Branching-coalescing particle systems

Siva R. Athreya Jan M. Swart*

February 17, 2004

Abstract

We study the ergodic behavior of systems of particles performing independent random walks, binary splitting, coalescence and deaths. Such particle systems are dual to systems of linearly interacting Wright-Fisher diffusions, used to model a population with resampling, selection and mutations. We use this duality to prove that the upper invariant measure of the particle system is the only homogeneous nontrivial invariant law and the limit started from any homogeneous nontrivial initial law.

Contents

*Research supported in part by the German Science Foundation.
MSC 2000. Primary: 60K35, 92D25; Secondary: 60J80, 60J60.
Keywords. First Schlögl model, reaction-diffusion process, autocatalytic reaction, branching, coalescence, resampling, selection, mutation, contact process.

1 Introduction and Main Results

1.1 Introduction

This paper studies 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 paper, 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

$$\begin{aligned} Gf(x) := &\sum_{ij} a(i,j)x(i)\{f(x+\delta_j-\delta_i)-f(x)\} + b\sum_i x(i)\{f(x+\delta_i)-f(x)\} \\ &+c\sum_i x(i)(x(i)-1)\{f(x-\delta_i)-f(x)\} + d\sum_i x(i)\{f(x-\delta_i)-f(x)\}, \end{aligned} \tag{1.1}$$

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 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.,

$$\Omega_d := \{i = (i_0, i_1, \ldots) : i_\alpha \in \{0, \ldots, d-1\} \ \forall \alpha \geq 0, \ i_\alpha \neq 0 \text{ finitely often}\}, \tag{1.2}$$

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.

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]):

$$
\begin{aligned}
\mathrm{d}\mathcal{X}_t(i) = & \sum_j a(j,i)(\mathcal{X}_t(j) - \mathcal{X}_t(i))\,\mathrm{d}t + b\mathcal{X}_t(i)(1-\mathcal{X}_t(i))\,\mathrm{d}t - d\mathcal{X}_t(i)\,\mathrm{d}t \\
& + \sqrt{2c\mathcal{X}_t(i)(1-\mathcal{X}_t(i))}\,\mathrm{d}B_t(i) \qquad\qquad (t \geq 0,\ i \in \Lambda).
\end{aligned}
\tag{1.3}
$$

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

$$\phi^x := \prod_i \phi(i)^{x(i)}, \tag{1.4}$$

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 1 (a) below)

$$E^\phi[(1-\mathcal{X}_t)^x] = E^x[(1-\phi)^{X_t^\dagger}]. \tag{1.5}$$

Formula (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 (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 (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 (1.5) in terms of potential ancestors. For some recent versions of this duality, see also [DK99, DG99, BES02]. A SDE that is dual to branching-*annihilating* random walks occurs in [BEM03, Lemma 2.1]. A SPDE version of (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 (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 [FS03b].) Thus, the duality relation (1.5) says that, loosely speaking, branching-coalescing particle systems have a random generating semigroup. The SDE (1.3) will be our main tool for studying branching-coalescing particle systems.

1.3 Preliminaries

In this section we introduce the notation and definitions that we will use throughout the paper.

(Inner product and norm notation) For $\phi, \psi \in [-\infty, \infty]^\Lambda$, we write

$$\langle \phi, \psi \rangle := \sum_i \phi(i)\psi(i) \qquad \text{and} \qquad |\phi| := \sum_i |\phi(i)|, \tag{1.6}$$

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

$$E[(1-\psi)^{\mathrm{Pois}(\phi)}] = E[e^{-\langle\phi,\psi\rangle}] \qquad (\psi \in [0,1]^\Lambda). \tag{1.7}$$

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

$$E[(1-\psi)^{\mathrm{Thin}_\phi(x)}] = E[(1-\phi\psi)^x] \qquad (\psi \in [0,1]^\Lambda). \tag{1.8}$$

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

$$\mathrm{Thin}_\psi(\mathrm{Thin}_\phi(x)) \stackrel{\mathcal{D}}{=} \mathrm{Thin}_{\psi\phi}(x) \quad \text{and} \quad \mathrm{Thin}_\psi(\mathrm{Pois}(\phi)) \stackrel{\mathcal{D}}{=} \mathrm{Pois}(\psi\phi), \tag{1.9}$$

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

$$\mathcal{S}(\mathcal{N}(\Lambda)) := \{f : \mathcal{N}(\Lambda) \to \mathbb{R} : |f(x)| \leq K|x|^k + M \text{ for some } K, M, k \geq 0\} \tag{1.10}$$

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

$$S_t f(x) := E^x[f(X_t)] \qquad (t \geq 0,\ x \in \mathcal{N}(\Lambda),\ f \in \mathcal{S}(\mathcal{N}(\Lambda))) \tag{1.11}$$

maps $\mathcal{S}(\mathcal{N}(\Lambda))$ into itself (see Proposition 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

$$\sum_i \gamma_i < \infty \quad \text{and} \quad \sum_j a_{\mathrm{s}}(i,j)\gamma_j \leq K\gamma_i \quad (i \in \Lambda) \tag{1.12}$$

for some $K < \infty$. We fix such $(\gamma_i)_{i\in\Lambda}$ throughout the paper and define the Liggett-Spitzer space (after [LS81])

$$\mathcal{E}_\gamma(\Lambda) := \{x \in \mathbb{N}^\Lambda : \|x\|_\gamma < \infty\}, \tag{1.13}$$

where for $x \in \mathbb{Z}^\Lambda$ we put

$$\|x\|_\gamma := \sum_i \gamma_i |x(i)|. \tag{1.14}$$

We let $\mathcal{C}_{\mathrm{Lip}}(\mathcal{E}_\gamma(\Lambda))$ 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 11 below) that for each $f \in \mathcal{C}_{\mathrm{Lip}}(\mathcal{E}_\gamma(\Lambda))$ and $t \geq 0$, the function $S_t f$ defined in (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

$$E^x[f(X_t)] = S_t f(x) \qquad (f \in \mathcal{C}_{\mathrm{Lip}}(\mathcal{E}_\gamma(\Lambda)),\ x \in \mathcal{E}_\gamma(\Lambda),\ t \geq 0). \tag{1.15}$$

We will show (in Proposition 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

$$P^x[X_t \neq 0\ \forall t \geq 0] > 0 \quad \text{for some} \quad x \in \mathcal{N}(\Lambda). \tag{1.16}$$

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

$$T_g x(j) := x(g^{-1}j) \qquad (i \in \Lambda,\ x \in \mathbb{N}^\Lambda,\ g \in \mathrm{Aut}(\Lambda, a)). \tag{1.17}$$

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.

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 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)

$$P^x[\mathrm{Thin}_\phi(X_t) = 0] = P^\phi[\mathrm{Thin}_{\mathcal{X}_t^\dagger}(x) = 0] \qquad (t \geq 0,\ \phi \in [0, 1]^\Lambda,\ x \in \mathcal{E}_\gamma(\Lambda)). \tag{1.18}$$

(b) (Self-duality) *Assume $c > 0$, then*

$$P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t\psi) = 0] = P^\psi[\mathrm{Pois}(\tfrac{b}{c}\phi\mathcal{X}_t^\dagger) = 0] \qquad (t \geq 0,\ \phi, \psi \in [0, 1]^\Lambda). \tag{1.19}$$

(c) (Poissonization) *Assume $c > 0$, then*

$$P^{\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\phi))}[X_t \in \cdot\,] = P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t) \in \cdot\,] \qquad (t \geq 0,\ \phi \in [0, 1]^\Lambda), \tag{1.20}$$

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 1 (a) is just a reformulation of the duality relation (1.5). Theorem 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 (1.19) can be rewritten as

$$E^\phi\big[e^{-\frac{b}{c}\langle \mathcal{X}_t, \psi\rangle}\big] = E^\psi\big[e^{-\frac{b}{c}\langle \phi, \mathcal{X}_t^\dagger\rangle}\big] \qquad (t \geq 0,\ \phi, \psi \in [0, 1]^\Lambda). \tag{1.21}$$

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 (1.18) and (1.19), which may feel a little uneasy in the beginning, is convenient, we give here the proof of the Poissonization formula (1.20).

