad\qquad (i,j,k \in \Lambda),\\
&\text{(ii)} \quad |a| := \sum_i a(0,i) < \infty,\\
&\text{(iii)} \quad \textstyle\bigcup_{n\geq 0,\ m\geq 0} A^n A^{-m} = \bigcup_{n\geq 0,\ m\geq 0} A^{-n} A^m = \Lambda,\\
&\qquad\ \text{where } A := \{i \in \Lambda : a(0,i) > 0\}.
\end{aligned} \tag{4.1.1}$$

Here we adopt the convention that sums over i,j,k always run over Λ, unless stated otherwise. For $i \in \Lambda$ and $A, B \subset \Lambda$ we put $AB := \{ij : i \in A,\ j \in B\}$, $iA := \{i\}A$, $Ai := A\{i\}$, $A^{-1} := \{i^{-1} : i \in A\}$, $A^0 := \{0\}$, $A^n := AA^{n-1}$ $(n \geq 1)$, and $A^{-n} := (A^{-1})^n = (A^n)^{-1}$. We let $|A|$ denote the cardinality of A. Note that property (4.1.1) (iii) is equivalent to the statement that for any two sites i,j there exists a site k from which both i and j can be infected, and a set k' that can be infected both from i and from j.

If Λ has a finite symmetric generating set Δ, then the (left) Cayley graph $\mathcal{G} = \mathcal{G}(\Lambda, \Delta)$ associated with Λ and Δ is the graph with vertex set $\mathcal{V}(\mathcal{G}) := \Lambda$ and edges $\mathcal{E}(\mathcal{G}) := \{\{i,j\} : i^{-1}j \in \Delta\}$. Examples of Cayley graphs are the d-dimensional integer lattice $\mathbb{Z}^d$ $(d \geq 1)$ with edges between points at distance one, or the regular tree T_d ($d \geq 2$) in which every vertex has $d+1$ neighbors. On Cayley graphs, one often considers *symmetric nearest-neighbor* infection rates of the form $a(i,j) = \lambda 1_{\{i^{-1}j \in \Delta\}}$, with $\lambda > 0$. In this case, λ is simply referred to as 'the' infection rate.

Let η_t be the set of all infected sites at time $t \geq 0$. Then $\eta = (\eta_t)_{t \geq 0}$ is a Markov process in the space $\mathcal{P}(\Lambda) := \{A : A \subset \Lambda\}$ of all subsets of Λ, called the contact process on Λ with infection rates $a = (a(i,j))_{i,j \in \Lambda}$ and recovery rate δ, or shortly the (Λ, a, δ)-*contact process*. If $\delta > 0$, then by rescaling time we may set $\delta = 1$, so it is customary so assume that $\delta = 1$. If $\delta = 0$ then η is a special case of first-passage percolation (see [Kes86]). We equip $\mathcal{P}(\Lambda) \cong \{0,1\}^\Lambda$ with the product topology and the associated Borel-σ-field $\mathcal{B}(\mathcal{P}(\Lambda))$, and let $\mathcal{P}_{\rm fin}(\Lambda) := \{A \subset \Lambda : |A| < \infty\}$ denote the subspace of finite subsets of Λ.

The contact process can be constructed with the help of Harris' [Har78] *graphical representation*. Let $\omega = (\omega^{\rm r}, \omega^{\rm i})$ be a pair of independent, locally finite random subsets of $\Lambda \times \mathbb{R}$ and $\Lambda \times \Lambda \times \mathbb{R}$, respectively, produced by Poisson point processes with intensity δ and local intensity $(j,k,t) \mapsto a(j,k)$, respectively. This is usually visualized by plotting $\Lambda \times \mathbb{R}$ with Λ horizontally and $\mathbb{R}$ vertically. Points $(i,s) \in \omega^{\rm r}$ and $(j,k,t) \in \omega^{\rm i}$ are marked with a recovery symbol $*$ at (i,s) and an infection arrow from (j,t) to (k,t), respectively. For $C, D \subset \Lambda \times \mathbb{R}$, say that there is a *path* from C to D, denoted by $C \rightsquigarrow D$, if there exist $n \geq 0$, $i_0, \ldots, i_n \in \Lambda$, and $t_0 \leq \cdots \leq t_{n+1}$ with $(i_0, t_0) \in C$ and $(i_n, t_{n+1}) \in D$, such that $\{i_k\} \times [t_k, t_{k+1}] \cap \omega^{\rm r} = \emptyset$ for all $k = 0, \ldots, n$ and $(i_{k-1}, i_k, t_k) \in \omega^{\rm i}$ for all $k = 1, \ldots, n$. Thus, a path must walk upwards in time, may follow arrows, and must avoid recoveries. For given $A \in \mathcal{P}(\Lambda)$ and $t_0 \in \mathbb{R}$, put

$$\eta_t^{A \times \{t_0\}} := \{i \in \Lambda : A \times \{t_0\} \rightsquigarrow (i, t_0 + t)\} \qquad (t \geq 0). \tag{4.1.2}$$

Then $\eta^{A \times \{t_0\}} = (\eta_t^{A \times \{t_0\}})_{t \geq 0}$ is a copy of the (Λ, a, δ)-contact process started in $\eta_0^{A \times \{t_0\}} = A$. For brevity, we put $\eta^A := \eta^{A \times \{0\}}$. The graphical representation couples processes with different initial states in such a way that

$$\eta_t^A \cup \eta_t^B = \eta_t^{A \cup B} \qquad (A, B \in \mathcal{P}(\Lambda),\ t \geq 0). \tag{4.1.3}$$

Define *reversed infection rates* $a^\dagger$ by $a^\dagger(i,j) := a(j,i)$ $(i,j \in \Lambda)$. Say that a is *symmetric* if $a = a^\dagger$. For $A \in \mathcal{P}(\Lambda)$ and $t_0 \in \mathbb{R}$, put

$$\eta_t^{\dagger\, A \times \{t_0\}} := \{i \in \Lambda : (i, t_0 - t) \rightsquigarrow A \times \{t_0\}\} \qquad (t \geq 0). \tag{4.1.4}$$

Then $\eta^{\dagger\, A \times \{t_0\}} = (\eta_t^{\dagger\, A \times \{t_0\}})_{t \geq 0}$ is a copy of the $(\Lambda, a^\dagger, \delta)$-contact process started in $\eta_0^{\dagger\, A \times \{t_0\}} = A$. For brevity, we put $\eta^{\dagger\, A} := \eta^{\dagger\, A \times \{0\}}$. Since for any $s \leq t$ and $A, B \in \mathcal{P}(\Lambda)$, the event

$$\big\{\eta_{u-s}^{A \times \{s\}} \cap \eta_{t-u}^{\dagger\, B \times \{t\}} = \emptyset\big\} = \big\{A \times \{s\} \not\rightsquigarrow B \times \{t\}\big\} \tag{4.1.5}$$

does not depend on $u \in [s,t]$, it follows that the (Λ, a, δ)-contact process and the $(\Lambda, a^\dagger, \delta)$-contact process are dual in the sense that

$$P[\eta_t^A \cap B = \emptyset] = P[A \cap \eta_t^{\dagger\, B} = \emptyset] \qquad (A, B \in \mathcal{P}(\Lambda),\ t \geq 0). \tag{4.1.6}$$

For any $C \subset \Lambda \times \mathbb{R}$, say that $C \rightsquigarrow \infty$ if there is an infinite path with times $t_k \uparrow \infty$ starting in C, and define $-\infty \rightsquigarrow D$ analogously. Instead of $\{(i,s)\} \rightsquigarrow$ and $\rightsquigarrow \{(j,t)\}$, simply write $(i,s) \rightsquigarrow$ and $\rightsquigarrow (j,t)$. We say that the (Λ, a, δ)-contact process η *survives* if

$$\rho(A) := P\big[\eta_t^A \neq \emptyset\ \forall t \geq 0\big] = P[A \times \{0\} \rightsquigarrow \infty] > 0 \tag{4.1.7}$$

for some, and hence for all $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$. If η does not survive then we say that it *dies out*. Set $\delta_{\rm c} = \delta_{\rm c}(\Lambda, a) := \sup\{\delta \geq 0 :$ the (Λ, a, δ)-contact process survives$\}$. Then the (Λ, a, δ)-contact process survives for $\delta < \delta_{\rm c}$ and dies out for $\delta > \delta_{\rm c}$. One has $\delta_{\rm c} \leq |a|$. If Λ is finitely generated, then moreover $\delta_{\rm c} > 0$ (see Section 4.3.4).

4.1.2 Long-time behavior

Since the (Λ, a, δ)-contact process is an attractive spin system, it has an *upper invariant law* $\overline{\nu}$, i.e., an invariant law that is maximal with respect to the stochastic order. It may be constructed as $\overline{\nu} = P[\overline{\eta}_0 \in \cdot\,]$, where

$$\overline{\eta}_t := \{i \in \Lambda : -\infty \rightsquigarrow (i,t)\} \qquad (t \in \mathbb{R}). \tag{4.1.8}$$

Note that

$$P\big[\overline{\eta}_0 \cap A \neq \emptyset\big] = \rho^\dagger(A) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.1.9}$$

where $\rho^\dagger$ denotes the survival probability of the $(\Lambda, a^\dagger, \delta)$-contact process. It is easy to see that $\overline{\nu}$ is nontrivial if and only if the $(\Lambda, a^\dagger, \delta)$-contact process survives. Here, we say that a probability law on $\mathcal{P}(\Lambda)$ is *nontrivial* if it gives zero probability to the empty set.

We say that a probability law μ on $\mathcal{P}(\Lambda)$ is *homogeneous* if μ is shift invariant with respect to the left action of the group, i.e., $\mu(\{iA : A \in \mathcal{A}\}) = \mu(\mathcal{A})$ for all $\mathcal{A} \in \mathcal{B}(\mathcal{P}(\Lambda))$. Using duality, it can be shown that

$$\int \mu(\mathrm{d}A) P[\eta_t^A \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \overline{\nu} \tag{4.1.10}$$

whenever the initial law μ is homogeneous and nontrivial (see [Har76], [Lig85, (VI.2.1)], and [Lig99, (I.1.10)]). Here $\Rightarrow$ denotes weak convergence of probability laws. In particular, (4.1.10) shows that if $\overline{\nu}$ is nontrivial, then it is the only nontrivial homogeneous invariant law.

The long-time behavior for nonhomogeneous initial laws is more subtle and depends on properties of the lattice Λ and the infection rates a, such as subexponential growth.

For the symmetric nearest-neighbor contact process on $\mathbb{Z}^d$ started in a finite initial state, the following picture has been rigorously verified. Either the process dies out in finite time, or in the long run there is a region in space with linearly growing diameter and deterministic limiting shape, such that most of the infected sites lie within this region and there the process is locally in the upper invariant law [BG90]. In particular, it has been shown that the symmetric nearest-neighbor process on $\mathbb{Z}^d$ exhibits *complete convergence*, i.e.,

$$P[\eta_t^A \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \rho(A)\overline{\nu} + (1-\rho(A))\delta_0 \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)). \tag{4.1.11}$$

Note that if complete convergence holds and $\overline{\nu}$ is nontrivial, then by monotonicity, it is the unique nontrivial invariant law. For other contact processes on $\mathbb{Z}^d$ the picture is supposedly similar, provided that the infection rates are symmetric and satisfy an appropriate tail condition. If the infection rates are not symmetric, there is probably still a linearly growing infected region with a limiting shape, but this region may walk out to infinity, so that complete convergence does not hold. (For results in the one-dimensional case, see [Sch86].)

The behavior of the symmetric nearest-neighbor process on regular trees T_d is known to be quite different. Here, there is a second critial value $\delta'_{\rm c} < \delta_{\rm c}$ such that for recovery rates $\delta \in [\delta'_{\rm c}, \delta_{\rm c})$, the process survives globally but not locally, i.e., $\rho(A) > 0$ but $P[\exists T \geq 0 \text{ s.t. } \eta^A_t \cap \{0\} = \emptyset \ \forall t \geq T] = 1$ for $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$. In this regime, there is a multitude of nontrivial invariant measures and complete convergence (obviously) does not hold [Lig99, Section I.4].

One would like to understand which properties of the lattices $\mathbb{Z}^d$ and T_d are responsible for the differences in the behavior of the contact process, and which types of behavior are possible on general lattices Λ. The proofs for $\mathbb{Z}^d$ and T_d use the structure of these lattices in an essential way, and are not easily generalized to other lattices.

It is known that (unoriented) percolation has quite different properties on $\mathbb{Z}^d$ and on T_d. Here, the important property of $\mathbb{Z}^d$, that T_d lacks, is *amenability*. For example, the Burton-Keane proof of the uniqueness of the infinite cluster [BK89] works on any amenable graph. Conversely, it is conjectured that on any nonamenable graph, there exists a range of the percolation parameter for which the infinite cluster is not unique. (See [BS01] and [LP05] some partial results in this direction.)

For our main theorem, we will need to assume that the expected number of infected sites in a contact process grows subexponentially. If Λ is finitely generated, then it turns out that the (Λ, a, δ)-contact process grows subexponentially if a satisfies an exponential moment condition and Λ itself has subexponential growth (see Proposition 4.1 (d) below). Here, by definition, a finitely generated group Λ has *subexponential growth* if

$$\lim_{n\to\infty} \frac{1}{n} \log |\Delta^n| = 0 \tag{4.1.12}$$

for some, and hence for all finite symmetric generating sets Δ. Observe that $\Delta^n = \{i : |i| \leq n\}$ where $|i|$ denotes the distance of i to the origin in the Cayley graph $\mathcal{G}(\Lambda, \Delta)$. Subexponential growth is stronger than amenability. An example of an amenable finitely generated group that does not have subexponential growth is the lamplighter group. (See [MW89, Section 5] for general facts about amenability and subexponential growth, and [LPP96] or [LP05, § 6.1] for a nice exposition of the lamplighter group.)

4.1.3 Results

It turns out that every (Λ, a, δ)-contact process has a well-defined exponential growth rate.

Proposition 4.1 (Exponential growth rate)
(a) *There exists a constant* $r = r(\Lambda, a, \delta) \in [-\delta, |a| - \delta]$ *such that the* (Λ, a, δ)*-contact process satisfies*

$$\lim_{t\to\infty} \tfrac{1}{t} \log E\big[|\eta^A_t|\big] = r \qquad (\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)). \tag{4.1.13}$$

(b) *If the* (Λ, a, δ)*-contact process survives, then* $r \geq 0$.

