Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 80, 1–32.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2003

Systems of branching, annihilating, and coalescing particles

Siva R. Athreya* Jan M. Swart†

Abstract

This paper studies systems of particles following independent random walks and subject to annihilation, binary branching, coalescence, and deaths. In the case without annihilation, such systems have been studied in our 2005 paper "Branching-coalescing particle systems". The case with annihilation is considerably more difficult, mainly as a consequence of the non-monotonicity of such systems and a more complicated duality. Nevertheless, we show that adding annihilation does not significantly change the long-time behavior of the process and in fact, systems with annihilation can be obtained by thinning systems without annihilation.

Keywords: Reaction-diffusion process; branching; coalescence; annihilation; thinning; Poissonization..
AMS MSC 2010: Primary: 82C22; Secondary: 60K35, 92D25.
Submitted to EJP on May 4, 2012, final version accepted on September 17, 2012.
Supersedes arXiv:1203.6477.

1 Results

1.1 Introduction

In [1], we studied systems of particles that perform independent random walks, branch binarily, coalesce, and die. Our motivation came from two directions. On the one hand, we were driven by the wish to study a population dynamic model that is more realistic than the usual branching particle systems, since the population at a given site cannot grow unboundedly but is instead controlled by an extra death term that is quadratic in the number of particles, which can be interpreted as extra deaths due to competition. On the other hand, such systems of branching and coalescing particles are known to be dual to certain systems of interacting diffusions, modelling gene frequencies in spatially structured populations subject to resampling, mutation, and selection [25]. In this context, the branching-coalecing particles can be interpreted as 'potential ancestors' [17].

*Indian Statistical Institute, Bangalore, India. E-mail: athreya@isibang.ac.in
†Institute of Information Theory and Automation of the ASCR (ÚTIA), Praha, Czech Republic.
E-mail: swart@utia.cas.cz

Apart from this duality, which was known, we showed in [1] that our particle systems are also related to resampling-selection processes by a Poissonization relation. Moreover, we proved that systems started with infinitely many particles on each site come down from infinity (a fact that had been proved before, with a less explicit bound, in [8]) and that systems on quite general spatially homogeneous lattices have at most one nontrivial homogeneous invariant law, which, if it exists, is the long-time limit law of the process started in any nontrivial homogeneous initial law.

In the present paper, we generalize all these results to systems where moreover, with some positive rate, pairs of particles on the same site *annihilate* each other, resulting in the disappearance of *both* particles. This my not seem like it should make a big difference with coalescence, where only one particle disappears -and indeed our results confirm this- but from the technical point of view annihilation has the huge disadvantage of making the system non-monotone, which means that many simple comparison arguments are not available. Some pioneering work on non-monotone systems can be found in, e.g., [5, 26, 10]. Despite progress in recent years, non-monotone particle systems are still generally less studied and worse understood than monotone ones.

As in the case without annihilation, our main tool is duality. In fact, it turns out that systems with annihilation are dual to the same Markov process (a system of interacting Wright-Fisher diffusions) as those without it, but with a different (and more complicated) duality function. As a result, we obtain Poissonization and thinning relations which show, among others, that systems with annihilation can be obtained from systems without it by independent thinning. We reported these duality and thinning relations before (without proof) in [29].

The paper is organized a follows. In Section 1.2 we define our model and the dual system of interacting diffusions. In Section 1.3 we state our duality result and show how this implies Poissonization and thinning relations. Section 1.4 presents our main results, showing that the system started with infinitely many particles comes down from infinity and that systems started in a spatially homogeneous, nontrivial invariant law converge to a unique homogeneous invariant law. Section 1.5 contains more discussion and an overview of our proofs, which are given in Section 2.

1.2 Definition of the models

Let Λ be a finite or countably infinite set and let $q(i,j) \geq 0$ $(i,j \in \Lambda,\ i \neq j)$ be the transition rates of a continuous time Markov process on Λ, the *underlying motion*, which jumps from site i to site j with rate $q(i,j)$. For notational convenience, we set $q(i,i) := 0$ $(i \in \Lambda)$. We assume that the rates $q(i,j)$ are uniformly summable and (in a weak sense) irreducible, and that the counting measure on Λ is an invariant law for the underlying motion, i.e.:

$$\begin{aligned}
&\text{(i)} \quad \sup_i \sum_j q(i,j) < \infty,\\
&\text{(ii)} \quad \forall \Delta \subset \Lambda,\ \Delta \neq \emptyset, \Lambda\ \ \exists i \in \Delta,\ j \in \Lambda\backslash\Delta \text{ such that } q(i,j) > 0 \text{ or } q(j,i) > 0,\\
&\text{(iii)} \quad \sum_j q^\dagger(i,j) = \sum_j q(i,j)\ \forall i \in \Lambda,\ \text{where } q^\dagger(i,j) := q(j,i).
\end{aligned} \tag{1.1}$$

Here and elsewhere sums and suprema over i,j always run over Λ, unless stated otherwise.