Proof of Theorem 1 (c) By (1.9) and the duality relations (1.18) and (1.19),

$$\begin{aligned} &P^{\mathcal{L}(\mathrm{Pois}(\frac{b}{c}\phi))}[\mathrm{Thin}_\psi(X_t)=0] = P^\psi[\mathrm{Thin}_{\mathcal{X}_t^\dagger}(\mathrm{Pois}(\tfrac{b}{c}\phi))=0] \\ &\quad = P^\psi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t^\dagger\phi)=0] = P^\phi[\mathrm{Pois}(\tfrac{b}{c}\psi\mathcal{X}_t)=0] = P^\phi[\mathrm{Thin}_\psi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t))=0]. \end{aligned} \tag{1.22}$$

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 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*

$$E[X_t^{(\infty)}(i)] \leq \begin{cases} \frac{r}{c(1-e^{-rt})} & \text{if } r\neq 0, \\ \frac{1}{ct} & \text{if } r=0 \end{cases} \qquad (i\in\Lambda,\ t>0). \tag{1.23}$$

(c) *If $X^{(n)}$ are (a,b,c,d)-braco-processes starting in initial states $x^{(n)}\in\mathcal{E}_\gamma(\Lambda)$ such that*

$$x^{(n)}(i)\uparrow\infty \quad \text{as } n\uparrow\infty \qquad (i\in\Lambda), \tag{1.24}$$

then

$$\mathcal{L}(X_t^{(n)})\uparrow\mathcal{L}(X_t^{(\infty)}) \quad \text{as } n\uparrow\infty \qquad (t>0). \tag{1.25}$$

(d) *There exists an invariant measure $\overline{\nu}$ of the (a,b,c,d)-braco-process such that*

$$\mathcal{L}(X_t^{(\infty)})\downarrow\overline{\nu} \quad \text{as } t\uparrow\infty. \tag{1.26}$$

(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*

$$\int\overline{\nu}(\mathrm{d}x)(1-\phi)^x = P^\phi[\exists t\geq 0 \text{ such that } \mathcal{X}_t^\dagger=0] \qquad (\phi\in[0,1]^\Lambda), \tag{1.27}$$

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 2 (f) holds, note that by Theorem 1 (a) and Theorem 2 (c),

$$P[\mathrm{Thin}_\phi(X_t^{(\infty)})=0] = \lim_{n\uparrow\infty}P^\phi[\mathrm{Thin}_{\mathcal{X}^\dagger}(x^{(n)})=0] = P^\phi[\mathcal{X}_t^\dagger=0] \qquad (\phi\in[0,1]^\Lambda,\ t>0). \tag{1.28}$$

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 (1.28) we arrive at (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 (The maximal resampling-selection process) *Let $\mathcal{X}^1$ denote the (a,b,c,d)-resem-process started in $\mathcal{X}^1_0(i) = 1$ $(i \in \Lambda)$. Then the following holds.*

(a) *There exists an invariant measure $\overline{\mu}$ of the (a,b,c,d)-resem process such that*

$$\mathcal{L}(\mathcal{X}^1_t) \downarrow \overline{\mu} \quad \text{as } t \uparrow \infty. \tag{1.29}$$

(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*

$$\int \overline{\mu}(\mathrm{d}\phi)(1-\phi)^x = P^x[\exists t \geq 0 \text{ such that } X^\dagger_t = 0] \qquad (x \in \mathcal{N}(\Lambda)), \tag{1.30}$$

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 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 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 4 (b) can be derived from Theorem 4 (a) by Poissonization, but not vice versa.

1.5 Methods

A key ingredient in the proofs of Theorem 3 (d) and Theorem 4 is the following property of resampling-selection processes, which is of some interest on its own.

Lemma 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*

$$\mathcal{X}_t = 0 \quad \text{for some } t \geq 0 \quad \text{or} \quad \lim_{t\to\infty} |\mathcal{X}_t| = \infty \quad \text{a.s.} \tag{1.31}$$

Note that by Theorem 1 (b),

$$E^\phi\big[e^{-\frac{b}{c}\langle \mathcal{X}_t, 1\rangle}\big] = E^1\big[e^{-\frac{b}{c}\langle \phi, \mathcal{X}^\dagger_t\rangle}\big] \geq e^{-\frac{b}{c}\langle \phi, 1\rangle} \qquad (\phi \in [0,1]^\Lambda), \tag{1.32}$$

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 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 5 is established the proof of Theorem 3 (d) is simple.

Proof of Theorem 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 (1.9), Theorem 1 (b), and Theorem 2 (f)

$$\begin{aligned} P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{Y}))=0] &= \lim_{t\to\infty} P^1[\mathrm{Pois}(\tfrac{b}{c}\phi\mathcal{X}_t)=0] = \lim_{t\to\infty} P^\phi[\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}^\dagger_t)=0] \\ &\overset{!}{=} P^\phi[\exists t\geq 0 \text{ such that } \mathcal{X}^\dagger_t=0] = P[\mathrm{Thin}_\phi(Y)=0], \end{aligned} \tag{1.33}$$

where we have used Lemma 5 in the equality marked with '!'. Since (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 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 (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 4, we need one more lemma.

Lemma 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 $t>0$*

$$\lim_{n\to\infty} P[\mathrm{Thin}_{\phi_n}(X_t)=0]=0, \tag{1.34}$$

for all $\phi_n\in[0,1]^\Lambda$ satisfying $|\phi_n|\to\infty$.

Proof of Theorem 4 (a) Let $\mathcal{X}^\dagger$ denote the $(a^\dagger,b,c,d)$-resem-process started in ϕ. By Theorem 1 (a), Lemmas 5 and 6, and Theorem 2 (f),

$$\begin{aligned} \lim_{t\to\infty} P[\mathrm{Thin}_\phi(X_t)=0] &= \lim_{t\to\infty} P[\mathrm{Thin}_{\mathcal{X}^\dagger_{t-1}}(X_1)=0] \\ &= P[\exists t\geq 0 \text{ such that } \mathcal{X}^\dagger_t=0] = \int \overline{\nu}(\mathrm{d}x)\,(1-\phi)^x. \end{aligned} \tag{1.35}$$

Since this holds for all $\phi\in[0,1]^\Lambda$, it follows that $\mathcal{L}(X_t)\Rightarrow\overline{\nu}$. ∎

Proof of Theorem 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 4 (a), $\mathcal{L}(X_t)\Rightarrow\mathcal{L}(X_\infty)$ as $t\to\infty$. Therefore, by Poissonization (Theorem 1 (c)) and by Theorem 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

$$\begin{aligned} &P\big[e^{-\frac{b}{c}\langle\mathcal{X}_t,\phi\rangle}\big] = P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_t))=0] \\ &\quad\Longrightarrow P[\mathrm{Thin}_\phi(\mathrm{Pois}(\tfrac{b}{c}\mathcal{X}_\infty))=0] = P\big[e^{-\frac{b}{c}\langle\mathcal{X}_\infty,\phi\rangle}\big] \qquad \text{as } t\to\infty. \end{aligned} \tag{1.36}$$

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 4 (a) is stronger than Theorem 4 (b).

Summarizing, all the hard work for getting Theorem 4 is in proving Lemmas 5 and 6, as well as the more basic Theorems 1 and 2. The heart of the proof of Theorem 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 23). We call this relation a subduality. Theorem 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 24, in particular formula (6.1)), which plays a key role in the proof of Lemma 5.

Our methods are similar to those of Shiga and Uchiyama [SU86]. Since they prove a version of our Theorem 4 (b), while our main focus is on proving the stronger Theorem 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 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 6. By adapting elements of their proof to our situation, we were able to simplify and considerably shorten our original proof of Lemma 6.

Our original proof of Lemma 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. This ergodic theorem seems to be new, and will be presented seperately in [Swa04].

1.6 Discussion

Generalizing our model, let X be a process in a Liggett-Spitzer subspace of $\mathbb{N}^\Lambda$, with local jump rates

$$\begin{array}{ll} x \mapsto x + \delta_j - \delta_i & \text{with rate } a(i,j) \\ x \mapsto x + \delta_i & \text{with rate } \sum_{n=0}^{k} b_n x^{(n)}, \\ x \mapsto x - \delta_i & \text{with rate } \sum_{n=1}^{k+1} c_n x^{(n)}, \end{array} \tag{1.37}$$

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 (1.37) are known as reaction-diffusion systems. It has been known for a long time that if the coefficients satisfy

$$a = a^\dagger \quad \text{and} \quad b_n = \lambda c_n \quad \text{for some} \quad \lambda \geq 0, \tag{1.38}$$

then $\mathcal{L}(\text{Pois}(\lambda))$ is a reversible equilibrium for the corresponding reaction-diffusion system. Note that the (a, b, c, d)-braco-process satisfies (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 (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 $\mathcal{L}(\text{Pois}(\frac{b}{c}))$. 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 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

$$P^x[X_t \in \cdot\,] \Rightarrow \rho(x)\overline{\nu} + (1 - \rho(x))\delta_0 \quad \text{as} \quad t \to \infty \qquad (x \in \mathcal{N}(\Lambda)), \tag{1.39}$$

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 (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 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 paper can be generalized to reaction-diffusion processes with higher-order branching and coalescence as in (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 (c) and (d) and Theorem 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.

1.7 Outline

We start in Section 2 with a few generalities about martingale problems that will be needed in our proofs. In Section 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 4 contains the proof of Theorem 1 and of the subduality between branching-coalescing particle systems and super random walks. In Section 5 we prove Theorems 2 and 3. In Section 6, finally, we prove Lemma 5 and Lemma 6, thereby completing the proof of Theorem 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.

2 Martingale problems

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 $f \in \mathcal{D}(A)$,

$$E\big[|f(X_t)|\big] < \infty \quad \text{and} \quad \int_0^t E\big[|Af(X_s)|\big]\mathrm{d}s < \infty \qquad (t \geq 0), \tag{2.1}$$

and the process $(M_t)_{t\geq 0}$ defined by

$$M_t := f(X_t) - \int_0^t Af(X_s)\mathrm{d}s \qquad (t \geq 0) \tag{2.2}$$

is a martingale with respect to the filtration generated by X.