(c) $r(\Lambda, a, \delta) = r(\Lambda, a^\dagger, \delta)$.

(d) *Assume that Λ is finitely generated. Let Δ be a finite symmetric generating set and let $|j|$ denote the distance of j to the origin in the Cayley graph $\mathcal{G}(\Lambda,\Delta)$. Assume that $\sum_j a(0,j)e^{\varepsilon|j|}<\infty$ for some $\varepsilon>0$ and that Λ has subexponential growth. Then $r\leq 0$.*

The proof of Proposition 4.1 will be given in Sections 4.2.2–4.2.3. Part (a) follows from subadditivity, part (b) is trivial, and part (c) is a consequence of duality. Part (d) follows from some basic large deviation estimates. The exponential moment condition on a appearing in part (d) can perhaps be weakened, but we conjecture that it cannot be dropped altogether. Indeed, it seems plausible that even on $\Lambda=\mathbb{Z}$, the exponential growth rate can be positive if a has a sufficiently heavy tail.

To formulate the main results of this chapter, we must describe the contact process as seen from a 'typical' infected site at a 'typical' late time. Assume that the exponential growth rate r from Proposition 4.1 satisfies $r\leq 0$. Recall the graphical construction of the (Λ,a,δ)-contact process (see Section 4.1.1). Let $(\Omega,\mathcal{F},P)$ be the probability space of the Poisson point processes used in the graphical representation. For $\lambda>r$, we define probability measures $\hat{P}^A_\lambda$ on $\Lambda\times\Omega\times\mathbb{R}_+$ by

$$\hat{P}^A_\lambda(\{i\}\times\{\mathrm{d}\omega\}\times\{\mathrm{d}t\}):=\frac{1}{\pi_\lambda(A)}1_{\{i\,\in\,\eta^A_t(\omega)\}}P(\mathrm{d}\omega)e^{-\lambda t}\mathrm{d}t, \tag{4.1.14}$$

where

$$\pi_\lambda(A):=\int_0^\infty E\big[|\eta^A_t|\big]e^{-\lambda t}\,\mathrm{d}t \qquad (A\in\mathcal{P}_{\rm fin}(\Lambda),\ \lambda>r) \tag{4.1.15}$$

is a normalizing constant. Using the fact that $\lambda>r$, it is easy to see that $0<\pi_\lambda(A)<\infty$, so $\hat{P}^A_\lambda$ is well-defined. Note that the projection of $\hat{P}^A_\lambda$ on $\Omega\times\mathbb{R}_+$ is given by

$$P^A_\lambda(\Lambda\times\{\mathrm{d}\omega\}\times\{\mathrm{d}t\})=\frac{1}{\pi_\lambda(A)}|\eta^A_t(\omega)|P(\mathrm{d}\omega)e^{-\lambda t}\mathrm{d}t \tag{4.1.16}$$

In other words, this projection is is obtained from the product measure $P()e^{-\lambda t}\mathrm{d}t$ on $\Omega\times\mathbb{R}_+$ by size-biasing with the number of infected sites $|\eta^A_t(\omega)|$. Let ι and τ denote the projections on Λ and $\mathbb{R}_+$, respectively. Then, under the law $\hat{P}^A_\lambda$, the random variable η^A_τ describes a size-biased contact process as a 'typical' time τ, and ι is a 'typical' infected site, chosen with equal probabilities from η^A_τ. The law $\hat{P}^A_\lambda[(\iota,\eta^A_\tau)\in\cdot\,]$ is a Campbell law, which is closely related to the more widely known Palm laws. (For the relation between Campbell and Palm laws, see [Eth00, Section 6.4].) The next lemma says that as λ decreases to r, under the laws P^A_λ, the 'typical' time τ tends in probability to ∞. Thus, the limit $\lambda\downarrow r$ corresponds to letting time to infinity.

Lemma 4.2 (Typical times) *For each $\emptyset\neq A\in\mathcal{P}_{\rm fin}(\Lambda)$,*

$$\hat{P}^A_\lambda\big[\tau\geq t\big]\underset{\lambda\downarrow r}{\longrightarrow}1 \qquad (t>0). \tag{4.1.17}$$

Note that $\iota^{-1}\eta_\tau$ is the process η_τ, viewed from the position of the typical infected site ι. The next theorem is the main result of this chapter. Recall the definition of $\overline{\eta}$ in (4.1.8).

Theorem 4.3 (The process seen from a typical infected site) *Assume that the upper invariant measure of the (Λ, a, δ)-contact process is nontrivial and that the exponential growth rate from Proposition 4.1 satisfies $r = 0$. Let $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$. Then*

(a) *One has*

$$\hat{P}^A_\lambda\big[\iota^{-1}\eta^A_\tau \in \cdot\,\big] \underset{\lambda\downarrow 0}{\Longrightarrow} P\big[\overline{\eta}_0 \in \cdot \,\big|\, 0 \in \overline{\eta}_0\big]. \tag{4.1.18}$$

(b) *Moreover,*

$$\hat{P}^A_\lambda\big[\iota^{-1}\eta^A_\tau \cap \Delta = \iota^{-1}\overline{\eta}_\tau \cap \Delta\big] \underset{\lambda\downarrow 0}{\longrightarrow} 1 \qquad (\Delta \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.1.19}$$

and the same holds with $\overline{\eta}_\tau$ replaced by η^Λ_τ.

Note that Theorem 4.3 holds when Λ is a general countable group, but we have only verified that its assumptions are satisfied for certain finitely generated groups (see Proposition 4.1 (d)). We remark that for fixed $\lambda > 0$, it is not at all obvious (and as far as we know *not true*) that the distribution $\hat{P}^A_\lambda[\iota^{-1}\overline{\eta}_\tau \in \cdot\,]$ should be the same as $P[\overline{\eta}_0 \in \cdot \,|\, 0 \in \overline{\eta}_0]$. Thus, none of the statements (4.1.18) and (4.1.19) trivially implies the other one.

As a result of our methods, we can also prove the following fact, which is of some interest on its own.

Proposition 4.4 (Typical particles descend from every surviving site) *Assume that the (Λ, a, δ)-contact process survives and that the exponential growth rate from Proposition 4.1 satisfies $r = 0$. Then*

$$\hat{P}^{\{i\}}_\lambda\big[(j,0) \rightsquigarrow (\iota,\tau) \,\big|\, (j,0) \rightsquigarrow \infty\big] \underset{\lambda\downarrow 0}{\longrightarrow} 1 \qquad (i,j \in \Lambda). \tag{4.1.20}$$

One of the original motivations of the present chapter was to answer the following question. Assuming survival and subexponential growth, is it true that for any $i, j \in \Lambda$,

$$P\big[\exists(k,t) \text{ s.t. } (i,0) \rightsquigarrow (k,t) \rightsquigarrow \infty \text{ and } (j,0) \rightsquigarrow (k,t) \,\big|\, (i,0) \rightsquigarrow \infty, (j,0) \rightsquigarrow \infty\big] = 1 \quad ? \tag{4.1.21}$$

This property may be interpreted as some sort of analogue of the uniqueness of the infinite cluster in (unoriented) percolation. Unfortunately, we do not know how to replace the size-biased law in (4.1.20) by a law conditioned on survival. Question (4.1.21) has been answered positively for oriented percolation on $\mathbb{Z}^d$ in [GH02]. As a further motivation for (4.1.21), we note that in the one-dimensional nearest-neighbor case, a considerably stronger statement holds.

Lemma 4.5 (Coupling of one-dimensional processes) *Consider a $(\mathbb{Z}, a, \delta)$-contact process with $a(i,j) = 0$ for $|i - j| \neq 1$. Assume that the process survives, and assume either $\delta > 0$ or $a(0,1) \wedge a(1,0) > 0$. Then, for any $i, j \in \mathbb{Z}$,*

$$P\big[\inf\{t \geq 0 : \eta^{\{i\}}_t = \eta^{\{j\}}_t\} < \infty \,\big|\, (i,0) \rightsquigarrow \infty,\ (j,0) \rightsquigarrow \infty\big] = 1. \tag{4.1.22}$$

4.1.4 Methods

In this section we describe the main line of our proof of Theorem 4.3 (a). The first ingredient is a chararacterization of the laws $\hat{P}^A_\lambda[\iota^{-1}\eta^A_\tau \in \cdot\,]$ and $P[\overline{\eta}_0 \in \cdot\,|\,0 \in \overline{\eta}_0]$ in terms of the dual $(\Lambda, a^\dagger, \delta)$-contact process $\eta^\dagger$. For simplicity, we only present the argument for $\hat{P}^{\{0\}}_\lambda$. Let $\pi_\lambda(A)$ be the normalizing constant in (4.1.15). Recall the definition of the survival probability ρ in (4.1.7). We write $\overline{\pi}_\lambda$ and $\overline{\rho}$ for the functions π_λ and ρ normalised to one in the point $\{0\}$:

$$\overline{\rho}(A) := \frac{\rho(A)}{\rho(\{0\})} \quad \text{and} \quad \overline{\pi}_\lambda(A) := \frac{\pi_\lambda(A)}{\pi_\lambda(\{0\})}. \tag{4.1.23}$$

We let $\rho^\dagger, \pi^\dagger_\lambda, \overline{\rho}^\dagger$, and $\overline{\pi}^\dagger_\lambda$ denote the analogues of $\rho, \pi_\lambda, \overline{\rho}$, and $\overline{\pi}_\lambda$ for the dual $(\Lambda, a^\dagger, \delta)$-contact process.

Lemma 4.6 (Characterization of laws seen from an infected site)
(a) *One has*

$$\hat{P}^{\{0\}}_\lambda\big[A \cap \iota^{-1}\eta^{\{0\}}_\tau = \emptyset\big] = \overline{\pi}^\dagger_\lambda(A \cup \{0\}) - \overline{\pi}^\dagger_\lambda(A) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ \lambda > r). \tag{4.1.24}$$

(b) *Moreover,*

$$P\big[A \cap \overline{\eta}_0 = \emptyset\,\big|\,0 \in \overline{\eta}_0\big] = \overline{\rho}^\dagger(A \cup \{0\}) - \overline{\rho}^\dagger(A) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)). \tag{4.1.25}$$

It is not hard to see that the law of a $\mathcal{P}(\Lambda)$-valued random variable η is uniquely characterized by all probabilities of the form $P[A \cap \eta = \emptyset]$ with $A \in \mathcal{P}_{\rm fin}(\Lambda)$. Therefore, by Lemma 4.6 and the compactness of $\mathcal{P}(\Lambda) \cong \{0,1\}^\Lambda$, in order to prove Theorem 4.3, it suffices to prove that under the assumptions there, $\overline{\pi}^\dagger_\lambda \to \overline{\rho}^\dagger$ pointwise as $\lambda \downarrow 0$. In order to reduce notation, we reverse the role of η and $\eta^\dagger$. Thus, we will prove that pointwise $\lim_{\lambda\downarrow 0} \overline{\pi}_\lambda = \overline{\rho}$, under the assumptions that the (Λ, a, δ)-contact process survives and its exponential growth rate is zero. (By (4.1.9) and Proposition 4.1 (c), this is equivalent to the $(\Lambda, a^\dagger, \delta)$-contact process having a nontrivial upper invariant law and exponential growth rate zero.)

It is not hard to show (see Section 4.2.1 below) that the (Λ, a, δ)-contact process started in a finite initial state solves the martingale problem for the operator

$$Gf(A) := \sum_{ij} a(i,j) 1_{\{i\in A\}}\{f(A \cup \{j\}) - f(A)\} + \delta \sum_i 1_{\{i \in A\}}\{f(A\backslash\{i\}) - f(A)\}, \tag{4.1.26}$$

with domain $\mathcal{D}(G) := \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, where

$$\mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda)) := \{f : \mathcal{P}_{\rm fin}(\Lambda) \to \mathbb{R} : |f(A)| \leq K|A|^k + M \text{ for some } K, M, k \geq 0\}. \tag{4.1.27}$$

It can be shown in a few lines that ρ is shift invariant, monotone (i.e., $A \subset B$ implies $\rho(A) \leq \rho(B)$), $\rho \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, and

$$G\rho = 0. \tag{4.1.28}$$

Formula (4.1.28) says that ρ is a harmonic function for the (Λ, a, δ)-contact process. It is not hard to see that π_λ shift invariant, monotone, $\pi_\lambda \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, and

$$G\pi_\lambda(A) = \lambda\pi_\lambda(A) - |A| \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ \lambda > r). \tag{4.1.29}$$

As a consequence, one obtains:

Lemma 4.7 (Cluster points of the rescaled expected population size) *The functions $(\overline{\pi}_\lambda)_{\lambda>r}$ are relatively compact with respect to the product topology on $\mathbb{R}^{\mathcal{P}_{\rm fin}(\Lambda)}$. Each pointwise limit*

$$\overline{\pi}_r(A) := \lim_{n\to\infty} \overline{\pi}_{\lambda_n}(A) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)) \tag{4.1.30}$$

along a sequence $\lambda_n \downarrow r$ is shift invariant, monotone in A, satisfies $\overline{\pi}_r \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, and

$$G\overline{\pi}_r = r\overline{\pi}_r. \tag{4.1.31}$$

In particular, if $r = 0$ and the (Λ, a, δ)-contact process survives, it turns out that Lemma 4.7 gives us enough information to determine $\overline{\pi}_0$ uniquely. Combined with the next proposition, Lemma 4.7 shows that $\overline{\pi}_\lambda \to \overline{\rho}$ pointwise as $\lambda \downarrow 0$, thereby completing the proof of Theorem 4.3.

Proposition 4.8 (Shift invariant monotone harmonic functions) *Assume that the (Λ, a, δ)-contact process survives. Assume that $f : \mathcal{P}_{\rm fin}(\Lambda) \to \mathbb{R}$ is shift invariant, monotone, $f(\emptyset) = 0$, $f \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, and $Gf = 0$. Then there exists a constant $c \geq 0$ such that $f = c\rho$.*

We note that if ν is a homogeneous invariant measure for the $(\Lambda, a^\dagger, \delta)$-contact process, then by duality, $f(A) := \nu(\{A : A \cap B \neq \emptyset\})$ defines a shift invariant, monotone, bounded harmonic function f for the (Λ, a, δ)-contact process. Therefore, in view of (4.1.9), Proposition 4.8 is a strengthening of the statement that all homogeneous invariant measures are convex combinations of $\overline{\nu}$ and δ_0.