Branching-annihilating particle systems. We let (Λ, q) be as above, fix rates $a,b,c,d \geq 0$, and consider systems of particles subject to the following dynamics.

1° Each particle jumps, independently of the others, from site i to site j with rate $q(i,j)$.

2° Each pair of particles, present on the same site, annihilates with rate $2a$, resulting in the disappearance of both particles.

3° Each particle splits with rate b into two new particles, created on the position of the old one.

4° Each pair of particles, present on the same site, coalesces with rate $2c$, resulting in the creation of one new particle on the position of the two old ones.

5° Each particle dies (disappears) with rate d.

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

$$\begin{aligned}Gf(x) := &\sum_{ij} q(i,j)x(i)\{f(x+\delta_j-\delta_i)-f(x)\} + a\sum_i x(i)(x(i)-1)\{f(x-2\delta_i)-f(x)\}\\ &+b\sum_i x(i)\{f(x+\delta_i)-f(x)\} + c\sum_i x(i)(x(i)-1)\{f(x-\delta_i)-f(x)\}\\ &+d\sum_i x(i)\{f(x-\delta_i)-f(x)\},\end{aligned} \tag{1.2}$$

where $\delta_i(j) := 1$ if $i = j$ and $\delta_i(j) := 0$ otherwise. We call X the (q,a,b,c,d)-*branco-process*.

The process X can be defined for finite initial states and also for some infinite initial states in an appropriate Liggett-Spitzer space. Following [19], we define

$$\mathcal{E}_\gamma(\Lambda) := \{x \in \mathbb{N}^\Lambda : \|x\|_\gamma < \infty\}, \quad \text{with} \quad \|x\|_\gamma := \sum_i \gamma_i |x(i)|, \tag{1.3}$$

where $\gamma = (\gamma_i)_{i\in\Lambda}$ are strictly positive constants satisfying

$$\sum_i \gamma_i < \infty \quad \text{and} \quad \sum_j (q(i,j) + q^\dagger(i,j))\gamma_j \leq K\gamma_i \quad (i \in \Lambda) \tag{1.4}$$

for some $K < \infty$. (Our assumptions on q imply the existence of a γ satisfying (1.4).)

Resampling selection processes. Let (Λ, q) be as before, let r, s, m be nonnegative constants, and let $\mathcal{X} = (\mathcal{X}_t)_{t\geq 0}$ be the $[0,1]^\Lambda$-valued Markov process given by the unique pathwise solutions to the infinite dimensional stochastic differential equation (SDE) (see [25, 1]):

$$\begin{aligned}\mathrm{d}\mathcal{X}_t(i) = &\sum_j q(j,i)(\mathcal{X}_t(j) - \mathcal{X}_t(i))\,\mathrm{d}t + s\mathcal{X}_t(i)(1-\mathcal{X}_t(i))\,\mathrm{d}t - m\mathcal{X}_t(i)\,\mathrm{d}t\\ &+\sqrt{2r\mathcal{X}_t(i)(1-\mathcal{X}_t(i))}\,\mathrm{d}B_t(i) \qquad (t\geq 0,\ i\in\Lambda),\end{aligned} \tag{1.5}$$

where $(B(i))_{i\in\Lambda}$ is a collection of independent Brownian motions. The process $\mathcal{X}$ is a system of linearly interacting Wright-Fisher diffusions, also known as stepping stone model, which can be used to model the spatial distribution of gene frequencies in the presence of resampling, selection, and mutation. Following [1], we call $\mathcal{X}$ the resampling-selection process with underlying motion (Λ, q), resampling rate r, selection rate s, and mutation rate m, or shortly the (q,r,s,m)-*resem-process*.

1.3 Duality, Poissonization, and thinning

We start with some 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. For any $\phi : \Lambda \to [-1,1]$ and $x : \Lambda \to \mathbb{N}$ we write

$$\phi^x := \prod_i \phi(i)^{x(i)} \quad \text{with} \quad 0^0 := 1 \tag{1.7}$$

whenever the infinite product converges and the limit does not depend on the order of the coordinates. The following proposition generalizes [1, Theorem 1 (a)].

Proposition 1.1. (Duality) *Assume that $a + c > 0$ and let*

$$\alpha = a/(a+c), \quad r = a + c, \quad s = (1+\alpha)b, \quad \textit{and} \quad m = \alpha b + d, \tag{1.8}$$

or equivalently

$$a = \alpha r, \quad b = s/(1+\alpha), \quad c = (1-\alpha)r, \quad \textit{and} \quad d = m - \alpha s/(1+\alpha). \tag{1.9}$$

Let X be a (q,a,b,c,d)-branco-process with $X_0 \in \mathcal{E}_\gamma(\Lambda)$ a.s. and let $\mathcal{X}^\dagger$ be a $(q^\dagger, r, s, m)$-resem-process, independent of X. Suppose that one or more of the following conditions are satisfied:

$$\text{(i) } \alpha < 1, \quad \text{(ii) } |X_0| < \infty \text{ a.s.}, \quad \text{(iii) } |\mathcal{X}^\dagger_0| < \infty \text{ a.s.} \tag{1.10}$$

Then

$$\mathbb{E}\big[(1-(1+\alpha)\mathcal{X}^\dagger_0)^{X_t}\big] = \mathbb{E}\big[(1-(1+\alpha)\mathcal{X}^\dagger_t)^{X_0}\big] \qquad (t \ge 0), \tag{1.11}$$

where the infinite products inside the expectation are a.s. well-defined.