2.2 Duality with error term

For later use in Section 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 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 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*

$$\Phi_1(x_1, x_2) := A_1\Psi(\cdot, x_2)(x_1) \quad \textit{and} \quad \Phi_2(x_1, x_2) := A_2\Psi(x_1, \cdot)(x_2) \qquad (x_1 \in E_1,\ x_2 \in E_2) \tag{2.3}$$

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

$$\int_0^T \mathrm{d}s \int_0^T \mathrm{d}t\ E\big[|\Phi_i(X^1_s, X^2_t)|\big] < \infty \qquad (T \geq 0,\ i = 1, 2). \tag{2.4}$$

Then

$$E[\Psi(X^1_T, X^2_0)] - E[\Psi(X^1_0, X^2_T)] = \int_0^T \mathrm{d}t\, E[R(X^1_t, X^2_{T-t})] \qquad (T \geq 0), \tag{2.5}$$

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

$$F(s,t) := E[\Psi(X^1_s, X^2_t)] \qquad (s,t \geq 0). \tag{2.6}$$

Then, for each $T > 0$,

$$\begin{aligned}\int_0^T \mathrm{d}t\, \big\{F(t,0) - F(0,t)\big\} &= \int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t) - F(T-t,t) + F(t,0)\big\} \\ &= \int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t)\big\} - \int_0^T \mathrm{d}t\, \big\{F(t,T-t) - F(t,0)\big\},\end{aligned} \tag{2.7}$$

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,

$$E\big[\Psi(X^1_{T-t}, x_2)\big] - E\big[\Psi(X^1_0, x_2)\big] = \int_0^{T-t} \mathrm{d}s\, E\big[\Phi_1(X^1_s, x_2)\big] \qquad (x_2 \in E_2), \tag{2.8}$$

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 (2.4), we find that

$$\begin{aligned}&\int_0^T \mathrm{d}t\, \big\{F(T-t,t) - F(0,t)\big\} = \int_0^T \mathrm{d}t\, \big\{E\big[\Psi(X^1_{T-t}, X^2_t)\big] - E\big[\Psi(X^1_0, X^2_t)\big]\big\} \\ &\quad = \int_0^T \mathrm{d}t \int_0^{T-t} \mathrm{d}s\, E\big[\Phi_1(X^1_s, X^2_t)\big] = \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_1(X^1_{t-s}, X^2_s)\big].\end{aligned} \tag{2.9}$$

Treating the second term in the right-hand side of (2.7) in the same way, we find that

$$\int_0^T \mathrm{d}t\, \big\{F(t,0) - F(0,t)\big\} = \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_1(X^1_{t-s}, X^2_s)\big] - \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, E\big[\Phi_2(X^1_{t-s}, X^2_s)\big]. \tag{2.10}$$

Differentiating with respect to T we arrive at (2.5). ∎

3 Construction and Comparison

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 (1.10) and of G from (1.1).

Proposition 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*

$$E^x\big[|X|_t^{\langle k\rangle}\big] \leq |x|^{\langle k\rangle} e^{kbt} \qquad (k = 1,2,\ldots,\ t \geq 0). \tag{3.1}$$

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 $f^k_t(x) := |x|^{\langle k\rangle} e^{-kbt}$. It is easy to see that

$$\{G + \tfrac{\partial}{\partial t}\} f^k_t(x) \leq kb|x|^{\langle k\rangle} e^{-kbt} - kb|x|^{\langle k\rangle} e^{-kbt} = 0. \tag{3.2}$$

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

$$M_t := f^k_{t\wedge\tau_N}(X_{t\wedge\tau_N}) - \int_0^{t\wedge\tau_N} \big(\{G + \tfrac{\partial}{\partial s}\} f^k_s\big)(X_s)\,\mathrm{d}s \qquad (t \geq 0) \tag{3.3}$$

is a martingale. By (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

$$E^x\big[|X_{t\wedge\tau_N}|^{\langle k\rangle}\big] \leq |x|^{\langle k\rangle} e^{kbt} \qquad (k = 1, 2, \ldots,\ t \geq 0). \tag{3.4}$$

In particular, setting $k = 1$, we see that

$$NP^x[\tau_N \leq t] \leq E^x\big[|X_{t\wedge\tau_N}|\big] \leq |x|e^{bt} \qquad (t \geq 0), \tag{3.5}$$

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 (3.4), using Fatou, we arrive at (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 $(M^N_t)_{t\geq 0}$ given by

$$M^N_t := f(X_{t\wedge\tau_N}) - \int_0^{t\wedge\tau_N} Gf(X_s)\mathrm{d}s \qquad (t \geq 0) \tag{3.6}$$

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 (3.1). Using (3.4), one can now check that for fixed $t \geq 0$, the random variables $\{M^N_t\}_{N\geq 1}$ are uniformly integrable. Taking the pointwise limit in (3.6), one can now check that X solves the martingale problem for G with domain $\mathcal{S}(\mathcal{N}(\Lambda))$.∎

3.2 Monotonicity and subadditivity

In this section we present two simple comparison results for finite branching-coalescing particle systems.

Lemma 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*

$$x \leq \tilde{x}, \quad b \leq \tilde{b}, \quad c \geq \tilde{c}, \quad d \geq \tilde{d}. \tag{3.7}$$

Then X and $\tilde{X}$ can be coupled in such a way that

$$X_t \leq \tilde{X}_t \qquad (t \geq 0). \tag{3.8}$$

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 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*

$$Z_t \leq X_t + Y_t \qquad (t \geq 0). \tag{3.9}$$

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 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 (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*

$$E^x[f(X_t)] = S_t f(x) \qquad (f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda)),\ x \in \mathcal{E}_\gamma(\Lambda),\ t \geq 0). \tag{3.10}$$

We start with the following lemma.

Lemma 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 (1.14), with Lipschitz constant L, and K is the constant from (1.12), then*

$$|S_t f(x) - S_t f(y)| \leq L e^{(K+b-d)t} \|x - y\|_\gamma \qquad (x, y \in \mathcal{N}(\Lambda),\ t \geq 0). \tag{3.11}$$

Proof It follows from Propostion 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

$$\begin{aligned} \tfrac{\partial}{\partial t} E^x[\|X_t\|_\gamma] = &\sum_{ij} a(i,j)(\gamma_j - \gamma_i) E[X_t(i)] + (b - d) E^x[\|X_t\|_\gamma] \\ &- c \sum_i \gamma_i E[X_t(i)(X_t(i) - 1)] \ \leq\ (K + b - d) E[\|X\|_\gamma], \end{aligned} \tag{3.12}$$

and therefore

$$E^x[\|X_t\|_\gamma] \leq e^{(K+b-d)t} \|x\|_\gamma \qquad (x \in \mathcal{N}(\Lambda)). \tag{3.13}$$

Let X^x denote the (a, b, c, d)-braco-process started in x. By Lemma 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

$$E[\|X^x_t - X^y_t\|_\gamma] \leq E[\|X^{x\vee y}_t - X^{x\wedge y}_t\|_\gamma]. \tag{3.14}$$

By Lemma 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 (3.14) and (3.13),

$$E[\|X^x_t - X^y_t\|_\gamma] \leq E[\|X^{|x-y|}_t\|_\gamma] \leq \|x-y\|_\gamma e^{(K+b-d)t}, \tag{3.15}$$

which implies that

$$|S_tf(x) - S_tf(y)| \leq E[|f(X^x_t) - f(X^y_t)|] \leq LE[\|X^x_t - X^y_t\|_\gamma] \leq L\|x-y\|_\gamma e^{(K+b-d)t}, \tag{3.16}$$

as required. ∎

Since Lipschitz functions on $\mathcal{N}(\Lambda)$ have a unique Lipschitz extension to $\mathcal{E}_\gamma(\Lambda)$, Lemma 12 implies that S_tf 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 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 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 (3.15),

$$E\big[\|X_t - X^{(n)}_t\|_\gamma\big] = \lim_{m\uparrow\infty} E\big[\|X^{(m)}_t - X^{(n)}_t\|_\gamma\big] \leq \|x - x^{(n)}\|_\gamma e^{(K+b-d)t}. \tag{3.17}$$

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 (3.17),

$$\begin{aligned} |E[f(X_t)] - E[f(X^{(n)}_t)]| &\leq E[|f(X_t) - f(X^{(n)}_t)|] \\ &\leq LE[\|X_t - X^{(n)}_t\|_\gamma] \leq L\|x - x^{(n)}\|_\gamma e^{(K+b-d)t}, \end{aligned} \tag{3.18}$$

and therefore

$$E[f(X_t)] = \lim_{n\uparrow\infty} E[f(X^{(n)}_t)] = \lim_{n\uparrow\infty} S_tf(x^{(n)}) = S_tf(x). \tag{3.19}$$

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 $\int P_t(x,\mathrm{d}y)f(y) = S_tf(x)$ 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

$$E\big[f_1(X^{(n)}_{t_1}) \cdots f_k(X^{(n)}_{t_k})\big] = S_{t_1}f_1S_{t_2-t_1}f_2 \cdots S_{t_k-t_{k-1}}f_k(x^{(n)}). \tag{3.20}$$