In order to prove Proposition 4.8, we need one more lemma.

Lemma 4.9 (Eventual domination of finite configurations) *Assume that the (Λ, a, δ)-contact process survives. Then*

$$\lim_{t\to\infty} P\big[\exists i \in \Lambda \text{ s.t. } \eta^A_t \geq iB \,\big|\, \eta^A_t \neq \emptyset\big] = 1 \qquad (A, B \in \mathcal{P}_{\rm fin}(\Lambda),\ A \neq \emptyset). \tag{4.1.32}$$

Formula (4.1.32) says that η exhibits a form of extinction versus unbounded growth. More precisely, either η_t gets extinct or η_t is eventually larger than a suitable shift (depending on η_t) of any finite configuration. We remark that Lemma 4.9 is no longer true if assumption (4.1.1) (iii) is replaced by the weaker assumption that $\{i \in \Lambda : a(0, i) > 0\}$ generates Λ.

Proof of Proposition 4.8 Since the (Λ, a, δ)-contact process solves the martingale problem for G, and $Gf = 0$, the process $f(\eta^A_t)$ is a martingale. In particular:

$$f(A) = E[f(\eta^A_t)] \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ t \geq 0). \tag{4.1.33}$$

Equip Λ with an arbitrary linear ordering, and for $A, B \in \mathcal{P}_{\rm fin}(\Lambda)$, put

$$\hat{\imath}_{A,B} := \begin{cases} \min\{i \in \Lambda : A \geq iB\} & \text{if } \{i \in \Lambda : A \geq iB\} \text{ is nonempty,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.1.34}$$

Since f is monotone and shift invariant, we have, using Lemma 4.9,

$$\begin{aligned} f(A) &= \lim_{t\to\infty} E[f(\eta^A_t)] \\ &\geq \limsup_{t\to\infty} E[1_{\{\exists i \in \Lambda \text{ s.t. } \eta^A_t \geq iB\}} f(\hat{\imath}_{\eta^A_t, B} B)] \\ &= f(B) \limsup_{t\to\infty} P[\exists i \in \Lambda \text{ s.t. } \eta^A_t \geq iB] \geq f(B)\rho(A) \qquad (A, B \in \mathcal{P}_{\rm fin}(\Lambda)). \end{aligned} \tag{4.1.35}$$

In particular, this shows that

$$f(B) \le \frac{f(\{0\})}{\rho(\{0\})} < \infty \qquad (B \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.1.36}$$

hence f is bounded. Now let $A_n, B_m \in \mathcal{P}_{\rm fin}(\Lambda)$ be sequences such that $\rho(A_n) \to 1$ and $\rho(B_n) \to 1$. Then, by (4.1.35),

$$\liminf_{n\to\infty} f(A_n) \ge \liminf_{n\to\infty} f(B_m)\rho(A_n) = f(B_m) \quad \forall m, \tag{4.1.37}$$

and therefore

$$\liminf_{n\to\infty} f(A_n) \ge \limsup_{m\to\infty} f(B_m). \tag{4.1.38}$$

This proves that the limit

$$\lim_{\rho(A_n)\to 1} f(A_n) =: f(\infty) \tag{4.1.39}$$

exists and does not depend on the choice of the sequence A_n with $\rho(A_n) \to 1$. By the Markov property and continuity of the conditional expectation with respect to increasing limits of σ-fields (see Complement 10(b) from [Loe63, Section 29] or [Loe78, Section 32]),

$$\rho(\eta_t^A) = P\big[\eta_s^A \neq 0\ \forall s \ge 0 \,\big|\, \eta_t^A\big] \to 1_{\{\eta_s^A \neq 0\ \forall s \ge 0\}} \quad \text{a.s.} \quad \text{as } t \to \infty. \tag{4.1.40}$$

We conclude that

$$f(A) = \lim_{t\to\infty} E[f(\eta_t^A)] = \rho(A) f(\infty) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.1.41}$$

which shows that f is a scalar multiple of ρ. ∎

4.1.5 Discussion and open problems

Palm and Campbell laws are standard tools in the study of (critical) spatial branching processes. In this context, they can be described by Kallenberg's backward tree technique; see, for example, [Kal77] or [GW91]. In the context of contact processes, it is less obvious that they should be of any use. For example, size-biasing with $|\eta_t^{\{0\}} \cap \{i\}|$ for fixed i and t is just the same as conditioning on $(0,0) \rightsquigarrow (i,t)$. In this case there seems to be no easy way to prove statements about $i^{-1}\eta_t^{\{0\}}$.

However, by looking at the process seen from a randomly chosen infected site rather than a fixed site, i.e., by looking at Campbell laws rather than Palm laws, we can make a connection with the growth of $E[|\eta_t^{\{0\}}|]$ as $t \to \infty$, and in this way obtain a result. A disadvantage of this approach is that one ends up with statements about size-biased laws, where one would probably be more interested in laws conditioned on survival. Nevertheless, it seems that the statements in Theorem 4.3 do catch a phenomenon that depends in a crucial way on a property of the underlying lattice, in this case, subexponential growth.

We next state some open problems and questions, and then comment on them.

1. **Problem** Replace the random time τ in by a deterministic time t and prove the analogue of Theorem 4.3 for $t \to \infty$.

2. **Problem** Study the contact process seen from a typical infected site in case the exponential growth rate is positive.

3. **Problem** Study the contact proces seen from a typical infected site chosen from a process conditioned to survive, instead of size-biased on the number of infected sites.

4. **Problem** Prove (4.1.21) assuming survival and subexponential growth.

5. **Problem** Assuming survival and subexponential growth, prove that conditional on $(i,0) \rightsquigarrow \infty$ and $(j,0) \rightsquigarrow \infty$, eventually most sites in $\eta_t^{\{i\}}$ are also in $\eta_t^{\{j\}}$.

6. **Question** With the same set-up as in the previous problem, is it even true that $\eta_t^{\{i\}}$ and $\eta_t^{\{j\}}$ are eventually equal? (Compare Lemma 4.5.)

7. **Problem** Prove that $\delta_c > 0$ for a contact process on a group Λ that is not finitely generated, for example on the hierarchical group.

8. **Problem** Give an example of a contact process on $\mathbb{Z}$ for which the exponential growth rate is positive.

9. **Question** Assuming that Λ has exponential growth, is it true that the (Λ, a, δ)-contact process survives if and only if $r(\Lambda, a, d) > 0$?

10. **Question** Does survival of the (Λ, a, δ)-contact process imply survival of the $(\Lambda, a^\dagger, \delta)$-contact process?

If one tries to solve Problem 1 in a naive way, by mimicking the techniques in this chapter, it seems one would have to strengthen Proposition 4.1 (a) in the sense that

$$\lim_{t\to\infty} \tfrac{\partial}{\partial t} \log E\big[|\eta_t^A|\big] = r \qquad (\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)). \tag{4.1.42}$$

Then it would follow that each cluster point $\overline{\pi}_\infty$ of the functions $\overline{\pi}_t(A) := E\big[|\eta_t^A|\big]/E\big[|\eta_t^{\{0\}}|\big]$ satisfies $G\overline{\pi}_\infty = 0$. However, (4.1.42) does not simply follow from subadditivity and seems hard to establish in general. Even random times τ that are uniformly distributed on intervals $[0,T]$ seem difficult to treat, since they would require that $\lim_{T\to\infty} \frac{\partial}{\partial T} \log \int_0^T E\big[|\eta_t^A|\big]\mathrm{d}t = r$.

In order to solve Problem 2, generalizing Proposition 4.8, one would like to show that the equation $G\overline{\pi}_r = r\overline{\pi}_r$ has a unique shift invariant, monotone solution $\overline{\pi}_r$ with $\overline{\pi}_r(\emptyset) = 0$ and $\overline{\pi}_r(\{0\}) = 1$ (perhaps also using that $\overline{\pi}_r$ is subadditive).

Problems 3–5 and Question 6 have been discussed before. The difficulty is to replace size-biased laws by laws conditioned on survival in statements like Proposition 4.4. Although size-biasing and conditioning are asymptotically equivalent in a 'local' sense (see Proposition 4.14 below), this does not seem easy. Note that if (4.1.21) holds for the $(\Lambda, a^\dagger, \delta)$-contact process, then the limit law in Theorem 4.3 (a) may also be written as $P[\hat\eta_0 \in \cdot \mid -\infty \rightsquigarrow (0,0)]$, where $\hat\eta_t := \{i \in \Lambda : \exists (j,s) \text{ s.t. } -\infty \rightsquigarrow (j,s) \rightsquigarrow (0,0) \text{ and } (j,s) \rightsquigarrow (i,t)\}$ $(t \in \mathbb{R})$. This construction is similar to Kallenberg's backward tree technique, and also somewhat reminiscent of the construction of the the second lowest extremal invariant measure of the contact process in [SS97, SS99].

Problem 7 seems interesting, since the hierarchical group has found applications in population biology, and the usual comparison with one-dimensional oriented percolation cannot work here.

Problem 8 and Question 9 are naturally motivated by Proposition 4.1 (d). Related to Question 9 is the more general question: what does the behavior of $E[|\eta_t|]$ for $t \to \infty$ tell us about survival? Especially for critical processes, it seems conceivable that $\lim_{t\to\infty} E[|\eta_t|] = \infty$ while the process dies out.

Related to this is Question 10, which has been asked before for branching-coalescing particle systems in [AS05]. For symmetric processes or for processes on abelian groups, the answer is obviously positive, but in general (Λ, a) and $(\Lambda, a^\dagger)$ need not be isomorphic. However, in formula (4.2.14) below, it is shown that $E[|\eta_t^{\{0\}}|] = E[|\eta_t^{\dagger\,\{0\}}|]$ for all $t \geq 0$. (On the other hand, dropping the assumption that Λ is a group, by considering contact processes on transitive graphs that are not unimodular, it is easy to construct examples where $E[|\eta_t^{\{0\}}|] \neq E[|\eta_t^{\dagger\,\{0\}}|]$ and where η survives but $\eta^\dagger$ dies out.) An example of a model on $\mathbb{Z}^2$ where nontriviality of the upper invariant law and survival are not equivalent is the NEC model due to A. Toom [BG85, DLSS91].

Related to Question 10 (compare also Question 6) is the following question: is it always true that $\inf\{t \geq 0 : \eta_t^{\{0\}} \subset \overline{\eta}_t\}$ is a.s. finite? Note that if the answer is positive, then extinction of the $(\Lambda, a^\dagger, \delta)$-contact process implies extinction of the (Λ, a, δ)-contact process, since in this case $\overline{\eta} \equiv 0$.

4.1.6 Outline

Section 4.2 is devoted to the proof of Theorem 4.3 (a). In Section 4.2.1 we prove that contact processes started in finite initial states solve the martingale problem for the operator G in (4.1.26). We establish Proposition 4.1 (a)–(c) in Section 4.2.2, and part (d) in Section 4.2.3. In Section 4.2.4, we establish Lemmas 4.2 and 4.6. In Section 4.2.5, we prove basic facts about the functions ρ and π_λ; in particular, formulas (4.1.28) and (4.1.29), and Lemma 4.7. In Section 4.2.6, we prove Lemma 4.9, thereby completing the proof of Theorem 4.3 in the case $A = \{0\}$. In Section 4.2.7 we show how the arguments may be generalized to arbitrary $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$.

Section 4.3 contains proofs of all results that are not directly needed for Theorem 4.3 (a). In Section 4.3.1, we prove that size-biasing and conditioning on survival are equivalent in a 'local' sense. Section 4.3.2 contains the proofs of Theorem 4.3 (b) and Proposition 4.4. Section 4.3.3 contains the proof of Lemma 4.5. For completeness, we prove in Section 4.3.4 the fact mentioned in the text that $\delta_{\rm c} > 0$ whenever Λ is finitely generated.

Acknowledgements The author thanks Geoffrey Grimmmett, Olle Häggström, Russel Lyons, and Roberto Schonmann for useful email conversations about the contact process, oriented percolation, and amenability.

4.2 The law seen from a typical particle

4.2.1 A martingale problem

In this section we prove that the (Λ, a, δ)-contact process started in finite initial states solves the martingale problem for the operator G in (4.1.26)–(4.1.27).

Proposition 4.10 (Martingale problem and moment estimate) *For each $f \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$ and $A \in \mathcal{P}_{\rm fin}(\Lambda)$, the process*

$$M_t := f(\eta_t^A) - \int_0^t Gf(\eta_s^A)\mathrm{d}s \qquad (t \geq 0) \tag{4.2.1}$$

is a martingale with respect to the filtration generated by η^A. Moreover, setting $z^{\langle k\rangle} := \prod_{i=0}^{k-1}(z+i)$, one has

$$E\big[|\eta_t^A|^{\langle k\rangle}\big] \leq |A|^{\langle k\rangle} e^{k(|a|-\delta)t} \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ k \geq 1,\ t \geq 0). \tag{4.2.2}$$

Proof The proof of [AS05, Proposition 8] can in a straightforward way be adapted to the present set-up. Set $f_k(A) := |A|^{\langle k\rangle}$. Then

$$\begin{aligned} Gf_k(A) &= \sum_{ij} a(i,j) 1_{\{i\in A\}} 1_{\{j\notin A\}} \{(|A|+1)^{\langle k\rangle} - |A|^{\langle k\rangle}\} + \delta \sum_i 1_{\{i\in A\}} \{(|A|-1)^{\langle k\rangle} - |A|^{\langle k\rangle}\}, \\ &\leq (|a|-\delta)|A|\{(|A|+1)^{\langle k\rangle} - |A|^{\langle k\rangle}\} = k(|a|-\delta)|A|^{\langle k\rangle}. \end{aligned} \tag{4.2.3}$$

Define stopping times $\tau_N := \inf\{t \geq 0 : |\eta_t^A| \geq N\}$. The stopped process $(\eta^A_{t\wedge\tau_N})_{t\geq 0}$ has bounded jump rates, and therefore standard theory tells us that for each $N \geq 1$ and $f \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, the process

$$M_t^N := f(\eta^A_{t\wedge\tau_N}) - \int_0^{t\wedge\tau_N} Gf(\eta_s^A)\mathrm{d}s \qquad (t \geq 0) \tag{4.2.4}$$

is a martingale. Moreover, it easily follows from (4.2.3) that

$$E\big[|\eta^A_{t\wedge\tau_N}|^{\langle k\rangle}\big] \leq |A|^{\langle k\rangle} e^{k(|a|-\delta)t} \qquad (k \geq 1,\ t \geq 0). \tag{4.2.5}$$

It is easy to see that $f \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$ implies $Gf \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$. Using this fact and (4.2.5) for some sufficiently high k (depending on f), one can show that for fixed $t \geq 0$, the random variables $(M_t^N)_{N\geq 1}$ are uniformly integrable. Therefore, letting $N \to \infty$ in (4.2.4), one finds that the process in (4.2.1) is a martingale. Letting $N \to \infty$ in (4.2.5) yields (4.2.2). ∎

4.2.2 The exponential growth rate

In this section we prove Proposition 4.1 (a)–(c).

Proof of Proposition 4.1 (a) By a slight abuse of notation, let us write (compare (4.1.15))

$$\pi_t(A) := E\big[|\eta_t^A|\big] \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ t \geq 0). \tag{4.2.6}$$

We start by showing that

$$\pi_{s+t}(\{0\}) \le \pi_s(\{0\})\pi_t(\{0\}) \qquad (s,t \ge 0). \tag{4.2.7}$$

By (4.1.3),

$$E\big[|\eta_t^A|\big] = E\Big[\big|\bigcup_{i\in A}\eta_t^{\{i\}}\big|\Big] \le \sum_{i\in A} E\big[|\eta_t^{\{i\}}|\big] = |A|E\big[|\eta_t^{\{0\}}|\big], \tag{4.2.8}$$

where in the last step we have used shift invariance. As a consequence,

$$\pi_{s+t}(\{0\}) = \int P[\eta_s^{\{0\}} \in \mathrm{d}A]E\big[|\eta_t^A|\big] \le \int P[\eta_s^{\{0\}} \in \mathrm{d}A]|A|E\big[|\eta_t^{\{0\}}|\big] = \pi_s(\{0\})\pi_t(\{0\}). \tag{4.2.9}$$

This proves (4.2.7). It follows that $t \mapsto \log \pi_t(\{0\})$ is subadditive and therefore, by [Lig99, Theorem B.22], the limit

$$\lim_{t\to\infty} \tfrac{1}{t}\log \pi_t(\{0\}) =: r \in [-\infty,\infty] \tag{4.2.10}$$

exists. By monotonicity and (4.2.8),

$$\pi_t(\{0\}) \le \pi_t(A) \le |A|\pi_t(\{0\}) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)). \tag{4.2.11}$$

Taking logarithms, dividing by t, and letting $t \to \infty$ we arrive at (4.1.13). Since η can be bounded from below by a simple death process and from above by a branching process (see (4.2.15) below), one has

$$e^{-\delta t} \le E\big[|\eta_t^{\{0\}}|\big] \le e^{(|a|-\delta)t} \qquad (t \ge 0), \tag{4.2.12}$$

which implies that $-\delta \le r \le |a| - \delta$. ∎

Proof of Proposition 4.1 (b) If the (Λ, a, δ)-contact process survives, then

$$\pi_t(\{0\}) \ge P[\eta_t^{\{0\}} \neq 0] \underset{t\to\infty}{\longrightarrow} P[\eta_s^{\{0\}} \neq 0\ \forall s \ge 0] > 0, \tag{4.2.13}$$

which implies that $r \ge 0$. ∎

Proof of Proposition 4.1 (c) By duality (formula (4.1.6)) and shift invariance,

$$\begin{aligned} E\big[|\eta_t^{\{0\}}|\big] &= \sum_i P\big[\eta_t^{\{0\}} \cap \{i\} \neq \emptyset\big] = \sum_i P\big[\{0\} \cap \eta_t^{\dagger\,\{i\}} \neq \emptyset\big] \\ &= \sum_i P\big[\{i^{-1}\} \cap \eta_t^{\dagger\,\{0\}} \neq \emptyset\big] = E\big[|\eta_t^{\dagger\,\{0\}}|\big], \end{aligned} \tag{4.2.14}$$

which implies that $r(\Lambda, a, \delta) = r(\Lambda, a^\dagger, \delta)$. ∎

4.2.3 Subexponential growth

Proof of Proposition 4.1 (d) Consider a branching process on Λ, started with one particle in the origin, where a particle at i produces a new particle at j with rate $a(i,j)$, and each particle dies with rate δ. Let $B_t(i)$ denote the number of particles at site $i\in\Lambda$ and time $t\geq 0$. It is not hard to see that $\eta^{\{0\}}$ and B may be coupled such that

$$1_{\eta^{\{0\}}_t} \leq B_t \qquad (t\geq 0). \tag{4.2.15}$$

Let $(\xi_t)_{t\geq 0}$ be a random walk on Λ that jumps from i to j with rate $a(i,j)$, started in $\xi_0=0$. Then it is not hard to see that (compare [Lig99, Proposition I.1.21])

$$E[B_t(i)] = P[\xi_t=i]e^{(|a|-\delta)t} \qquad (i\in\Lambda,\ t\geq 0). \tag{4.2.16}$$

Let $\gamma>0$ be a constant, to determined later. It follows from (4.2.15) and (4.2.16) that

$$\begin{aligned} E\big[|\eta^{\{0\}}_t|\big] &\leq \sum_i \big(1\wedge P[\xi_t=i]e^{(|a|-\delta)t}\big) \\ &= \big|\{i\in\Lambda: |i|\leq\gamma t\}\big| + P[|\xi_t|>\gamma t]e^{(|a|-\delta)t} \qquad (t\geq 0). \end{aligned} \tag{4.2.17}$$

Let $(Y_i)_{i\geq 1}$ be i.i.d. $\mathbb{N}$-valued random variables with $P[Y_i=k]=\frac{1}{|a|}\sum_{j:\ |j|=k}a(0,j)$ $(k\geq 0)$, let N be a Poisson-distributed random variable with mean $|a|$, independent of the $(Y_i)_{i\geq 1}$, and let $(X_m)_{m\geq 1}$ be i.i.d. random variables with law $P[X_m\in\cdot\,]=P[\sum_{i=1}^N Y_i\in\cdot\,]$. Since the random walk ξ makes jumps whose sizes are distributed in the same way as the Y_i, and the number of jumps per unit of time is Poisson distributed with mean $|a|$, it follows that

$$P[|\xi_t|>\gamma t] \leq P\Big[\frac{1}{\lceil t\rceil}\sum_{m=1}^{\lceil t\rceil} X_m > \gamma\frac{t}{\lceil t\rceil}\Big] \qquad (t>0), \tag{4.2.18}$$

where $\lceil t\rceil$ denotes t rounded up to the next integer. By our assumptions,

$$E\big[e^{\varepsilon X_m}\big] = E\big[e^{\varepsilon\sum_{i=1}^N Y_k}\big] = e^{-|a|}\sum_{n=0}^\infty \frac{|a|^n}{n!}E\big[e^{\varepsilon Y_1}\big]^n = e^{-|a|(1-E[e^{\varepsilon Y_1}])} < \infty, \tag{4.2.19}$$

for some $\varepsilon>0$. Therefore, by [DZ98, Theorem 2.2.3 and Lemma 2.2.20], for each $R>0$ there exists a $\gamma>0$ and $K<\infty$ such that

$$P\Big[\frac{1}{n}\sum_{m=1}^n X_m > \gamma\Big] \leq Ke^{-nR} \qquad (n\geq 1). \tag{4.2.20}$$

Choosing γ such that (4.2.20) holds for some $R>|a|-\delta$ yields, by (4.2.18)

$$\lim_{t\to\infty} P\big[|\xi_t|>\gamma t\big]e^{(|a|-\delta)t} = 0. \tag{4.2.21}$$

Inserting this into (4.2.17) we find that the exponential growth rate $r=r(\Lambda,a,\delta)$ satisfies

$$r \leq \limsup_{t\to\infty}\frac{1}{t}\log\big|\{i\in\Lambda: |i|\leq\gamma t\}\big| = 0, \tag{4.2.22}$$

where we have used that Λ is subexponential. ∎

4.2.4 Duality and Campbell laws

Proof of Lemma 4.6 (a) This follows by writing

$$
\begin{aligned}
\hat P^{\{0\}}_\lambda\big[A\cap\iota^{-1}\eta^{\{0\}}_\tau=\emptyset\big] &\overset{(1)}{=} \pi_\lambda(\{0\})^{-1}\sum_i\int_0^\infty P\big[i\in\eta^{\{0\}}_t,\ A\cap i^{-1}\eta^{\{0\}}_t=\emptyset\big]e^{-\lambda t}\,\mathrm{d}t\\
&\overset{(2)}{=} \pi_\lambda(\{0\})^{-1}\sum_i\int_0^\infty P\big[0\in\eta^{\{i^{-1}\}}_t,\ A\cap\eta^{\{i^{-1}\}}_t=\emptyset\big]e^{-\lambda t}\,\mathrm{d}t\\
&\overset{(3)}{=} \pi_\lambda(\{0\})^{-1}\sum_j\int_0^\infty \Big\{P\big[(A\cup\{0\})\cap\eta^{\{j\}}_t\neq\emptyset\big]-P\big[A\cap\eta^{\{j\}}_t\neq\emptyset\big]\Big\}e^{-\lambda t}\,\mathrm{d}t\\
&\overset{(4)}{=} \pi^\dagger_\lambda(\{0\})^{-1}\sum_j\int_0^\infty \Big\{P\big[\eta^{\dagger\,A\cup\{0\}}_t\cap\{j\}\neq\emptyset\big]-P\big[\eta^{\dagger\,A}_t\cap\{j\}\neq\emptyset\big]\Big\}e^{-\lambda t}\,\mathrm{d}t\\
&\overset{(5)}{=} \pi^\dagger_\lambda(\{0\})^{-1}\int_0^\infty \Big\{E\big[|\eta^{\dagger\,A\cup\{0\}}_t|\big]-E\big[|\eta^{\dagger\,A}_t|\big]\Big\}e^{-\lambda t}\,\mathrm{d}t\\
&\overset{(6)}{=} \pi^\dagger_\lambda(\{0\})^{-1}\big\{\pi^\dagger_\lambda(A\cup\{0\})-\pi^\dagger_\lambda(A)\big\}\overset{(7)}{=}\overline{\pi}^\dagger_\lambda(A\cup\{0\})-\overline{\pi}^\dagger_\lambda(A).
\end{aligned}
\tag{4.2.23}
$$

Here, in step (2) we have used shift invariance, in step (3) we have changed the summation order and used that $\{0\in\eta^{\{j\}}_\tau,\ A\cap\eta^{\{j\}}_\tau=\emptyset\}=\{(A\cup\{0\})\cap\eta^{\{j\}}_\tau\neq\emptyset\}\backslash\{A\cap\eta^{\{j\}}_\tau\neq\emptyset\}$, and in step (4) we have used duality (formula (4.1.6)) and formula (4.2.14). ∎

Proof of Lemma 4.6 (b) We have

$$
\begin{aligned}
P\big[A\cap\overline{\eta}_0=\emptyset\,\big|\,0\in\overline{\eta}_0\big] &\overset{(1)}{=} P\big[0\in\overline{\eta}_0\big]^{-1}P\big[0\in\overline{\eta}_0,\ A\cap\overline{\eta}_0=\emptyset\big]\\
&\overset{(2)}{=} P\big[\{0\}\cap\overline{\eta}_0\neq\emptyset\big]^{-1}\big\{P\big[(A\cup\{0\})\cap\overline{\eta}_0\neq\emptyset\big]-P\big[A\cap\overline{\eta}_0\neq\emptyset\big]\big\}\\
&\overset{(3)}{=} \rho^\dagger(\{0\})\big\{\rho^\dagger(A\cup\{0\})-\rho^\dagger(A)\big\}\overset{(4)}{=}\overline{\rho}^\dagger(A\cup\{0\})-\overline{\rho}^\dagger(A),
\end{aligned}
\tag{4.2.24}
$$

where in step (3) we have used (4.1.9). ∎

As a preparation for the proof of Lemma 4.2, we prove:

Lemma 4.11 (Expected population size) *One has* $\lim_{\lambda\downarrow r}\pi_\lambda(A)=\infty$ *for all* $\emptyset\neq A\in\mathcal{P}_{\rm fin}(\Lambda)$.

Proof We start with the case $A=\{0\}$. Recall that Proposition 4.1 (a) is a consequence of the subadditivity of the function $t\mapsto\log E[|\eta^{\{0\}}_t|]$. In fact, subadditivity gives us a little more. By [Lig99, Theorem B.22],

$$
\lim_{t\to\infty}\tfrac{1}{t}\log E\big[|\eta^{\{0\}}_t|\big]=\inf_{t>0}\tfrac{1}{t}\log E\big[|\eta^{\{0\}}_t|\big]=r,
\tag{4.2.25}
$$

where $r=r(\Lambda,a,\delta)\in[-\delta,|a|-\delta]$ is the exponential growth rate. Formula (4.2.25) says that $E[|\eta^{\{0\}}_t|]=e^{r_t t}$ where $\lim_{t\to\infty}r_t=\inf_{t>0}r_t=r$. Thus, for every $\varepsilon>0$, there exists a $T_\varepsilon<\infty$ such that

$$
e^{rt}\le E\big[|\eta^{\{0\}}_t|\big]\le e^{(r+\varepsilon)t}\qquad(t\ge T_\varepsilon).
\tag{4.2.26}
$$

It follows from the lower bound in (4.2.26) and monotone convergence that

$$\lim_{\lambda\downarrow r}\pi_\lambda(\{0\})=\int_0^\infty E\big[|\eta_t^{\{0\}}|\big]\,e^{-rt}\,\mathrm{d}t=\infty. \tag{4.2.27}$$

The generalization to arbitrary $\emptyset\neq A\in\mathcal{P}_{\mathrm{fin}}(\Lambda)$ is immediate, since π_λ is monotone. ∎

Proof of Lemma 4.2 By Lemma 4.11,

$$\hat P_\lambda^A\big[\tau<t\big]=\frac{\int_0^t E\big[|\eta_s^A|\big]\,e^{-\lambda s}\,\mathrm{d}s}{\int_0^\infty E\big[|\eta_s^A|\big]\,e^{-\lambda s}\,\mathrm{d}s}\leq\frac{\int_0^t E\big[|\eta_s^A|\big]\,e^{-rs}\,\mathrm{d}s}{\pi_\lambda(A)}\underset{\lambda\downarrow r}{\longrightarrow}0. \tag{4.2.28}$$

for any $t>0$. ∎

4.2.5 Harmonic functions

In this section we prove formulas (4.1.28) and (4.1.29), and Lemma 4.7.