It follows from (3.17) that

$$\big|E\big[f_1(X_{t_1}) \cdots f_k(X_{t_k})\big] - E\big[f_1(X^{(n)}_{t_1}) \cdots f_k(X^{(n)}_{t_k})\big]\big| \leq \|x - x^{(n)}\|_\gamma \sum_{i=1}^k L_ie^{(K+b-d)t_k} \prod_{j\neq i} \|f_j\|_\infty, \tag{3.21}$$

where L_i is the Lipschitz constant of f_i. Taking the limit $n \uparrow \infty$ in (3.20), using (3.21), we see that

$$E\big[f_1(X_{t_1})\cdots f_k(X_{t_k})\big] = S_{t_1} f_1 S_{t_2-t_1} f_2 \cdots S_{t_k-t_{k-1}} f_k(x), \tag{3.22}$$

i.e., X is the Markov process with semigroup $(S_t)_{t\geq 0}$. ∎

Proof of Proposition 11 We need to show that the process X from Lemma 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 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

$$\begin{aligned}
&G_{X,Y} f(x,y) := \\
&\sum_{ij} a(i,j)x(i)\{f(x+\delta_j-\delta_i, y+\delta_i) - f(x,y)\} + \sum_{ij} a(i,j)y(i)\{f(x,y+\delta_j) - f(x,y)\} \\
&+b\sum_i x(i)\{f(x+\delta_i,y) - f(x,y)\} + b\sum_i y(i)\{f(x,y+\delta_i) - f(x,y)\} \\
&+c\sum_i x(i)(x(i)-1)\{f(x-\delta_i,y+\delta_i) - f(x,y)\} + d\sum_i x(i)\{f(x-\delta_i,y+\delta_i) - f(x,y)\}.
\end{aligned} \tag{3.23}$$

and initial state $(X^{(n)}_0, Y^{(n)}_0) = (x^{(n)}, 0)$. 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

$$G_Z f(z) := \sum_{ij} a(i,j)z(i)\{f(z+\delta_j) - f(z)\} + b\sum_i z(i)\{f(z+\delta_i) - f(z)\} \tag{3.24}$$

and initial state $Z^{(n)}_0 = x^{(n)}$. In analogy with (3.13) it is easy to check that

$$E^z[\|Z^{(n)}_t\|_\gamma] \leq \|x^{(n)}\|_\gamma e^{(K+b)t} \qquad (z\in\mathcal{N}(\Lambda),\ t\geq 0). \tag{3.25}$$

$Z^{(n)}$ 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

$$|\{t\in[0,1] : X^{(n)}_{t-}(i) \neq X^{(n)}_t(i)\}| \leq x^{(n)}(i) + 2Z^{(n)}_1(i). \tag{3.26}$$

Summing with respect to the γ_i, taking expectations, using (3.25), we see that

$$\sum_i \gamma_i\, E\big[|\{t\in[0,1] : X^{(n)}_{t-}(i) \neq X^{(n)}_t(i)\}|\big] \leq \|x^{(n)}\|_\gamma (1 + 2e^{K+b}). \tag{3.27}$$

Let Z be the increasing limit of the processes $Z^{(n)}$. It follows from (3.25) that $Z_1 \in \mathcal{E}_\gamma(\Lambda)$ a.s. Now

$$X_t, X_{t-} \leq Z_t \leq Z_1 \qquad \forall t\in[0,1] \quad \text{a.s.}, \tag{3.28}$$

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,

$$|\{t\in[0,1] : X^{(n)}_{t-}(i) \neq X^{(n)}_t(i)\}| \uparrow |\{t\in[0,1] : X_{t-}(i) \neq X_t(i)\}| \quad \text{as } n\uparrow\infty. \tag{3.29}$$

Thus, taking the limit $n \uparrow \infty$ in (3.27) we see that

$$\sum_i \gamma_i \, E\big[|\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\}|\big] \leq \|x\|_\gamma (1 + 2e^{K+b}). \tag{3.30}$$

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,

$$\|X_{t_n} - X_t\|_\gamma = \sum_i \gamma_i |X_{t_n}(i) - X_t(i)| \to 0. \tag{3.31}$$

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 11 yields a useful corollary.

Corollary 14 (Locally finite number of jumps) *The (a,b,c,d)-braco-process X satisfies*

$$\sum_i \gamma_i \, E^x\big[|\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\}|\big] \leq \|x\|_\gamma (1 + 2e^{K+b}). \tag{3.32}$$

We can now prove two approximation lemmas.

Lemma 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*

$$\lim_{n \to \infty} \|x_n - x\|_\gamma = 0. \tag{3.33}$$

Then, for all $0 \leq t_1 < \cdots < t_k$, one has

$$(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_k}) \Rightarrow (X_{t_1}, \ldots, X_{t_k}) \qquad \text{as } n \to \infty. \tag{3.34}$$

Proof Use (3.22) for x_n and then let $n \to \infty$. ∎

Lemma 16 (Monotonicities for infinite systems) *Lemmas 9 and 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*

$$X^{x_n}_t(i) \uparrow X^x_t(i) \quad \text{as } n \uparrow \infty \quad \forall i \in \Lambda, \; t \geq 0 \quad \text{a.s.} \tag{3.35}$$

Proof The proof of Proposition 11 shows that (3.35) holds if the x_n are finite. To generalize Lemma 9 to infinite initial states $x, \tilde{x}$, it therefore suffices to note that if $x \leq \tilde{x}$, then there exist finite $x_n \leq \tilde{x}_n$ such that $x_n \uparrow x$ and $\tilde{x}_n \uparrow \tilde{x}$, and then take the limit $n \uparrow \infty$ in (3.8) using (3.35). Lemma 10 can be generalized to infinite x, y by approximation with finite x_n, y_n in the same way. Finally, to see that (3.35) remains valid if the x_n are infinite, note that by Lemma 9 (which has now been proved in the infinite case), the processes X^{x_n} can be coupled such that $X^{x_n}_t(i) \leq X^{x_{n+1}}_t(i)$ for all $i \in \Lambda$ and $t \geq 0$. Denote the increasing limit of the X^{x_n} by X^x. Lemma 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 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. ∎

3.4 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 $\mathcal{C}^2_{\rm fin}([0,1]^\Lambda)$ we denote the space of $\mathcal{C}^2$ functions on $[0,1]^\Lambda$ depending on finitely many coordinates. By definition, $\mathcal{C}^2_{\rm sum}([0,1]^\Lambda)$ 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

$$\phi\mapsto\big(\tfrac{\partial}{\partial\phi(i)}f(\phi)\big)_{i\in\Lambda}\quad\text{and}\quad\phi\mapsto\big(\tfrac{\partial^2}{\partial\phi(i)\partial\phi(j)}f(\phi)\big)_{i,j\in\Lambda}\tag{3.36}$$

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 $\mathcal{G}:\mathcal{C}^2_{\rm sum}([0,1]^\Lambda)\to\mathcal{C}([0,1]^\Lambda)$ by

$$\begin{aligned}\mathcal{G}f(\phi):=&\sum_{ij}a(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi)+b\sum_i\phi(i)(1-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi)\\&+c\sum_i\phi(i)(1-\phi(i))\tfrac{\partial^2}{\partial\phi(i)^2}f(\phi)-d\sum_i\phi(i)\tfrac{\partial}{\partial\phi(i)}f(\phi)\qquad(\phi\in[0,1]^\Lambda).\end{aligned}\tag{3.37}$$

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 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 (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 (1.3). Thus, existence of (weak) solutions to (1.3) follows from Lemma 17. Distribution uniqueness of solutions to (1.3) follows from pathwise uniqueness, which is in turn implied by the following comparison result.

Lemma 18 (Monotone coupling of linearly interacting diffusions) *Let $I\subset\mathbb{R}$ be a closed interval, let $\sigma:I\to\mathbb{R}$ be Hölder-$\frac12$-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*

$$\mathrm{d}\mathcal{X}^\alpha_t(i)=\sum_j a(j,i)(\mathcal{X}^\alpha_t(j)-\mathcal{X}^\alpha_t(i))\mathrm{d}t+b_\alpha(\mathcal{X}^\alpha_t(i))\mathrm{d}t+\sigma(\mathcal{X}^\alpha_t(i))\mathrm{d}B_t(i).\tag{3.38}$$

$(i \in \Lambda,\ t \geq 0,\ \alpha = 1,2)$. *Then*

$$\mathcal{X}_0^1 \leq \mathcal{X}_0^2 \quad \textit{implies} \quad \mathcal{X}_t^1 \leq \mathcal{X}_t^2 \quad \forall t \geq 0 \quad \text{a.s.} \tag{3.39}$$

Proof (sketch) Set $\Delta_t(i) := \mathcal{X}_t^1(i) - \mathcal{X}_t^2(i)$ 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

$$E[\|\Delta_t^+\|_\gamma] \leq (K+L) \int_0^t E[\|\Delta_s^+\|_\gamma]\mathrm{d}s, \tag{3.40}$$

where $\|\cdot\|_\gamma$ is the norm from (1.14), K is the constant from (1.12), and L is the Lipschitz-constant of b_2. The result now follows from Gronwall's inequality. ∎

Corollary 19 (Comparison of resampling-selection processes) *Assume that $\mathcal{X}, \tilde{\mathcal{X}}$ are solutions to the SDE (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*

$$\phi \leq \tilde{\phi}, \quad d - b \geq \tilde{d} - \tilde{b}, \quad d \geq \tilde{d}. \tag{3.41}$$

Then

$$\mathcal{X}_t \leq \tilde{\mathcal{X}}_t \quad \forall t \geq 0 \quad \text{a.s.} \tag{3.42}$$

Proof Immediate from Lemma 18 and the fact that by (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 (3.43) is not very good if $b - d < 0$, but it suffices for our purposes.

Lemma 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*

$$E^x\big[|\mathcal{X}_t|\big] \leq |x|e^{rt} \qquad (t \geq 0), \tag{3.43}$$

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 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,

$$\mathrm{d}\mathcal{Y}_t(i) = \sum_j a(j,i)(\mathcal{Y}_t(j) - \mathcal{Y}_t(i))\,\mathrm{d}t - be^{-rt}\mathcal{X}_t(i)^2\mathrm{d}t + e^{-rt}\sqrt{c\mathcal{X}_t(i)(1-\mathcal{X}_t(i))}\,\mathrm{d}B_t(i). \tag{3.44}$$

Set $\tau_N := \inf\{t \geq 0 : |\mathcal{X}_t| \geq N\}$. Integrate (3.44) up to $t \wedge \tau_N$ and sum over i. The motion terms yield

$$\begin{aligned} &\int_0^{t\wedge\tau_N} \sum_{ij} a(j,i)(\mathcal{Y}_s(j) - \mathcal{Y}_s(i))\,\mathrm{d}s \\ &\quad = \int_0^{t\wedge\tau_N} \sum_j \big(\sum_i a(j,i)\big)\mathcal{Y}_s(j)\,\mathrm{d}s - \int_0^{t\wedge\tau_N} \sum_i \big(\sum_j a^\dagger(i,j)\big)\mathcal{Y}_s(i)\,\mathrm{d}s = 0, \end{aligned} \tag{3.45}$$

where the infinite sums converge in a bounded pointwise way since $|Y_s| \leq N$ for $s \leq \tau_N$. It follows that

$$|\mathcal{Y}_{t\wedge\tau_N}| = |x| - b\sum_i \int_0^{t\wedge\tau_N} \mathcal{X}_s(i)^2 e^{-rs}\mathrm{d}s + \sum_i \int_0^{t\wedge\tau_N} \sqrt{c\mathcal{X}_s(i)(1-\mathcal{X}_s(i))}\, e^{-rs}\mathrm{d}B_s(i), \quad (3.46)$$

provided we can show that the infinite sum of stochastic integrals converges. Indeed, for any finite $\Delta \subset \Lambda$, by the Itô isometry,

$$\begin{aligned}
&\sum_{i\in\Delta} E\Big[\Big|\int_0^{t\wedge\tau_N} \sqrt{c\mathcal{X}_s(i)(1-\mathcal{X}_s(i))}\, e^{-rs}\mathrm{d}B_s(i)\Big|^2\Big] \\
&\qquad = c\sum_{i\in\Delta} E\Big[\int_0^{t\wedge\tau_N} \mathcal{X}_s(i)(1-\mathcal{X}_s(i))e^{-2rs}\mathrm{d}s\Big] \leq cE\Big[\int_0^{t\wedge\tau_N} |\mathcal{X}_s|\mathrm{d}s\Big] \leq ctN,
\end{aligned} \quad (3.47)$$

which shows that the stochastic integrals in (3.46) are absolutely summable in L^2-norm. It follows from (3.46) that

$$E^x[|\mathcal{X}_{t\wedge\tau_N}|]e^{-rt} \leq E^x[|\mathcal{X}_{t\wedge\tau_N}|e^{-r(t\wedge\tau_N)}] = E^x[|\mathcal{Y}_{t\wedge\tau_N}|] \leq |x|. \quad (3.48)$$

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 (3.48) we arrive at (3.43). ∎

We conclude this section with two results on the continuity of $\mathcal{X}$ in its initial state.

Lemma 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 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 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*

$$x^{(n)} \uparrow x \qquad \textit{as} \quad n \uparrow \infty. \quad (3.49)$$

Then $\mathcal{X}^{(n)}, \mathcal{X}$ may be defined on the same probablity space such that

$$\mathcal{X}^{(n)}_t(i) \uparrow \mathcal{X}_t(i) \qquad \forall i \in \Lambda,\ t \geq 0 \quad \textit{as} \quad n \uparrow \infty \quad \text{a.s.} \quad (3.50)$$

Proof Let $\mathcal{X}^{(n)}, \mathcal{X}$ be solutions of the SDE (1.3) relative to the same system of Brownian motions. By Corollary 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 18 shows that

$$\mathrm{d}\|\Delta^{(n)}_t\|_\gamma \leq (K+b)\|\Delta^{(n)}_t\|\mathrm{d}t \quad + \quad \text{martingale terms.} \quad (3.51)$$

It follows that

$$E\big[\|\Delta^{(n)}_{t\wedge\tau^{(n)}_\varepsilon}\|_\gamma\big] \leq \|x - x^{(n)}\|_\gamma e^{(K+b)t}. \quad (3.52)$$

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 $\varepsilon > 0$. ∎

4 Duality

4.1 Duality and self-duality

Proof of Theorem 1 (a) We first prove the statement for finite x. We apply Theorem 7. Our duality function is

$$\Psi(x,\phi) := (1-\phi)^x \qquad (x \in \mathcal{N}(\Lambda),\ \phi \in [0,1]^\Lambda). \tag{4.1}$$

We need to check that the right-hand side in (2.5) is zero, i.e., that

$$G\Psi(\cdot,\phi)(x) = \mathcal{G}^\dagger \Psi(x,\cdot)(\phi) \qquad (\phi \in [0,1]^\Lambda,\ x \in \mathcal{N}(\Lambda)), \tag{4.2}$$

where G be the generator of the (a,b,c,d)-braco-process, defined in (1.1), and $\mathcal{G}^\dagger$ is the generator of the $(a^\dagger,b,c,d)$-resem-process, defined in (3.37). Note that since x is finite, $\Psi(x,\cdot) \in \mathcal{C}^2_{\rm fin}([0,1]^\Lambda)$. We check that

$$\begin{aligned}
&G\Psi(\cdot,\phi)(x)\\
&= \sum_{ij} a(i,j)x(i)\{(1-\phi(j)) - (1-\phi(i))\}(1-\phi)^{x-\delta_i} + b\sum_i x(i)\{(1-\phi(i)) - 1\}(1-\phi)^x\\
&\quad + c\sum_i x(i)(x(i)-1)\{1-(1-\phi(i))\}(1-\phi)^{x-\delta_i} + d\sum_i x(i)\{1-(1-\phi(i)\}(1-\phi)^{x-\delta_i}\\
&= -\sum_{ij} a^\dagger(j,i)(\phi(j)-\phi(i))x(i)(1-\phi)^{x-\delta_i} - b\sum_i \phi(i)(1-\phi(i))x(i)(1-\phi)^{x-\delta_i}\\
&\quad + c\sum_i \phi(i)(1-\phi(i))x(i)(x(i)-1)(1-\phi)^{x-2\delta_i} + d\sum_i \phi(i)x(i)(1-\phi)^{x-\delta_i}\\
&\quad = \mathcal{G}^\dagger\Psi(x,\cdot)(\phi) \qquad (\phi \in [0,1]^\Lambda,\ x \in \mathcal{N}(\Lambda)).
\end{aligned} \tag{4.3}$$

Set

$$\Phi(x,\phi) := G\Psi(\cdot,\phi)(x) = \mathcal{G}^\dagger\Psi(x,\cdot)(\phi) \qquad (\phi \in [0,1]^\Lambda,\ x \in \mathcal{N}(\Lambda)). \tag{4.4}$$

It is not hard to see that there exists a constant K such that

$$|\Phi(x,\phi)| \le K\Big(1+|x|^2\Big) \qquad (\phi \in [0,1]^\Lambda,\ x \in \mathcal{N}(\Lambda)). \tag{4.5}$$

Therefore, condition (2.4) is satisfied by (3.1).