Proof of (4.1.28) The shift invariance and monotonicity of ρ follow from the corresponding properties of the contact process. Since ρ is bounded, obviously $\rho\in\mathcal{S}(\mathcal{P}_{\mathrm{fin}}(\Lambda))$. Since η^A solves the martingale problem for G, for any $f\in\mathcal{S}(\mathcal{P}_{\mathrm{fin}}(\Lambda))$, one has

$$\int_0^t E[Gf(\eta_s^A)]\mathrm{d}s=E[f(\eta_t^A)]-f(A)\qquad(A\in\mathcal{P}_{\mathrm{fin}}(\Lambda)), \tag{4.2.29}$$

and therefore

$$Gf(A)=\lim_{t\to 0}t^{-1}\big\{E[f(\eta_t^A)]-f(A)\big\}\qquad(A\in\mathcal{P}_{\mathrm{fin}}(\Lambda)). \tag{4.2.30}$$

By the Markov property,

$$\rho(\eta_t^A)=E\big[\eta_s^A\neq 0\ \forall s\geq 0\,\big|\,\eta_t^A\big]=E\big[\eta_s^A\neq 0\ \forall s\geq 0\,\big|\,\mathcal{F}_t^A\big], \tag{4.2.31}$$

where $(\mathcal{F}_t^A)_{t\geq 0}$ denotes the filtration generated by η^A. It follows that $\rho(\eta_t^A)$ is a martingale, and therefore, by (4.2.30), $G\rho=0$. ∎

Proof of (4.1.29) The shift invariance and monotonicity of π_λ follow from the corresponding properties of the contact process. It follows from (4.1.3) that $\pi_\lambda(A)\leq\pi_\lambda(\{0\})|A|$, which shows that $\pi_\lambda\in\mathcal{S}(\mathcal{P}_{\mathrm{fin}}(\Lambda))$. Moreover,

$$\begin{aligned}
&t^{-1}\big\{E[\pi_\lambda(\eta_t^A)]-\pi_\lambda(A)\big\}\\
&=t^{-1}\int_0^\infty\Big\{E\big[|\eta_{t+s}^A|\big]-E\big[|\eta_s^A|\big]\Big\}e^{-\lambda s}\,\mathrm{d}s\\
&=t^{-1}\Big\{\int_t^\infty E\big[|\eta_s^A|\big]e^{-\lambda(s-t)}\,\mathrm{d}s-\int_0^\infty E\big[|\eta_s^A|\big]e^{-\lambda s}\,\mathrm{d}s\Big\}\\
&=t^{-1}(e^{\lambda t}-1)\int_0^\infty E\big[|\eta_s^A|\big]e^{-\lambda s}\,\mathrm{d}s-e^{\lambda t}t^{-1}\int_0^t E\big[|\eta_s^A|\big]e^{-\lambda s}\,\mathrm{d}s.
\end{aligned} \tag{4.2.32}$$

Letting $t\to 0$, using (4.2.30), it follows that

$$G\pi_\lambda(A)=\lambda\pi_\lambda(A)-|A|\qquad(A\in\mathcal{P}_{\mathrm{fin}}(\Lambda),\ \lambda>r), \tag{4.2.33}$$

as desired. ∎

Proof of Lemma 4.7 It follows from (4.1.24) that $\overline{\pi}_\lambda(A) \le |A|$, which shows that the functions $(\overline{\pi}_\lambda)_{\lambda > r}$ are relatively compact, and each pointwise limit $\overline{\pi}_\infty$ along a sequence $\lambda_n \downarrow r$ satisfies $\overline{\pi}_\infty \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$. Since each $\overline{\pi}_{\lambda_n}$ is shift invariant an monotone, the same is true for $\overline{\pi}_\infty$. If $f_n, f \in \mathcal{S}(\mathcal{P}_{\rm fin}(\Lambda))$, $f_n \to f$ pointwise, and the f_n are uniformly bounded on sets of the form $\{A \in \mathcal{P}_{\rm fin}(\Lambda) : |A| \le K\}$, then it is not hard to see that pointwise

$$\lim_{n\to\infty} Gf_n = Gf. \tag{4.2.34}$$

Applying this to the functions $\overline{\pi}_{\lambda_n}$, which satisfy the uniform bound $\overline{\pi}_{\lambda_n}(A) \le |A|$, using (4.1.29) and Lemma 4.11, we find that

$$G\overline{\pi}_\infty(A) = \lim_{n\to\infty} \frac{\lambda_n \pi_{\lambda_n}(A) - |A|}{\pi_{\lambda_n}(\{0\})} = \lim_{n\to\infty} \lambda_n \overline{\pi}_{\lambda_n}(A) - \frac{|A|}{\pi_{\lambda_n}(\{0\})} = r\overline{\pi}_r(A) \qquad (A \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.2.35}$$

as required. ∎

4.2.6 Eventual domination of finite configurations

In this section we prove Lemma 4.9. We start with two preparatory lemmas.

Lemma 4.12 (Local creation of finite configurations) *For each $B \in \mathcal{P}_{\rm fin}(\Lambda)$ and $t > 0$, there exists a finite $\Delta \subset \Lambda$ and $j \in \Lambda$ such that*

$$\varepsilon := P\big[\eta_t^{\{0\}} \supset jB \ \textit{and}\ \eta_s^{\{0\}} \subset \Delta\ \forall 0 \le s \le t\big] > 0. \tag{4.2.36}$$

Proof It follows from assumption (4.1.1) (iii) that there exists a site $j^{-1} \in \Lambda$ with $P\big[\eta_t^{\{j^{-1}\}} \supset B\big] > 0$, and therefore $P\big[\eta_t^{\{0\}} \supset jB\big] > 0$. Since $\bigcup_{0 \le s \le t} \eta_s^{\{0\}}$ is a.s. finite, we can choose a finite but large enough Δ such that (4.2.36) holds. ∎

Lemma 4.13 (Domination of finite configurations) *For each $B \in \mathcal{P}_{\rm fin}(\Lambda)$, $t > 0$, and $A_n \in \mathcal{P}_{\rm fin}(\Lambda)$ satisfying $\lim_{n\to\infty} |A_n| = \infty$, one has*

$$\lim_{n\to\infty} P[\exists i \in \Lambda \ \textit{s.t.}\ \eta_t^{A_n} \ge iB] = 1. \tag{4.2.37}$$

Proof Let Δ, j, and ε be as in Lemma 4.12. We can find $\tilde{A}_n \subset A_n$ such that $|\tilde{A}_n| \to \infty$ as $n \to \infty$, and for fixed n, the sets $(k\Delta)_{k \in \tilde{A}_n}$ are disjoint. It follows that

$$\begin{aligned} &P[\exists i \in \Lambda \text{ s.t. } \eta_t^{A_n} \ge iB] \\ &\quad \ge 1 - \prod_{k \in \tilde{A}_n} \big(1 - P\big[\eta_t^{\{k\}} \supset kjB \text{ and } \eta_s^{\{k\}} \subset k\Delta\ \forall 0 \le s \le t\big]\big) \\ &\quad = 1 - (1-\varepsilon)^{|\tilde{A}_n|} \underset{n\to\infty}{\longrightarrow} 1, \end{aligned} \tag{4.2.38}$$

where we have used (4.2.36) and the fact that events concerning the graphical representation in disjoint parts of space are independent. ∎

Proof of Lemma 4.9 If $\delta = 0$, then obviously $\lim_{t\to\infty} |\eta_t^A| = \infty$ a.s. If $\delta > 0$, then it is easy to see that $\inf\{\rho(A) : |A| \leq M\} < 1$ for all $M < \infty$. Therefore, by (4.1.40),

$$\eta_t^A = \emptyset \text{ for some } t \geq 0 \quad \text{or} \quad |\eta_t^A| \underset{t\to\infty}{\longrightarrow} \infty \qquad \text{a.s.} \tag{4.2.39}$$

Fix $\emptyset \neq B \in \mathcal{P}_{\rm fin}(\Lambda)$ and set $\psi_t(A) := P[\exists i \in \Lambda \text{ s.t. } \eta_t^A \geq iB] \quad (A \in \mathcal{P}_{\rm fin}(\Lambda),\ t \geq 0)$. Then, for each $t > 0$,

$$\lim_{T\to\infty} P[\exists i \in \Lambda \text{ s.t. } \eta_T^A \supset iB] = \lim_{T\to\infty} E[\psi_t(\eta_{T-t}^A)] = \rho(A), \tag{4.2.40}$$

where we have used Lemma 4.13 and (4.2.39). ∎

4.2.7 Generalization to arbitrary initial states

In this section, we show how the proof of Theorem 4.3 (a) must be adapted to cover general initial states $\emptyset \neq A \in \mathcal{P}(\Lambda)$.

Proof of Theorem 4.3 (a) for general initial states For $A, B \in \mathcal{P}_{\rm fin}(\Lambda)$ with $A \neq \emptyset$, we observe that $i \in BA^{-1} \Leftrightarrow B \cap iA \neq \emptyset$, and therefore

$$|BA^{-1}| = \sum_i 1_{\{B \cap iA \neq \emptyset\}}. \tag{4.2.41}$$

We define

$$\pi_{A,\lambda}(B) := \int_0^\infty E\big[|\eta_t A^{-1}|\big] e^{-\lambda t}\,\mathrm{d}t \quad \text{and} \quad \overline{\pi}_{A,\lambda}(B) := \frac{\pi_{A,\lambda}(B)}{\pi_{A,\lambda}(\{0\})}, \tag{4.2.42}$$

and let $\pi^\dagger_{A,\lambda}$ and $\overline{\pi}^\dagger_{A,\lambda}$ denote the analogues of $\pi_{A,\lambda}$ and $\overline{\pi}_{A,\lambda}$ for the $(\Lambda, a^\dagger, \delta)$-contact process. Generalizing the proof of Lemma 4.6 (a), we find that

$$\hat{P}^A_\lambda\big[B \cap \iota^{-1}\eta^A_\tau = \emptyset\big] = \overline{\pi}^\dagger_{A,\lambda}(B \cup \{0\}) - \overline{\pi}^\dagger_{A,\lambda}(B). \tag{4.2.43}$$

Since $|B| \leq |BA^{-1}| \leq |A|\,|B|$ for any $A, B \in \mathcal{P}_{\rm fin}(\Lambda)$ with $A \neq \emptyset$, it follows that

$$\lim_{t\to\infty} \tfrac{1}{t} \log E\big[|\eta_t^B A^{-1}|\big] = r \qquad (\emptyset \neq B \in \mathcal{P}_{\rm fin}(\Lambda)), \tag{4.2.44}$$

where r is the exponential growth rate from Proposition 4.1. The proofs of (4.1.29) and Lemma 4.7 now carry over to the functions $(\overline{\pi}_{A,\lambda})_{\lambda > r}$ without a change, and therefore the arguments in Section 4.1.4 show that Theorem 4.3 (a) holds for general initial states $\emptyset \neq A \in \mathcal{P}(\Lambda)$. ∎

4.3 Proofs of further results

Recall that $\omega = (\omega^{\mathrm{r}}, \omega^{\mathrm{i}})$ is the pair of Poisson point processes used in the graphical representation. We construct ω on the canonical probability space $\Omega := \mathcal{P}_{\mathrm{loc}}(\Lambda \times \mathbb{R}) \times \mathcal{P}_{\mathrm{loc}}(\Lambda \times \Lambda \times \mathbb{R})$, where $\mathcal{P}_{\mathrm{loc}}(\Lambda \times \mathbb{R})$ and $\mathcal{P}_{\mathrm{loc}}(\Lambda \times \Lambda \times \mathbb{R})$ denote the spaces of locally finite subsets of $\Lambda \times \mathbb{R}$ and $\Lambda \times \Lambda \times \mathbb{R}$, respectively. These spaces can in a natural way be identified with subspaces of the spaces of locally finite counting measures on $\Lambda \times \mathbb{R}$ and $\Lambda \times \Lambda \times \mathbb{R}$, respectively. Using this identification, we equip $\mathcal{P}_{\mathrm{loc}}(\Lambda \times \mathbb{R})$ and $\mathcal{P}_{\mathrm{loc}}(\Lambda \times \Lambda \times \mathbb{R})$ with the vague topology. We equip Ω with the product topology and the associated Borel-σ-field $\mathcal{F}$, and let P be the probability measure on $(\Omega, \mathcal{F})$ such that under P, the coordinate functions $\omega^{\mathrm{r}}, \omega^{\mathrm{i}}$ are Poisson point processes as described in the introduction.

We equip $\Lambda \times \mathbb{R}$ and $\Lambda \times \Lambda \times \mathbb{R}$ with a group structure by putting $(i,s)(j,t) := (ij, s+t)$ and $(i,j,s)(k,l,t) := (ik, jl, s+t)$, respectively. In line with our earlier notation, for any subset $\alpha \subset \Lambda \times \mathbb{R}$, we write $(i,s)\alpha := \{(ij, s+t) : (j,t) \in \alpha\}$. For $\beta \subset \Lambda \times \Lambda \times \mathbb{R}$, we define $(i,j,s)\beta$ analogously. We define shift operators $\theta_{i,t} : \Omega \to \Omega$ by

$$\theta_{i,t}(\alpha, \beta) := ((i,t)\alpha, (i,i,t)\beta) \tag{4.3.1}$$

$(i \in \Lambda,\ t \in \mathbb{R},\ (\alpha, \beta) \in \Omega)$. Thus, $\theta_{i,t}$ shifts a graphical representation by left-multiplication with i and increasing all times by t.

4.3.1 Conditioning and size-biasing

In this section, we prove that size-biasing and conditioning on survival are asymptotically equivalent in a 'local' sense. Let

$$\omega_t := \big(\omega^{\mathrm{r}} \cap \Lambda \times (-\infty, t]\,,\ \omega^{\mathrm{i}} \cap \Lambda \times \Lambda \times (-\infty, t]\big) \tag{4.3.2}$$

denote the restriction of the Poisson point processes used in the graphical representation to the time interval $(-\infty, t]$.

Proposition 4.14 (Conditioning and size-biasing) *Assume that the (Λ, a, δ)-contact process survives and that the exponential growth rate satisfies $r(\Lambda, a, \delta) = 0$. Then, for any $\emptyset \neq A \in \mathcal{P}_{\mathrm{fin}}(\Lambda)$,*

$$\hat{P}^A_\lambda\big[\omega_t \in \cdot\,\big] \underset{\lambda \downarrow 0}{\longrightarrow} P\big[\omega_t \in \cdot\,\big|\, A \times \{0\} \rightsquigarrow \infty\big] \qquad (t \in \mathbb{R}). \tag{4.3.3}$$

Proof It suffices to prove the claims for $t > 0$. For any $\mathcal{A} \in \mathcal{F}$, write

$$\hat{P}^A_\lambda\big[\omega_t \in \mathcal{A}\big] = \hat{P}^A_\lambda\big[\omega_t \in \mathcal{A}\,\big|\, \tau \geq t\big]\, \hat{P}^A_\lambda\big[\tau \geq t\big] + \hat{P}^A_\lambda\big[\omega_t \in \mathcal{A},\ \tau < t\big], \tag{4.3.4}$$

and observe that

$$\begin{aligned}
\hat{P}^A_\lambda\big[\omega_t \in \mathcal{A}\,\big|\, \tau \geq t\big] &= \frac{\int_0^\infty E[|\eta^A_{t+s}| 1_{\{\omega_t \in \mathcal{A}\}}] e^{-\lambda s}\,\mathrm{d}s}{\int_0^\infty E[|\eta^A_{t+s}|] e^{-\lambda s}\,\mathrm{d}s} = \frac{E[\int_0^\infty E[|\eta^A_{t+s}|\,|\,\omega_t] e^{-\lambda s}\,\mathrm{d}s 1_{\{\omega_t \in \mathcal{A}\}}]}{E[\int_0^\infty E[|\eta^A_{t+s}|\,|\,\omega_t] e^{-\lambda s}\,\mathrm{d}s]} \\
&= \frac{E[\pi_\lambda(\eta^A_t) 1_{\{\omega_t \in \mathcal{A}\}}]}{E[\pi_\lambda(\eta^A_t)]} = \frac{E[\overline{\pi}_\lambda(\eta^A_t) 1_{\{\omega_t \in \mathcal{A}\}}]}{E[\overline{\pi}_\lambda(\eta^A_t)]} \underset{\lambda \downarrow 0}{\longrightarrow} \frac{E[\overline{\rho}(\eta^A_t) 1_{\{\omega_t \in \mathcal{A}\}}]}{E[\overline{\rho}(\eta^A_t)]},
\end{aligned} \tag{4.3.5}$$

where we have used that $\overline{\pi}_\lambda \to \overline{\rho}$ pointwise as $\lambda \downarrow 0$ by Lemma 4.7 and Proposition 4.8, and bounded convergence, using the uniform bound $\pi_\lambda \leq |\cdot|$. Since

$$\begin{aligned}\frac{E[\overline{\rho}(\eta^A_t)1_{\{\omega_t\in\mathcal{A}\}}]}{E[\overline{\rho}(\eta^A_t)]} &= \frac{E[\rho(\eta^A_t)1_{\{\omega_t\in\mathcal{A}\}}]}{E[\rho(\eta^A_t)]} \\ &= \frac{E[P[A\times\{0\}\rightsquigarrow\infty\,|\,\omega_t]1_{\{\omega_t\in\mathcal{A}\}}]}{E[P[A\times\{0\}\rightsquigarrow\infty\,|\,\omega_t]]} = P\big[\omega_t\in\mathcal{A}\,\big|\,A\times\{0\}\rightsquigarrow\infty\big],\end{aligned} \tag{4.3.6}$$

formula (4.3.3) follows from Lemma 4.2, (4.3.4), and (4.3.5). ∎

4.3.2 Coupling to the maximal process

In this section we prove Theorem 4.3 (b) and Proposition 4.4. In analogy with (4.1.14), we put

$$\hat{P}^{\dagger A}_\lambda(\{i\}\times\{\mathrm{d}\omega\}\times\{\mathrm{d}t\}) := \pi^\dagger_\lambda(A)^{-1}\,1_{\{i\,\in\,\eta^{\dagger A}_t(\omega)\}}P(\mathrm{d}\omega)e^{-\lambda t}\mathrm{d}t, \tag{4.3.7}$$

which is well-defined for any $\emptyset\neq A\in\mathcal{P}_{\rm fin}(\Lambda)$ and $\lambda>r$. Recall that $\eta^{\dagger A}_\tau=\{i\in\Lambda:(i,-\tau)\rightsquigarrow A\times\{0\}\}$. We can view $\eta^{\dagger A}_t$ as the set of all 'ancestors' at time $-t$ of the set A at time 0. As before, let ι and τ denote the projections on Λ and $\mathbb{R}_+$, respectively. Then, under the law $\hat{P}^{\dagger A}_\lambda$, the random variables ι and τ describe a 'typical' ancestor of A and a 'typical' time $-\tau$.

In the next lemma, we shift the graphical representation ω in such a way that the 'typical' infected site and time (ι,τ), chosen with respect to $\hat{P}^{\{0\}}_\lambda$, are mapped to the point $(0,0)$. Note that under such a shift, the origin is mapped to ι^{-1}. Thus, the next lemma can be described by saying that if we start the contact process with only the origin infected, then seen from a typical infected site, the origin is a typical ancestor.

Lemma 4.15 (Origin seen from a typical infected site) *Assume that* $r(\Lambda,a,\delta)\leq 0$. *Then*

$$\hat{P}^{\{0\}}_\lambda\big[(\iota^{-1},\theta_{\iota^{-1},-\tau}\omega,\tau)\in\cdot\,\big] = \hat{P}^{\dagger\{0\}}_\lambda\big[(\iota,\omega,\tau)\in\cdot\,\big]. \tag{4.3.8}$$

Proof Let us write $(i,s)\overset{\omega}{\rightsquigarrow}(j,t)$ when (i,s) can be connected to (j,t) along a path in the graphical representation ω. Then

$$\begin{aligned}&\hat{P}^{\{0\}}_\lambda\big[\iota^{-1}=j,\ \theta_{\iota^{-1},-\tau}\omega\in\mathcal{A},\ \tau\in(a,b)\big] = \hat{P}^{\{0\}}_\lambda\big[\iota=j^{-1},\ \theta_{\iota^{-1},-\tau}\omega\in\mathcal{A},\ \tau\in(a,b)\big]\\ &\quad=\pi_\lambda(\{0\})^{-1}\int_a^b P\big[j^{-1}\in\eta^{\{0\}}_t,\ \theta_{j,-t}\omega\in\mathcal{A}\big]e^{-\lambda t}\,\mathrm{d}t\\ &\quad=\pi_\lambda(\{0\})^{-1}\int_a^b P\big[(0,0)\overset{\omega}{\rightsquigarrow}(j^{-1},t),\ \theta_{j,-t}\omega\in\mathcal{A}\big]e^{-\lambda t}\,\mathrm{d}t\\ &\quad=\pi_\lambda(\{0\})^{-1}\int_a^b P\big[(j,-t)\overset{\theta_{j,-t}\omega}{\rightsquigarrow}(0,0),\ \theta_{j,-t}\omega\in\mathcal{A}\big]e^{-\lambda t}\,\mathrm{d}t\\ &\quad=\pi_\lambda(\{0\})^{-1}\int_a^b P\big[(j,-t)\overset{\omega}{\rightsquigarrow}(0,0),\ \omega\in\mathcal{A}\big]e^{-\lambda t}\,\mathrm{d}t\\ &\quad=\pi^\dagger_\lambda(\{0\})^{-1}\int_a^b P\big[j\in\eta^{\dagger\{0\}}_t,\ \omega\in\mathcal{A}\big]e^{-\lambda t}\,\mathrm{d}t = \hat{P}^{\dagger\{0\}}_\lambda\big[\iota=j,\ \omega\in\mathcal{A},\ \tau\in(a,b)\big],\end{aligned} \tag{4.3.9}$$

where we have used (4.2.14). ∎

In order to prove Theorem 4.3 (b), we need two more lemmas.

Lemma 4.16 (Large populations) *Assume that the (Λ, a, δ)-contact process survives and that the exponential growth rate satisfies $r(\Lambda, a, \delta) \leq 0$. Then, for any $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$,*

$$\hat{P}^A_\lambda\big[|\eta^A_\tau| \geq K\big] \underset{\lambda\downarrow 0}{\longrightarrow} 1 \qquad (K < \infty). \tag{4.3.10}$$

Proof Let τ_λ be an exponentially distributed reandom variable with mean $1/\lambda$, independent of the Poisson processes used in the graphical representation. Then

$$\begin{aligned}\hat{P}^A_\lambda\big[|\eta^A_\tau| \geq K\big] &= \frac{E\big[|\eta^A_{\tau_\lambda}| 1_{\{|\eta^A_{\tau_\lambda}| \geq K\}}\big]}{E\big[|\eta^A_{\tau_\lambda}|\big]} \\ &= \frac{E\big[|\eta^A_{\tau_\lambda}| 1_{\{|\eta^A_{\tau_\lambda}| \geq K\}} \,\big|\, \eta^A_{\tau_\lambda} \neq \emptyset\big]}{E\big[|\eta^A_{\tau_\lambda}| \,\big|\, \eta^A_{\tau_\lambda} \neq \emptyset\big]} \geq E\big[1_{\{|\eta^A_{\tau_\lambda}| \geq K\}} \,\big|\, \eta^A_{\tau_\lambda} \neq \emptyset\big] \underset{\lambda\downarrow 0}{\longrightarrow} 1,\end{aligned} \tag{4.3.11}$$

where we have used (4.2.39), and the fact that $|\eta^A_{\tau_\lambda}|$ and $1_{\{\eta^A_{\tau_\lambda} \geq K\}}$ are positively correlated since the functions $z \mapsto z$ and $z \mapsto 1_{\{z \geq K\}}$ are nondecreasing. ∎

Recall that in the proof (in Section 4.1.4) of Proposition 4.8, sequences $A_n \in \mathcal{P}_{\rm fin}(\Lambda)$ such that $\rho(A_n) \to 1$ played an important role. Although we did not need this fact there, the next lemma implies that for $\delta > 0$, actually $\rho(A_n) \to 1$ if and only if $|A_n| \to \infty$.

Lemma 4.17 (High survival probabilities) *Assume that the (Λ, a, δ)-contact process survives, and $A_n \in \mathcal{P}_{\rm fin}(\Lambda)$. Then $|A_n| \to \infty$ implies $\rho(A_n) \to 1$.*

Proof By (4.1.40) there exist $B_m \in \mathcal{P}_{\rm fin}(\Lambda)$ with $\rho(B_m) \to 1$. Now if $A_n \in \mathcal{P}_{\rm fin}(\Lambda)$ satisfy $|A_n| \to \infty$, then by Lemma 4.13,

$$\begin{aligned}&\liminf_{n\to\infty} \rho(A_n) \\ &\quad\geq \liminf_{n\to\infty} P\big[\eta^{A_n}_s \neq \emptyset\ \forall s \geq t \,\big|\, \exists i \in \Lambda \text{ s.t. } \eta^{A_n}_t \geq iB_m\big] P\big[\exists i \in \Lambda \text{ s.t. } \eta^{A_n}_t \geq iB_m\big] \\ &\quad\geq \rho(B_m),\end{aligned} \tag{4.3.12}$$

for each $t > 0$ and m. Letting $m \to \infty$ yields the claim. ∎

We now first prove Theorem 4.3 (b) in the case $A = \{0\}$, and then indicate how the arguments may be generalised to $\emptyset \neq A \in \mathcal{P}_{\rm fin}(\Lambda)$. We will obtain Proposition 4.4 as a corollary to our proofs in the case $A = \{0\}$.

Proof of Theorem 4.3 (b) in the case $A = \{0\}$ By Lemma 4.15, we must show that for fixed $\Delta \in \mathcal{P}_{\rm fin}$, the sets $\{j \in \Delta : (\iota, -\tau) \rightsquigarrow (j, 0)\}$, $\{j \in \Delta : -\infty \rightsquigarrow (j, 0)\}$, and $\{j \in \Delta : \Lambda \times \{-\tau\} \rightsquigarrow (j, 0)\}$ are asymptotically equal under the laws $\hat{P}^{\dagger\,\{0\}}_\lambda$ as $\lambda \downarrow 0$. It suffices to show that for any $j \in \Lambda$,

$$\hat{P}^{\dagger\,\{0\}}_\lambda\big[(\iota, -\tau) \rightsquigarrow (j, 0) \,\big|\, \Lambda \times \{-\tau\} \rightsquigarrow (j, 0)\big] \underset{\lambda\downarrow 0}{\longrightarrow} 1. \tag{4.3.13}$$

and

$$\hat P^{\dagger\{0\}}_{\lambda}\big[-\infty\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\underset{\lambda\downarrow 0}{\longrightarrow}1. \tag{4.3.14}$$

Reversing the direction of time and interchanging the roles of η and $\eta^\dagger$, this then yields Proposition 4.4 as a corollary.

For any $t>0$, by Proposition 4.14,

$$\begin{aligned}\hat P^{\dagger\{0\}}_{\lambda}\big[\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]&=\hat P^{\dagger\{0\}}_{\lambda}\big[\eta^{\dagger\{j\}}_{\tau}\neq\emptyset\big]\\ &\leq\hat P^{\dagger\{0\}}_{\lambda}\big[\eta^{\dagger\{j\}}_{t}\neq\emptyset\big]+\hat P^{\dagger\{0\}}_{\lambda}\big[\tau<t\big]\underset{\lambda\downarrow 0}{\longrightarrow}P\big[\eta^{\dagger\{j\}}_{t}\neq\emptyset\,\big|-\infty\rightsquigarrow(0,0)\big].\end{aligned} \tag{4.3.15}$$

Letting $t\to\infty$ yields

$$\limsup_{\lambda\downarrow 0}\hat P^{\dagger\{0\}}_{\lambda}\big[\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\leq P\big[-\infty\rightsquigarrow(j,0)\,\big|-\infty\rightsquigarrow(0,0)\big]=:\phi(j). \tag{4.3.16}$$

By Lemma 4.15 and Theorem 4.3 (a),

$$\lim_{\lambda\downarrow 0}\hat P^{\dagger\{0\}}_{\lambda}\big[(\iota,-\tau)\rightsquigarrow(j,0)\big]=\lim_{\lambda\downarrow 0}\hat P^{\{0\}}_{\lambda}\big[j\in\iota^{-1}\eta^{\{0\}}_{\tau}\big]=P\big[j\in\overline{\eta}_0\,\big|\,0\in\overline{\eta}_0\big]=\phi(j). \tag{4.3.17}$$

Combining (4.3.16) and (4.3.17) we arrive at (4.3.13).