To generalize the statement from finite x to general $x \in \mathcal{E}_\gamma(\Lambda)$, we apply Lemma 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 $\phi \in [0,1]^\Lambda$,

$$E^\phi[(1-\mathcal{X}_t)^{x^{(n)}}] \downarrow E^\phi[(1-\mathcal{X}_t)^x] \quad \text{as } n \uparrow \infty, \tag{4.6}$$

and

$$E[(1-\phi)^{X^{(n)}_t}] \downarrow E[(1-\phi)^{X_t}] \quad \text{as } n \uparrow \infty, \tag{4.7}$$

where we used the continuity of the function $x \mapsto (1-\phi)^x$ with respect to increasing sequences.■

Proof of Theorem 1 (b) We first prove the statement under the additional assumption that ϕ and ψ are summable. Recall that by Lemma 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 7. Our duality function is

$$\Psi(\phi,\psi) := e^{-\frac{b}{c}\langle\phi,\psi\rangle} \qquad (\phi,\psi \in S). \tag{4.8}$$

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 (3.37), respectively. We need to show that the right-hand side in (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

$$
\begin{aligned}
\mathcal{G}\Psi(\cdot,\psi)(\phi) =& \Big\{ \sum_{ij} a(j,i)(\phi(j)-\phi(i))(-\tfrac{b}{c})\psi(i) + b\sum_i \phi(i)(1-\phi(i))(-\tfrac{b}{c})\psi(i) \\
&+ c\sum_i \phi(i)(1-\phi(i))(-\tfrac{b}{c})^2\psi(i)^2 - d\sum_i \phi(i)(-\tfrac{b}{c})\psi(i)\Big\} e^{-\frac{b}{c}\langle\phi,\psi\rangle} \\
=& -\tfrac{b}{c}\Big\{ \sum_{ij} a(j,i)\phi(j)\psi(i) - \Big(\sum_j a(j,i)\Big)\sum_i \phi(i)\psi(i) \\
&+ b\sum_i \phi(i)(1-\phi(i))\psi(i)(1-\psi(i)) - d\sum_i \phi(i)\psi(i)\Big\} e^{-\frac{b}{c}\langle\phi,\psi\rangle} \\
=& \,\mathcal{G}^\dagger\Psi(\phi,\cdot)(\psi).
\end{aligned}
\tag{4.9}
$$

It is not hard to see that there exists a constant K such that

$$
|\mathcal{G}\Psi(\cdot,\psi)(\phi)| \leq K|\phi|\,|\psi| \qquad (\phi,\psi \in S). \tag{4.10}
$$

Therefore, condition (2.4) is implied by Lemma 20, and Theorem 7 is applicable. To generalize the result to general $\phi,\psi \in [0,1]^\Lambda$, we apply Lemma 22. ∎

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

$$
\mathrm{d}\mathcal{Y}_t(i) = \sum_j a(j,i)(\mathcal{Y}_t(j)-\mathcal{Y}_t(i))\,\mathrm{d}t + \beta\mathcal{Y}_t(i)\,\mathrm{d}t + \sqrt{2\gamma\mathcal{Y}_t(i)}\,\mathrm{d}B_t(i) \tag{4.11}
$$

($t \geq 0$, $i \in \Lambda$). 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]

$$
E^\phi\big[e^{-\langle\mathcal{Y}_t,\psi\rangle}\big] = e^{-\langle\phi,\mathcal{U}_t\psi\rangle} \tag{4.12}
$$

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

$$
\tfrac{\partial}{\partial t}u_t(i) = \sum_j a(j,i)(u_t(j)-u_t(i)) + \beta u_t(i) - \gamma u_t(i)^2 \qquad (i \in \Lambda,\ t \geq 0) \tag{4.13}
$$

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 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*

$$E^x\big[e^{-\langle\phi,\,X_t\rangle}\big] \geq E^\phi\big[e^{-\langle\mathcal{Y}_t,\,x\rangle}\big] \qquad (x\in\mathcal{E}_\gamma(\Lambda),\ \phi\in\mathcal{M}(\Lambda)). \tag{4.14}$$

Proof We first prove the statement for finite x. We apply Theorem 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

$$\begin{aligned}\mathcal{H}f(\phi) :=& \sum_{ij} a^\dagger(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi) + (b-d+c)\sum_i \phi(i)\tfrac{\partial}{\partial\phi(i)}f(\phi)\\ &+c\sum_i \phi(i)\tfrac{\partial^2}{\partial\phi(i)^2}f(\phi) \qquad (\phi\in[0,1]^\Lambda),\end{aligned} \tag{4.15}$$

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

$$\begin{aligned}G\Psi(\cdot,\phi)(x) =& \Big\{\sum_{ij} a(i,j)x(i)\big(e^{\phi(i)-\phi(j)}-1\big) + b\sum_i x(i)\big(e^{-\phi(i)}-1\big)\\ &+c\sum_i x(i)(x(i)-1)\big(e^{\phi(i)}-1\big) + d\sum_i x(i)\big(e^{\phi(i)}-1\big)\Big\}e^{-\langle\phi,\,x\rangle},\end{aligned} \tag{4.16}$$

and

$$\begin{aligned}\mathcal{H}\Psi(x,\cdot)(\phi) =& \Big\{\sum_{ij} a^\dagger(j,i)x(i)(\phi(i)-\phi(j)) - (b-d+c)x(i)\phi(i)\\ &+c\sum_i x(i)^2\phi(i)\Big\}e^{-\langle\phi,\,x\rangle}\end{aligned} \tag{4.17}$$

$(x\in\mathcal{N}(\Lambda),\ \phi\in\mathcal{M}(\Lambda))$. It is not hard to see that there exists a constant K such that

$$|G\Psi(\cdot,\phi)(x)| \leq K|x|^2 \quad\text{and}\quad |\mathcal{H}\Psi(x,\cdot)(\phi)| \leq K|x|^2\,|\phi| \qquad (x\in\mathcal{N}(\Lambda),\ \phi\in\mathcal{M}(\Lambda)). \tag{4.18}$$

and therefore condition (2.4) is implied by (3.1) and the elementary estimate $E[|\mathcal{Y}_t|] \leq e^{(b-d+c)t}|\phi|$. One has

$$\begin{aligned}G\Psi(\cdot,\phi)(x) - \mathcal{H}\Psi(x,\cdot)(\phi) =& \Big\{\sum_{ij} a(i,j)x(i)\big(e^{\phi(i)-\phi(j)}-1-(\phi(i)-\phi(j))\big)\\ &+b\sum_i x(i)\big(e^{-\phi(i)}-1+\phi(i)\big) + c\sum_i x(i)(x(i)-1)\big(e^{\phi(i)}-1-\phi(i)\big)\\ &+d\sum_i x(i)\big(e^{\phi(i)}-1-\phi(i)\big)\Big\}e^{-\langle\phi,\,x\rangle} \geq 0,\end{aligned} \tag{4.19}$$

and therefore, for finite x, (4.14) is implied by Theorem 7. The general case follows by approximation, using Lemma 16. ∎

5 The Maximal Processes

5.1 The maximal branching-coalescing process

Using Proposition 23 we can now prove Theorem 2.

Proof of Theorem 2 Choose $x^{(n)} \in \mathcal{E}_\gamma(\Lambda)$ such that $x^{(n)}(i) \uparrow \infty$ for all $i \in \Lambda$. By Lemma 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 23 and (4.12),

$$E\big[1 - e^{-\langle \varepsilon\delta_i, X^{(n)}_t\rangle}\big] \leq 1 - e^{-\langle \varepsilon\delta_i, \mathcal{U}_t x^{(n)}\rangle} \qquad (t, \varepsilon \geq 0,\ i \in \Lambda). \tag{5.1}$$

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

$$E[X^{(n)}_t(i)] = \lim_{\varepsilon\downarrow 0} \varepsilon^{-1} E\big[1 - e^{-\langle \varepsilon\delta_i, X^{(n)}_t\rangle}\big] \leq \lim_{\varepsilon\downarrow 0} \varepsilon^{-1}\big(1 - e^{-\langle \varepsilon\delta_i, \mathcal{U}_t x^{(n)}\rangle}\big) = \mathcal{U}_t x^{(n)}(i) \tag{5.2}$$

$(t \geq 0,\ i \in \Lambda)$. Using the explicit solution of (4.13) for constant initial conditions, it is easy to see that $\mathcal{U}_t x^{(n)} \uparrow \mathcal{U}_t\infty$, where

$$\mathcal{U}_t\infty := \begin{cases} \frac{r}{c(1-e^{-rt})} & \text{if } r \neq 0, \\ \frac{1}{ct} & \text{if } r = 0. \end{cases} \tag{5.3}$$

(See, for example, [FS03a, formula (32)].) Letting $n \uparrow \infty$ in (5.2) we arrive at Theorem 2 (b). Moreover, we see that

$$E\big[\|X^{(\infty)}_t(i)\|_\gamma\big] \leq \mathcal{U}_t\infty \sum_i \gamma_i < \infty \qquad (t > 0), \tag{5.4}$$

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 16. Using Theorem 1 (a) and the continuity of the function $x \mapsto (1-\phi)^x$ with respect to increasing sequences, reasoning as in (1.28), we see that

$$P[\mathrm{Thin}_\phi(X^{(\infty)}_t) = 0] = P^\phi[\mathcal{X}^\dagger_t = 0] \qquad (\phi \in [0,1]^\Lambda,\ t \geq 0), \tag{5.5}$$

where $\mathcal{X}^\dagger$ denotes the $(a^\dagger, b, c, d)$-resem-process. Since formula (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 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 $\mathcal{L}(X^{(\infty)}_t) \downarrow \overline{\nu}$ for some probability measure $\overline{\nu}$ on $\mathcal{E}_\gamma(\Lambda)$. Set $\rho := \mathcal{L}(X^{(\infty)}_1)$ and let $(S_t)_{t\geq 0}$ denote the semigroup of the (a,b,c,d)-braco-process. Recall the definition of $\mathcal{C}_{\mathrm{Lip,b}}(\mathcal{E}_\gamma(\Lambda))$ above (3.20). One has

$$\int \overline{\nu}(\mathrm{d}x) f(x) = \lim_{t\to\infty} \int \rho(\mathrm{d}x) S_t f(x) \tag{5.6}$$

for every $f \in \mathcal{C}_{\mathrm{Lip,b}}(\mathcal{E}_\gamma(\Lambda))$. Therefore, since S_t maps $\mathcal{C}_{\mathrm{Lip,\,b}}(\mathcal{E}_\gamma(\Lambda))$ into itself,

$$\int \overline{\nu}(\mathrm{d}x) S_s f(x) = \lim_{t\to\infty} \int \rho(\mathrm{d}x) S_t S_s f(x) = \int \overline{\nu}(\mathrm{d}x) f(x) \qquad (s \geq 0), \tag{5.7}$$

for every $f \in \mathcal{C}_{\mathrm{Lip,\,b}}(\mathcal{E}_\gamma(\Lambda))$, 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. ∎

5.2 The maximal resampling-selection process

The proof of Theorem 3 (a)–(c) is similar to the proof of Theorem 2, but easier. Recall that Theorem 3 (d) is proved in Section 1.5.