Since conditional on $\eta^{\dagger\{0\}}_{\tau}$, the typical site ι is chosen with equal probabilities from the sites in $\eta^{\dagger\{0\}}_{\tau}$,

$$\hat P^{\dagger\{0\}}_{\lambda}\big[(\iota,-\tau)\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]=\hat E^{\dagger\{0\}}_{\lambda}\Big[\frac{|\eta^{\dagger\{j\}}_{\tau}\cap\eta^{\dagger\{0\}}_{\tau}|}{|\eta^{\dagger\{0\}}_{\tau}|}\,\Big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\Big]. \tag{4.3.18}$$

Therefore, (4.3.13) and Lemma 4.16 imply that

$$\lim_{\lambda\downarrow 0}\hat P^{\dagger\{0\}}_{\lambda}\big[|\eta^{\dagger\{j\}}_{\tau}|\geq K\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]=1\qquad(K<\infty), \tag{4.3.19}$$

which by Lemma 4.17 implies (4.3.14). ∎

Generalization to arbitrary initial states In analogy with (4.3.7), we define, for any $\emptyset\neq A,B\in\mathcal{P}_{\rm fin}(\Lambda)$,

$$\hat P^{\dagger B}_{A,\lambda}(\{i\}\times\{{\rm d}\omega\}\times\{{\rm d}t\}):=\pi^{\dagger}_{A,\lambda}(B)^{-1}1_{\{\eta^{\dagger B}_t(\omega)\cap iA\neq\emptyset\}}P({\rm d}\omega)e^{-\lambda t}{\rm d}t, \tag{4.3.20}$$

where $\pi^{\dagger}_{A,\lambda}(B)^{-1}$ is defined below (4.2.42). Note that this is a probability measure by (4.2.41). As before, let ι denote the projection on Λ. Then, under the law $\hat P^{\dagger B}_{A,\lambda}$, the random variable ι describes a 'typical' site such that $\iota A\times\{-\tau\}\rightsquigarrow B\times\{0\}$. By an obvious analogue of Lemma 4.15, we must prove the following generalisations of (4.3.13) and (4.3.14):

$$\begin{aligned}&\text{(i)}\quad\hat P^{\dagger\{0\}}_{A,\lambda}\big[\iota A\times\{-\tau\}\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\underset{\lambda\downarrow 0}{\longrightarrow}1,\\ &\text{(ii)}\quad\hat P^{\dagger\{0\}}_{A,\lambda}\big[-\infty\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\underset{\lambda\downarrow 0}{\longrightarrow}1.\end{aligned} \tag{4.3.21}$$

Define a measure $\tilde P^{\dagger\{0\}}_{A,\lambda}$ on $\Lambda\times\Lambda\times\Omega\times\mathbb{R}_+$ by

$$\tilde P^{\dagger\{0\}}_{A,\lambda}(\{k\}\times\{i\}\times\{\mathrm{d}\omega\}\times\{\mathrm{d}t\}) := \hat E^{\dagger\{0\}}_{A,\lambda}\Big[|\eta^{\dagger\{0\}}_\tau\cap iA|^{-1}1_{\{k\,\in\,\eta^{\dagger\{0\}}_\tau\,\cap\, iA\}}1_{\{i\}\,\times\,\{\mathrm{d}\omega\}\,\times\,\{\mathrm{d}t\}}\Big]. \tag{4.3.22}$$

Let $\kappa,\iota:\Lambda\times\Lambda\times\Omega\times\mathbb{R}_+\to\Lambda$ denote the projections on the first and second coordinate, respectively. Then, under the law $\tilde P^{\dagger\{0\}}_{A,\lambda}$, the random variable κ describes a site chosen with equal probabilities from $\eta^{\dagger\{0\}}_\tau\cap\iota A$. Therefore, in order to prove (4.3.21), it suffices to prove:

$$\begin{aligned}&\text{(i)}\quad \tilde P^{\dagger\{0\}}_{A,\lambda}\big[(\kappa,-\tau)\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\underset{\lambda\downarrow 0}{\longrightarrow}1,\\ &\text{(ii)}\quad \tilde P^{\dagger\{0\}}_{A,\lambda}\big[-\infty\rightsquigarrow(j,0)\,\big|\,\Lambda\times\{-\tau\}\rightsquigarrow(j,0)\big]\underset{\lambda\downarrow 0}{\longrightarrow}1.\end{aligned} \tag{4.3.23}$$

We claim that $\tilde P^{\dagger\{0\}}_{A,\lambda}\big[(\kappa,\omega,\tau)\in\cdot\,\big]$ has a density with respect to $\hat P^{\dagger\{0\}}_{\lambda}\big[(\iota,\omega,\tau)\in\cdot\,\big]$ that is uniformly bounded away from 0 and ∞, and therefore (4.3.23) follows from (4.3.13) and (4.3.14). Indeed, by (4.3.20) and (4.3.22),

$$\begin{aligned}&\tilde P^{\dagger\{0\}}_{A,\lambda}(\{k\}\times\Lambda\times\{\mathrm{d}\omega\}\times\{\mathrm{d}t\})\\ &\quad=\pi^\dagger_{A,\lambda}(\{0\})^{-1}\sum_i E\big[|\eta^{\dagger\{0\}}_\tau\cap iA|^{-1}1_{\{k\,\in\,\eta^{\dagger\{0\}}_\tau\,\cap\, iA\}}1_{\{\eta^{\dagger\{0\}}_\tau\,\cap\, iA\,\neq\,\emptyset\}}1_{\{\mathrm{d}\omega\}}\big]e^{-\lambda t}\,\mathrm{d}t\\ &\quad=Z\pi^\dagger_\lambda(\{0\})^{-1}E\big[F1_{\{k\,\in\,\eta^{\dagger\{0\}}_\tau\}}1_{\{\mathrm{d}\omega\}}\big]e^{-\lambda t}\,\mathrm{d}t=\hat E^{\dagger\{0\}}_\lambda\big[ZF(k)1_{\{k\}\,\times\,\{\mathrm{d}\omega\}\,\times\,\{\mathrm{d}t\}}\big],\end{aligned} \tag{4.3.24}$$

where $Z:=\pi^\dagger_\lambda(\{0\})/\pi^\dagger_{A,\lambda}(\{0\})$ satisfies $|A|^{-1}\le Z\le 1$ and

$$F(k):=\sum_i|\eta^{\dagger\{0\}}_\tau\cap iA|^{-1}1_{\{k\,\in\, iA\}}=\sum_{i\in kA^{-1}}|\eta^{\dagger\{0\}}_\tau\cap iA|^{-1} \tag{4.3.25}$$

satisfies $1\le F(k)\le|A|$. ∎

4.3.3 Coupling of one-dimensional processes

Proof of Lemma 4.5 For any point (i,s) such that $(i,s)\rightsquigarrow\infty$, set

$$r_{s,t}(i):=\max\{j\in\mathbb{Z}:(i,s)\rightsquigarrow(j,t)\rightsquigarrow\infty\}\qquad(t\ge s). \tag{4.3.26}$$

Then $(r_{s,t}(i))_{t\ge s}$ is the right-most path to infinity starting at (i,s). By symmetry and the nearest-neighbor property, it suffices to show that for any (i,s) and (j,s) such that $(i,s)\rightsquigarrow\infty$ and $(j,s)\rightsquigarrow\infty$, there exists a $t\ge s$ such that $r_{s,t}(i)=r_{s,t}(j)$. Imagine that this is not the case. Then, for any $i\in\mathbb{Z}$ and $s\le t$, the maximum

$$R_{s,t}(i):=\max\{j\in\mathbb{Z}:r_{t,u}(j)=r_{s,u}(i)\text{ for some }u\ge t\} \tag{4.3.27}$$

exists. Set

$$\chi_s:=\{i\in\mathbb{Z}:R_{s,s}(i)=i\}\qquad(s\in\mathbb{R}). \tag{4.3.28}$$

It is not hard to see that $R_{s,t}$ maps $\mathbb{Z}$ into χ_t and that $R_{s,t} : \chi_s \to \chi_t$ is one-to-one. We claim that $R_{s,t} : \chi_s \to \chi_t$ is with positive probability not surjective if $s < t$. Indeed, since we are assuming that $\delta > 0$ or $a(0,1) \wedge a(1,0) > 0$, it is easy to see that with positive probability there exist $i, j, k \in \chi_t$ with $i < j < k$ such that

$$\max\{i' \in \mathbb{Z} : (0,s) \rightsquigarrow (i',t)\} = i \quad \text{and} \quad \max\{k' \in \mathbb{Z} : (1,s) \rightsquigarrow (k',t)\} = k. \tag{4.3.29}$$

It follows that $R_{s,t}(0) = i$ and $R_{s,t}(1) = k$, and therefore, since $R_{s,t}$ is monotone, there is no $n \in \mathbb{Z}$ with $R_{s,t}(n) = j$.

This 'obviously' violates stationarity. More formally, fix $s < t$ and define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ by

$$f(i,j) := P[i \in \chi_s, j \in \chi_t,\ j = R_{s,t}(i)]. \tag{4.3.30}$$

Then

$$\begin{aligned} \sum_j f(0,j) &= P[0 \in \chi_s, \exists j \in \chi_t \text{ s.t. } j = R_{s,t}(0)] \\ &= P[0 \in \chi_s] > P[\exists i \in \chi_s \text{ s.t. } 0 \in \chi_t,\ 0 = R_{s,t}(i)] = \sum_i f(i,0). \end{aligned} \tag{4.3.31}$$

Since $\sum_j f(0,j) = \sum_j f(-j,0) = \sum_i f(i,0)$ (this equality is a special case of the mass transport principle; see [Hag97], [BLPS99, Section 3], or [LP05, Chapter 7]), we arrive at a contradiction. ∎

4.3.4 Survival on finitely generated groups

In this section we prove:

Lemma 4.18 (Survival for low recovery rates) *If Λ is finitely generated, then $\delta_{\rm c} > 0$.*

Proof Let Δ be a finite generating set for Λ. Since $\{i : a(0,i) > 0\}$ generates Λ, there exists a finite subset $A \subset \{i : a(0,i) > 0\}$ that generates Δ, and thereby all of Λ. Therefore, we can find $i_0, i_1, \ldots \in \Lambda$, all different, such that $\inf_{k \geq 0} a(i_k, i_{k+1}) > 0$. We will use comparison to oriented site percolation to show that $P((i_0, 0) \rightsquigarrow \infty) > 0$ if δ is sufficiently small. Fix $T > 0$. Call a point (n,m) with $n, m \in \mathbb{N}^2$ good if in the grapical representation, in the time interval $[Tm, T(m+1))$, there is an arrow from i_n to i_{n+1} and there are no recoveries in i_n and i_{n+1}. By choosing T large enough and δ small enough, the probability that a point is good can be made arbitrarily high, uniformly in n. If this probability is larger than the critical parameter for independent 2-dimensional oriented site percolation, then with positive probability there is an upward path along good points, and therefore the contact process survives. ∎

Bibliography

[Apo69] T.M. Apostol. *Calculus, Vol. II.* Wiley, 1969.

[AS05] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. *Prob. Theory Relat. Fields.* 131(3), 376–414, 2005.

[BCGH95] J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. I. The compact case. *Canad. J. Math.* 47(1): 3–27, 1995.

[BCGH97] J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. II. The non-compact case. *J. Funct. Anal.* 146: 236–298, 1997.

[BEM03] J. Blath, A.M. Etheridge, and M.E. Meredith. Coexistence in locally regulated competing populations. University of Oxford, preprint, 2003.

[BES04] N.H. Barton, A.M. Etheridge, and A.K. Sturm. Coalescence in a random background. *Ann. Appl. Probab.* 14(2), 754–785, 2004.

[BG85] C. Bennett and G. Grinstein. Role of irreversibility in stabilizing complex and nonergodic behavior in local interacting discrete systems. *Phys. Rev. Lett.* 55, 657–660, 1985.

[BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.* 18(4), 1462–1482, 1990.

[BGN91] D.J. Barsky, G.R. Grimmett, and C.M. Newman. Percolation in half-spaces: Equality of critical densities and continuity of the percolation probability. *Probab. Theory Relat. Fields* 90(1), 111–148, 1991.

[BK89] R.M. Burton and M. Keane. Density and uniqueness in percolation. *Commun. Math. Phys.* 121, 501–505, 1989.

[BLPS99] I. Benjamini, R. Lyons, Y. Peres, O. Schramm. Group-invariant percolation on graphs. *Geom. Funct. Anal.* 9(1), 29–66, 1999.

[BS01] I. Benjamini and O. Schramm. Percolation in the hyperbolic plane. *J. Am. Math. Soc.* 14, 487–507, 2001.

[CDG04] J.T. Cox, D.A. Dawson, and A. Greven. Mutually catalytic super branching random walks: Large finite systems and renormalization analysis. *Mem. Am. Math. Soc.* 809, 2004.

[CFG96] J.T. Cox, K. Fleischmann, and A. Greven. Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Relat. Fields* 105, 513–528, 1996.

[CG86] J.T. Cox and D. Griffeath. Diffusive clustering in the two dimensional voter model. *Ann. Probab.* 14(2), 347–370, 1986.

[CG94] J.T. Cox and A. Greven. Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.* 22(2), 833–853, 1994.

[Che87] M.F. Chen. Existence theorems for interacting particle systems with non-compact state space. *Sci. China Ser. A* 30, 148–156, 1987.

[Daw77] D.A. Dawson. The critical measure diffusion process. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 40, 125–145, 1977.

[Daw93] D.A. Dawson. Measure-valued Markov processes. *Ecole d'Et de probabilits de Saint-Flour XXI. Lect. Notes Math.* 1541, 1-260, Springer, Berlin, 1993.

[DDL90] W. Ding, R. Durrett, and T.M. Liggett. Ergodicity of reversible reaction diffusion processes. *Probab. Theory Relat. Fields* 85(1), 13–26, 1990.

[DE68] D.A. Darling and P. Erdős. On the recurrence of a certain chain. *Proc. Am. Math. Soc.* 19(1): 336-338, 1968.

[DEFMPX02a] D.A. Dawson, A. Etheridge, K. Fleischmann, L. Mytnik, E.A. Perkins, and J. Xiong. Mutually catalytic branching in the plane: Finite measure states. *Ann. Probab.* 30(4), 1681–1762, 2002.