Proof of Theorem 3 (a)–(c) Part (a) can be proved in the same way as Theorem 2 (d), using Lemma 22. The proof of part (b) goes analogue to the proof of Theorem 2 (e). To see why (1.30) holds, note that for any $\phi \in [0,1]^\Lambda$, by Theorem 1 (a),

$$\int \overline{\mu}(\mathrm{d}\phi)(1-\phi)^x = \lim_{t\to\infty} P^1[\mathrm{Thin}_{\mathcal{X}_t}(x) = 0] = \lim_{t\to\infty} P^x[\mathrm{Thin}_1(X_t^\dagger) = 0]. \tag{5.8}$$

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 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 (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. ∎

6 Convergence to the Upper Invariant Measure

6.1 Extinction versus unbounded growth

In this section we prove Lemma 5. It has already been proved in Section 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 25 below), and that $\mathcal{X}$ gets extinct in finite time if the limit is zero (Lemma 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 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 1 (a) and Theorem 2 (b),

$$\begin{aligned} P^\phi[\mathcal{X}_t \neq 0] &= \lim_{n\uparrow\infty} P^\phi[\mathrm{Thin}_{\mathcal{X}_t}(x^{(n)}) \neq 0] = \lim_{n\uparrow\infty} P[\mathrm{Thin}_\phi(X_t^{(n)\dagger}) \neq 0] \\ &= P[\mathrm{Thin}_\phi(X_t^{(\infty)\dagger}) \neq 0] \leq E\big[|\mathrm{Thin}_\phi(X_t^{(\infty)\dagger})|\big] = \langle \phi, E[X_t^{(\infty)\dagger}] \rangle \leq |\phi| \mathcal{U}_t \infty, \end{aligned} \tag{6.1}$$

where $\mathcal{U}_t\infty$ is the function on the right-hand side in (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 (6.1),

$$\tfrac{1}{2}1_{\{|\mathcal{X}_t| \leq \varepsilon\}} \leq P[\mathcal{X}_{t+t_0} = 0|\mathcal{F}_t] \leq P[\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0|\mathcal{F}_t]. \tag{6.2}$$

Now

$$1_{\{\lim_{s\to\infty} \mathcal{X}_s = 0\}} \leq \liminf_{t\to\infty} 1_{\{|\mathcal{X}_t| \leq \varepsilon\}}, \tag{6.3}$$

while

$$P[\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0|\mathcal{F}_t] \to 1_{\{\exists s \geq 0 \text{ s.t. } \mathcal{X}_s = 0\}} \quad \text{as } t \to \infty \quad \text{a.s.}, \tag{6.4}$$

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 (see, for example, [FS03a, Lemma A.1]). Letting $t \to \infty$ in (6.2), using (6.3) and (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 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 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 26 (Integrable fluctuations) *One has*

$$\int_0^\infty \sum_i \mathcal{X}_t(i)(1 - \mathcal{X}_t(i))\,\mathrm{d}t < \infty \tag{6.5}$$

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 $e^{-\langle\cdot,\psi\rangle} \in \mathcal{C}^2_{\rm sum}([0,1]^\Lambda)$ and (compare (4.9))

$$\begin{aligned} \mathcal{G}e^{-\langle\cdot,\,\psi\rangle}(\phi) = \Big\{ &-\sum_i \phi(i) \sum_j a^\dagger(j,i)(\psi(j) - \psi(i)) \\ &+ \sum_i \phi(i)(1-\phi(i))\big(c\psi(i)^2 - b\psi(i)\big) + d\sum_i \phi(i)\psi(i)\Big\} e^{-\langle\phi,\,\psi\rangle}. \end{aligned} \tag{6.6}$$

Since $\mathcal{X}$ solves the martingale problem for $\mathcal{G}$,

$$E\Big[\int_0^t \mathcal{G}e^{-\langle\cdot,\,\psi\rangle}(\mathcal{X}_s)\mathrm{d}s\Big] = E\big[e^{-\langle\mathcal{X}_t,\,\psi\rangle}\big] - e^{-\langle\phi,\,\psi\rangle} \qquad (t \geq 0). \tag{6.7}$$

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 (6.7), using Lemma 20, we find that

$$E\Big[\int_0^t \sum_i \Big\{\mu\mathcal{X}_s(i)(1-\mathcal{X}_s(i)) + \lambda d\mathcal{X}_s(i)\Big\} e^{-\lambda|\mathcal{X}_s|}\mathrm{d}s\Big] = E\big[e^{-\lambda|\mathcal{X}_t|}\big] - e^{-\langle\phi,\,\psi\rangle}. \tag{6.8}$$

Letting $t \uparrow \infty$, using the fact that the right-hand side of (6.8) is bounded by one, we see that

$$\int_0^\infty \sum_i \Big\{\mu\mathcal{X}_t(i)(1-\mathcal{X}_t(i)) + \lambda d\mathcal{X}_t(i)\Big\} e^{-\lambda|\mathcal{X}_t|}\,\mathrm{d}t < \infty \quad \text{a.s.}, \tag{6.9}$$

which implies (6.5). ∎

Lemma 27 (Process not started with only zeros and ones) *For every $0<\varepsilon<\frac{1}{4}$ there exists a $\delta, r>0$ such that*

$$P^{\phi}\big[\mathcal{X}_t(i)\in(\varepsilon,1-\varepsilon)\ \forall t\in[0,r]\big]\geq\delta\qquad(i\in\Lambda,\ \phi\in[0,1]^{\Lambda},\ \phi(i)\in(2\varepsilon,1-2\varepsilon)).\tag{6.10}$$

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 (6.10) is just a standard calculation, which we skip. ∎

Lemma 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*

$$\lim_{t\to\infty}\inf_{i\in\Delta}\mathcal{X}_t(i)=1\quad\textit{and}\quad\lim_{t\to\infty}\sup_{i\in\Lambda\backslash\Delta}\mathcal{X}_t(i)=0.\tag{6.11}$$

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 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 26. ∎

Lemma 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 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 $\lim_{t\to\infty}|\mathcal{X}_t|>0$,

$$\int_T^{\infty}\sum_i\mathcal{X}_t(i)(1-\mathcal{X}_t(i))\mathrm{d}t\geq\frac{1}{2}\int_T^{\infty}|\mathcal{X}_t|=\infty.\tag{6.12}$$

We arrive at a contradiction with Lemma 26. ∎

Proof of Proposition 25 Let Δ be the random set from Lemma 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 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 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 (1.3). (We skip the details.) ∎

6.2 Convergence to the upper invariant measure

In this section we complete the proof of Theorem 4, started in Section 1.5, by proving Lemma 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 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

$$\Gamma^\varepsilon_i := \{j \in \Lambda : P^i[\xi^{\mathrm{s}}_1 = j] \geq \varepsilon\} \qquad (i \in \Lambda). \tag{6.13}$$

We can choose $\varepsilon > 0$ small enough such that

$$j \notin \Gamma^\varepsilon_i \quad \text{implies} \quad g_i(\Delta) \cap g_j(\Delta) = \emptyset \quad (i, j \in \Lambda). \tag{6.14}$$

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

$$\phi_n \text{ assumes its maximum over } \Lambda \backslash \bigcup_{l=1}^{k} \Gamma^\varepsilon_{i_l} \text{ in } i_{k+1}. \tag{6.15}$$

Then $g_{i_1}(\Delta), g_{i_2}(\Delta), \ldots$ are disjoint by (6.14). Since $K := |\Gamma^\varepsilon_i|$ is finite and does not depend on i,

$$\sum_{l=1}^{\infty} \phi_n(i_l) \geq \frac{|\phi_n|}{K}, \tag{6.16}$$

and we can choose k_n such that

$$\lambda_n := \sum_{l=1}^{k_n} \phi_n(i_l) \underset{n \to \infty}{\longrightarrow} \infty. \tag{6.17}$$

Setting

$$\pi_n := \frac{1}{\lambda_n} \phi_n 1_{\{i_1, \ldots, i_{k_n}\}} \tag{6.18}$$

yields λ_n and π_n with the desired properties. ∎

Let $(\xi_t)_{t \geq 0}$ and $(\xi^\dagger_t)_{t \geq 0}$ 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

$$R\Delta := \{i \in \Lambda : P^i[\xi_t \in \Delta] > 0\} \quad \text{and} \quad R^\dagger\Delta := \{i \in \Lambda : P^i[\xi^\dagger_t \in \Delta] > 0\} \quad (t > 0) \tag{6.19}$$

of points from which ξ and $\xi^\dagger$ can enter Δ do not depend on $t > 0$. Indeed

$$R\Delta = \big\{i : \exists n \geq 0,\ i_0, \ldots, i_n \text{ s.t. } i_0 = i,\ i_n \in \Delta,\ a(i_{l-1}, i_l) > 0\ \forall l = 1, \ldots, n\big\} \tag{6.20}$$

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 31 (Points from which 0 can be reached) *If μ is a G-homogeneous and nontrivial probability measure on $\mathbb{N}^\Lambda$, then*

$$\mu\big(\{x : x|_{R\{0\}} = 0\}\big) = 0. \tag{6.21}$$

Proof Let Y be a $\mathbb{N}^\Lambda$-valued random variable with law μ. We will show that for any $\Delta \subset \Lambda$,

$$P\big[Y|_{R^\dagger R\Delta} = 0\big] = P\big[Y|_{R\Delta} = 0\big]. \tag{6.22}$$

Assume that (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 $(\xi^\dagger_t)_{t\geq 0}$ cannot escape from $R\Delta$ this implies that for any $t > 0$

$$P^i\big[Y(\xi^\dagger_0) \neq 0,\ Y(\xi^\dagger_s) = 0\ \forall s \geq t\big] > 0, \tag{6.23}$$

which contradicts the fact that $(Y(\xi^\dagger_t))_{t\geq 0}$ is stationary. This proves (6.22). Continuing this process, we see that

$$P\big[Y|_{R\{0\}} = 0\big] = P\big[Y|_{R^\dagger R\{0\}} = 0\big] = P\big[Y|_{RR^\dagger R\{0\}} = 0\big] = \cdots \tag{6.24}$$

By irreducibility, the sets $R\{0\}, R^\dagger R\{0\}, RR^\dagger R\{0\}, \ldots$ increase to Λ, and therefore, since μ is nontrivial,

$$P\big[Y|_{R\{0\}} = 0\big] = P\big[Y|_{\Lambda} = 0\big] = 0. \tag{6.25}$$

∎

Proof of Lemma 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 (1.1))

$$\begin{aligned} G^\Delta f(x) := & \sum_{i,j\in\Delta} a(i,j)x(i)\{f(x+\delta_j-\delta_i) - f(x)\} + \sum_{i\in\Delta, j\in\Lambda\backslash\Delta} a(i,j)x(i)\{f(x-\delta_i) - f(x)\} \\ & + b\sum_{i\in\Delta} x(i)\{f(x+\delta_i) - f(x)\} + c\sum_{i\in\Delta} x(i)(x(i)-1)\{f(x-\delta_i) - f(x)\} \\ & + d\sum_{i\in\Delta} x(i)\{f(x-\delta_i) - f(x)\}. \end{aligned} \tag{6.26}$$

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

$$X_t \leq \sum_{i=1}^n X^{\Delta_i}_t \qquad (t \geq 0) \tag{6.27}$$

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

$$x|_\Delta \neq 0 \quad \Rightarrow \quad P^x[X^\Delta_t(0) > 0] > 0. \tag{6.28}$$

Choose λ_n, π_n, and $\{g_i\}_{i\in\mathrm{supp}(\pi_n)}$ as in Lemma 30. Then, for deterministic $x \in \mathcal{E}_\gamma(\Lambda)$, we can estimate

$$\begin{aligned}
P^x\big[\mathrm{Thin}_{\phi_n}(X_t)=0\big] &\leq P^x\big[\mathrm{Thin}_{\lambda_n\pi_n}(X_t)=0\big] \\
&\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^x\big[\mathrm{Thin}_{\lambda_n\pi_n(i)}(X_t^{g_i(\Delta)}(i))=0\big] \\
&\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-\lambda_n\pi_n(i)X_t^\Delta(i)}\big] \\
&\leq \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-X_t^\Delta(i)}\big]^{\lambda_n\pi_n(i)},
\end{aligned} \tag{6.29}$$

where the $T_{g_i^{-1}}$ are shift operators as in (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 (6.29) and Hölder's inequality,

$$\begin{aligned}
P\big[\mathrm{Thin}_{\phi_n}(X_t)=0\big] &\leq \int P[X_0\in\mathrm{d}x] \prod_{i\in\mathrm{supp}(\pi_n)} P^{T_{g_i^{-1}}x}\big[e^{-X_t^\Delta(i)}\big]^{\lambda_n\pi_n(i)} \\
&\leq \prod_{i\in\mathrm{supp}(\pi_n)} \Big(\int P[X_0\in\mathrm{d}x]\, P^{T_{g_i^{-1}}x}\big[e^{-X_t^\Delta(i)}\big]^{\lambda_n}\Big)^{\pi_n(i)} \\
&= \int P[X_0\in\mathrm{d}x]\, P^x\big[e^{-X_t^\Delta(0)}\big]^{\lambda_n},
\end{aligned} \tag{6.30}$$

and therefore, by (6.28) and the fact that $\lambda_n\to\infty$,

$$\limsup_{n\to\infty} P\big[\mathrm{Thin}_{\phi_n}(X_t)=0\big] \leq P\big[X_0|_\Delta=0\big]. \tag{6.31}$$

Put

$$\Delta_k := \bigcup_{n=0}^{k} \big\{i : \exists i_0,\ldots,i_n \text{ s.t. } i_0=i,\ i_n=0,\ a(i_{l-1},i_l)>\tfrac{1}{k}\ \forall l=1,\ldots,n\big\}. \tag{6.32}$$

Then the Δ_k satisfy (6.28) and $\Delta_k\uparrow R\{0\}$ as $k\uparrow\infty$, where $R\{0\}$ is defined in (6.20). Therefore, inserting $\Delta=\Delta_k$ in (6.31) and taking the limit $k\uparrow\infty$, using Lemma 31, we arrive at (1.34).∎

References

[BEM03] J. Blath, A.M. Etheridge, and M.E. Meredith. Coexistence in locally regulated competing populations. University of Oxford, preprint, 2003.

[BES02] N.H. Barton, A.M. Etheridge, and A.K. Sturm. Coalescence in a random background. WIAS Berlin, Preprint No. 756, 2002. To appear in *Ann. Appl. Probab.*

[Che87] M.F. Chen. Existence theorems for interacting particle systems with non-compact state space. *Sci. China Ser. A* 30, 148–156, 1987.

[Daw93] D.A. Dawson. Measure-valued Markov processes. In P.L. Hennequin, editor, *École d'été de probabilités de Saint Flour XXI–1991*, volume 1541 of *Lecture Notes in Mathematics*, pages 1–260. Springer, Berlin, 1993.

[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.

[DDL90] W. Ding, R. Durrett, and T.M. Liggett. Ergodicity of reversible reaction diffusion processes. *Probab. Theory Relat. Fields* 85(1), 13–26, 1990.

[DK99] P. Donelly and T.G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.* 9, 1091-1148, 1999.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes; Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[FS03a] K. Fleischmann and J.M. Swart. Extinction versus exponential growth in a supercritical super-Wright-Fisher diffusion. *Stochastic Processes Appl.* 106(1), 141-165, 2003.

[FS03b] K. Fleischmann and J.M. Swart. Trimmed trees and embedded particle systems. WIAS Berlin, Preprint No. 793, 2002. To appear in *Ann. Probab.*

[Kal83] O. Kallenberg. *Random measures*, 3rd rev. and enl. ed. Akademie-Verlag, Berlin, 1983.

[KN97] S.M. Krone and C. Neuhauser. Ancestral processes with selection. *Theor. Popul. Biol.* 51(3), 210–237, 1997.

[Lig99] T.M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Process.* Springer-Verlag, Berlin, 1999.

[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.

[Neu90] C. Neuhauser. An ergodic theorem for Schlögl models with small migration. *Probab. Theory Relat. Fields* 85(1), 27–32, 1990.

[Pat88] A.L.T. Paterson. *Amenability.* American Mathematical Soc., Providence, 1988.

[Sch72] F. Schlögl. Chemical reaction models and non-equilibrium phase transitions. *Z. Phys.* 253, 147–161, 1972.

[Shi81] T. Shiga. Diffusion processes in population genetics. *J. Math. Kyoto Univ.* 21, 133–151, 1981.

[SS80] T. Shiga and A. Shimizu. Infinite dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* 20, 395–416, 1980.

[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_j/largspscb.pdf.

[Swa00] J.M. Swart. Clustering of linearly interacting diffusions and universality of their long-time limit distribution. *Probab. Theory Relat. Fields* 118, 574–594, 2000.

[Swa04] J.M. Swart. A spatial ergodic theorem for nonamenable groups. In preparation 2004.

[YW71] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* 11, 155–167, 1971.

Siva R. Athreya

Stat. Math. Unit
Indian Statistical Institute
Delhi Centre 7
S.J.S. Sansanwal Marg
New Delhi -110016, India

e-mail: athreya@isid.ac.in

Jan M. Swart

Mathematical Institute
University Erlangen-Nuremberg
Bismarckstraße $1\frac{1}{2}$
D–91054 Erlangen, Germany

e-mail: swart@math.tu-berlin.de