[DEFMPX02b] D.A. Dawson, A. Etheridge, K. Fleischmann, L. Mytnik, E.A. Perkins, and J. Xiong. Mutually catalytic branching in the plane: Infinite measure states. *Electron. J. Probab.* 7, paper no. 15, 61 pp., 2002.

[Deu89] J.-D. Deuschel. Invariance principle and empirical mean large deviations of the critical Ornstein-Uhlenbeck process. *Ann. Probab.* 17(1), 74–90, 1989.

[DF97a] D.A. Dawson and K. Fleischmann. A continuous super-Brownian motion in a super-Brownian medium. *J. Theoret. Probab.* 10(1), 213–276, 1997.

[DF97b] D.A. Dawson and K. Fleischmann. Longtime behavior of a branching process controlled by branching catalysts. *Stoch. Process. Appl.* 71(2), 241–257, 1997.

[DF02] D.A. Dawson and K. Fleischmann. Catalytic and mutually catalytic super-Brownian motions In R.C. Dalang (ed.) *Seminar on stochastic analysis, random fields and applications III. Proceedings of the 3rd seminar, Ascona, Switzerland, September 20-24, 1999.* Prog. Probab. 52, 89–110, Birkhäuser, Basel, 2002.

[DFMPX03] D.A. Dawson, K. Fleischmann, L. Mytnik, E.A. Perkins, and J. Xiong. Mutually catalytic branching in the plane: Uniqueness. *Ann. Inst. Henri Poincar, Probab. Stat.* 39(1), 135–191, 2003.

[DG93a] D.A. Dawson and A. Greven. Hierarchical models of interacting diffusions: Multiple time scale phenomena, phase transition and pattern of cluster-formation. *Probab. Theory Related Fields* 96(4): 435–473, 1993.

[DG93b] D.A. Dawson and A. Greven. Multiple time scale analysis of interacting diffusions. *Probab. Theory Related Fields* 95(4): 467–508, 1993.

[DG96] D.A. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.* 1, paper no. 14, 84 pp., 1996.

[DG99] D.A. Dawson and A. Greven. Hierarchically interacting Fleming-Viot processes with selection and mutation: Multiple space time scale analysis and quasi-equilibria. *Electron. J. Probab.* 4, Paper no. 4, 81 p., 1999.

[DGV95] D.A. Dawson, A. Greven and J. Vaillancourt. Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.* 347(7): 2277–2360, 1995.

[DK96] P. Donnelly and T.G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* 24(2), 698–742, 1996.

[DK99] P. Donelly and T.G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.* 9, 1091–1148, 1999.

[DLSS91] B. Derrida, J.L. Lebowitz, E.R. Speer, and H. Spohn. Dynamics of an anchored Toom interface. *J. Phys. A: Math. Gen.* 24, 4805–4834, 1991.

[DP98] D.A. Dawson and E.A. Perkins. Long-time behavior and coexistence in a mutually catalytic branching model. *Ann. Probab.* 26(3), 1088-1138, 1998.

[DS95] R. Durrett and R. Schinazi. Intermediate phase for the contact process on a tree. *Ann. Probab.* 23(2), 668–673, 1995.

[Dyn02] E. Dynkin. *Diffusions, superdiffusions and partial differential equations.* Vol. 50 of American Mathematical Society Colloquium Publications. AMS, Providence, RI, 2002.

[DZ98] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications. 2. ed.* Springer, New York, 1998.

[EF98] A. Etheridge and K. Fleischmann. Persistence of a two-dimensional super-Brownian motion in a catalytic medium. *Probab. Theory Relat. Fields* 110(1), 1–12, 1998.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes; Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[EK04] J. Engländer and A.E. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.* 32(1A), 78–99, 2004.

[EP99] J. Engländer and R.G. Pinsky. On the construction and support properties of measure-valued diffusions on $D \subseteq \mathbb{R}^d$ with spatially dependent branching. *Ann. Probab.*, 27(1): 684–730, 1999.

[ER91] N. El Karoui and S. Roelly. Propriétés de martingales, explosion et représentation de Lévy- Khintchine d'une classe de processus de branchement à valeurs mesures. *Stoch. Proc. Appl.* 38(2): 239–266, 1991.

[ET02] J. Engländer and D. Turaev. A scaling limit for a class of superdiffussions. *Ann. Probab.* 30(2), 683–722, 2002.

[Eth00] A. Etheridge. *An Introduction to Superprocesses.* University Lecture Series 20, AMS, Providence, 2000.

[EU48] C.J. Everett and S. Ulam. Multiplicative systems in several variables, I. Los Alamos Scientific Laboratory, LA-683, 1948.

[Ewe04] W.J. Ewens. *Mathematical Population Genetics. I: Theoretical Introduction. 2nd ed.* Interdisciplinary Mathematics 27. Springer, New York, 2004.

[FG94] K. Fleischmann and A. Greven. Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Relat. Fields* 98(4), 517–566, 1994.

[Fit88] P.J. Fitzsimmons. Construction and regularity of measure-valued branching processes. *Isr. J. Math.* 64(3): 337–361, 1988.

[Fit91] P.J. Fitzsimmons. Correction to "Construction and regularity of measure-valued branching processes". *Isr. J. Math.*, 73(1): 127, 1991.

[Fit92] P.J. Fitzsimmons. On the martingale problem for measure-valued Markov branching processes. pp 39–51 in: *Seminar on Stochastic Processes, 1991.* Progress in Probability 29, Birkhäuser, Boston, 1992.

[FK99] K. Fleischmann and A. Klenke. Smooth density field of catalytic super-Brownian motion. *Ann. Appl. Probab.* 9, 298–318, 1999.

[FS03] K. Fleischmann and J.M. Swart. Extinction versus exponential growth in a supercritical super-Wright-Fischer diffusion. *Stoch. Proc. Appl.* 106(1): 141–165, 2003.

[FS04] K. Fleischmann and J.M. Swart. Trimmed trees and embedded particle systems. *Ann. Probab.* 32(3a): 2179–2221, 2004.

[GH02] G. Grimmett and P. Hiemer. Directed percolation and random walk. Pages 273–297 in: V. Sidoravicius (ed.), *In and Out of Equilibrium.* Prog. Probab. 51. Birkhäuser, Boston, 2002

[GKW99] A. Greven, A. Klenke, and A. Wakolbinger. The longtime behavior of branching random walk in a catalytic medium. *Electron. J. Probab.* 4, paper no. 12, 80 pp., 1999.

[GKW01] A. Greven, A. Klenke, and A. Wakolbinger. Interacting Fisher-Wright diffusions in a catalytic medium. *Probab. Theory Related Fields* 120(1): 85–117, 2001.

[GLW05] A. Greven, V. Limic, and A. Winter. Representation theorems for interacting Moran models and interacting Fisher-Wright models, and applications. *Electron. J. Probab.* 10, paper no. 39, 1286–1358, 2005.

[GM90] G.R. Grimmett and J.M. Marstrand. The supercritical phase of percolation is well behaved. *Proc. R. Soc. Lond., Ser. A* 430, no. 1879, 439–457, 1990.

[GW91] L.G. Gorostiza and A. Wakolbinger. Persistence criteria for a class of critical branching particle systems in continuous time. *Ann. Probab.* 19(1), 266–288, 1991.

[Hag97] O. Häggström. Infinite clusters in dependent automorphism invariant percolation on trees. *Ann. Probab.* 25(3), 1423–1436, 1997.

[Har63] T.E. Harris. *The theory of branching processes.* Springer, Berlin, 1963.

[Har76] T.E. Harris. On a class of set-valued Markov processes. *Ann. Probab.* 4, 175–194, 1976.

[Har78] T.E. Harris. Additive set-valued Markov processes and graphical methods. *Ann. Probab.* 6, 355–378, 1978.

[HS98] F. den Hollander and J.M. Swart. Renormalization of hierarchically interacting isotropic diffusions. *J. Stat. Phys.* 93: 243–291, 1998.

[Jir64] M. Jiřina. Branching processes with measure-valued states. In *Trans. Third Prague Conf. Information Theory, Statist. Decision Functions, Random Processes (Liblice, 1962)*, pages 333–357, Czech. Acad. Sci., Prague, 1964.

[Kal76] O. Kallenberg. *Random Measures.* Akademie-Verlag, Berlin, 1976.

[Kal77] O. Kallenberg. Stability of critical cluster fields. *Math. Nachr.* 77, 7–43, 1977.

[Kal83] O. Kallenberg. *Random measures*, 3rd rev. and enl. ed. Akademie-Verlag, Berlin, 1983.

[Kes86] H. Kesten. Aspects of first passage percolation. Pages 125–264 in: P.L. Hennequin (ed.), *École d'été de probabilités de Saint-Flour XIV - 1984.* Lect. Notes Math. 1180, Springer, Berlin, 1986.

[Kle96] A. Klenke. Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.* 24(2): 660–697, 1996.

[KN97] S.M. Krone and C. Neuhauser. Ancestral processes with selection. *Theor. Popul. Biol.* 51(3), 210–237, 1997.

[Kol33] A. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Ergebnisse der Mathematik, 1933.

[Law05] G.F. Lawler. *Conformally Invariant Processes in the Plane.* AMS, 2005.

[Lie81] A. Liemant. Kritische Verzweigungsprozesse mit allgemeinem Phasenraum. IV. *Math. Nachr.* 102: 235–254, 1981.

[Lig85] T.M. Liggett. *Interacting Particle Systems.* Springer, New York, 1985.

[Lig96] T.M. Liggett. Branching random walks and contact processes on homogeneous trees. *Probab. Theory Relat. Fields* 106(4), 495–519, 1996.

[Lig99] T.M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Process.* Springer, Berlin, 1999.

[Loe63] M. Loève. *Probability Theory* 3rd ed. Van Nostrand, Princeton, 1963.

[Loe78] M. Loève. *Probability Theory II* 4th ed. Graduate Texts in Mathematics 46. Springer, New York, 1978.

[LP05] R. Lyons and Y. Peres. *Probability on trees and networks.* Draft available from http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html, 2005.

[LPP96] R. Lyons, R. Pemantle, and Y. Peres. Random walks on the lamplighter group. *Ann. Probab.* 24(4), 1993–2006, 1996.

[LS81] T.M. Liggett and F. Spitzer. Ergodic theorems for coupled random walks and other systems with locally interacting components. *Z. Wahrsch. verw. Gebiete* 56, 443–468, 1981.

[Mou92] T.S. Mountford. The ergodicity of a class of reversible reaction-diffusion processes. *Probab. Theory Relat. Fields* 92(2), 259–274, 1992.

[MT95] C. Müller and R. Tribe. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Relat. Fields* 102(4), 519–545, 1995.

[MW89] B. Mohar and W. Woess. A survey on spectra of infinite graphs. *Bull. Lond. Math. Soc.* 21(3), 209–234, 1989.

[Neu90] C. Neuhauser. An ergodic theorem for Schlögl models with small migration. *Probab. Theory Relat. Fields* 85(1), 27–32, 1990.

[NS80] M. Notohara and T. Shiga. Convergence to genetically uniform state in stepping stone models of population genetics. *J. Math. Biology* 10, 281–294, 1980.

[Pat88] A.L.T. Paterson. *Amenability.* AMS, Providence, 1988.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York, 1983.

[Pem92] R. Pemantle. The contact process on trees. *Ann. Probab.* 20(4), 2089–2116, 1992.

[Pen04] C. Penßel. *Interacting Catalytic Feller Diffusions: Finite System Scheme and Renormalisation.* Logos, Berlin, 2004.

[Rev84] D. Revuz. *Markov chains. 2nd rev. ed.* North-Holland Mathematical Library, Vol 11. North-Holland, Amsterdam, 1984.

[RW87] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales, Volume 2: Ito Calculus.* Wiley, Chichester, 1987.

[Sch72] F. Schlögl. Chemical reaction models and non-equilibrium phase transitions. *Z. Phys.* 253, 147–161, 1972.

[Sch86] R.H. Schonmann. The asymmetric contact process. *J. Stat. Phys.* 44, 505–534, 1986.

[Sch98] F. Schiller. *Application of the Multiple Space-Time Scale Analysis on a System of R-valued, Hierarchically Interacting, Stochastic Differential Equations.* Master thesis, Universtity Erlangen-Nürnberg, 1998.

[SF83] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.* 20: 1–10, 1983.

[Sha88] M. Sharpe. *General Theory of Markov Processes.* Academic Press, Boston, 1988.

[Shi80a] T. Shiga. An interacting system in population genetics. *J. Math. Kyoto Univ.* 20(2), 213–242, 1980.

[Shi80b] T. Shiga. An interacting system in population genetics, II. *J. Math. Kyoto Univ.* 20(4), 723–733, 1980.

[Shi81] T. Shiga. Diffusion processes in population genetics. *J. Math. Kyoto Univ.* 21, 133–151, 1981.

[Shi92] T. Shiga. Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems. *Osaka J. Math.* 29, 789–807, 1992.

[Smo83] J. Smoller. *Shock Waves and Reaction-diffusion Equations.* Vol. 258 of Grundlehren Math. Wiss., Springer, New York, 1983.

[SS80] T. Shiga and A. Shimizu. Infinite dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* 20, 395–416, 1980.

[SS97] M. Salzano and R.H. Schonmann. The second lowest extremal invariant measure of the contact process. *Ann. Probab.* 25(4), 1846–1871, 1997.

[SS99] M. Salzano and R.H. Schonmann. The second lowest extremal invariant measure of the contact process. II. *Ann. Probab.* 27(2), 845–875, 1999.

[Sta96] A.M. Stacey. The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* 24(4), 1711–1726, 1996.

[SU86] T. Shiga and K. Uchiyama. Stationary states and their stability of the stepping stone model involving mutation and selection. *Probab. Theory Relat. Fields* 73, 87–117, 1986.

[Swa99] J.M. Swart. *Large Space-Time Scale Behavior of Linearly Interacting Diffusions.* PhD thesis, Katholieke Universiteit Nijmegen, 1999. http://helikon.ubn.kun.nl/mono/s/swart_/largspscb.pdf.

[Swa00] J.M. Swart. Clustering of linearly interacting diffusions and universality of their long-time limit distribution. *Prob. Theory Related Fields* 118: 574–594, 2000.

[WG74] H.W. Watson and F. Galton. On the probability of the extinction of families. *J. Anthropol. Inst. Great Britain and Ireland* 4, 138–144, 1874.

[WK74] K.G. Wilson and J. Kogut. The renormalization group and the ε-expansion. *Phys. Rep.* 12C, 75–200, 1974.

[YW71] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* 11: 155–167, 1971.