Proposition 1.1, together with a self-duality for (q,r,s,m)-resem-processes described in [1, Theorem 1 (b)], implies that (q,a,b,c,d)-branco-processes can be obtained as Poissonizations of resampling-selection processes, and as thinnings of each other, as we explain now. (These thinning relations will prove useful several times in what will follow. On the other hand, we have no application of the Poissonization relations, but since they are very similar and closely related, we treat them here as well.)

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

$$\mathbb{E}\big[(1-\psi)^{\mathrm{Pois}(\phi)}\big] = \mathbb{E}\big[e^{-\langle \phi, \psi\rangle}\big] \qquad (\psi \in [0,1]^\Lambda), \tag{1.12}$$

where we allow for the case that $e^{-\langle \phi, \psi\rangle} = e^{-\infty} := 0$. In particular, if ϕ is nonrandom, then the components $(\mathrm{Pois}(\phi)(i))_{i\in\Lambda}$ are independent Poisson distributed random variables with intensity $\phi(i)$.

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

$$\mathbb{E}\big[(1-\psi)^{\mathrm{Thin}_\phi(x)}\big] = \mathbb{E}\big[(1-\phi\psi)^x\big] \qquad (\psi \in [0,1]^\Lambda). \tag{1.13}$$

In particular, when x and ϕ are nonrandom and $x = \sum_n \delta_{i_n}$, then a ϕ-thinning of x can be constructed as $\mathrm{Thin}_\phi(x) := \sum_n \chi_n \delta_{i_n}$ where the χ_n are independent $\{0,1\}$-valued random variables with $\mathbb{P}[\chi_n = 1] = \phi(i_n)$. More generally, if x and ϕ are random, then we may construct $\mathrm{Thin}_\phi(x)$ in such a way that its conditional law given x and ϕ is as in the deterministic case. It is not hard to check that (1.13) holds more generally for any $\psi \in [0,2]^\Lambda$ provided $(1-\psi)^{\mathrm{Thin}_\phi(x)}$ is a.s. well-defined.

Proposition 1.2. (Poissonization and thinning) *Fix* $s, m \geq 0$, $r > 0$, *and* $0 \leq \beta \leq \alpha \leq 1$ *such that* $m - \frac{\beta}{1+\beta} s \geq 0$. *Let* X *and* $\overline{X}$ *be the* $(q, \alpha r, \frac{1}{1+\alpha} s, (1-\alpha) r, m - \frac{\alpha}{1+\alpha} s)$*-branco-process and* $(q, \beta r, \frac{1}{1+\beta} s, (1-\beta) r, m - \frac{\beta}{1+\beta} s)$*-branco-process, respectively, and let* $\mathcal{X}$ *be the* (q, r, s, m)*-resem-process. Then*

$$\mathbb{P}\big[X_0 \in \cdot\,\big] = \mathbb{P}\big[\mathrm{Pois}(\tfrac{s}{(1+\alpha)r} \mathcal{X}_0) \in \cdot\,\big] \quad \textit{implies} \quad \mathbb{P}\big[X_t \in \cdot\,\big] = \mathbb{P}\big[\mathrm{Pois}(\tfrac{s}{(1+\alpha)r} \mathcal{X}_t) \in \cdot\,\big] \tag{1.14}$$

$(t \geq 0)$, *and*

$$\mathbb{P}\big[X_0 \in \cdot\,\big] = \mathbb{P}\big[\mathrm{Thin}_{\frac{1+\beta}{1+\alpha}}(\overline{X}_0) \in \cdot\,\big] \quad \textit{implies} \quad \mathbb{P}\big[X_t \in \cdot\,\big] = \mathbb{P}\big[\mathrm{Thin}_{\frac{1+\beta}{1+\alpha}}(\overline{X}_t) \in \cdot\,\big] \qquad (t \geq 0). \tag{1.15}$$

Proof. Formula (1.14) has been proved in case $\alpha = 0$ in [1]. The general case can be derived along the same lines. Alternatively, this can be derived from the case $\alpha = 0$ using the fact that $\mathbb{P}[\mathrm{Thin}_{\frac{1}{1+\alpha}}(\mathrm{Pois}(\frac{s}{r}\mathcal{X}_t)) \in \cdot\,] = \mathbb{P}[\mathrm{Pois}(\frac{s}{(1+\alpha)r}\mathcal{X}_t) \in \cdot\,]$, and formula (1.15), which we prove now.

If the initial laws of X and $\overline{X}$ are related as in (1.15) and $\mathcal{X}^\dagger$ is a $(q^\dagger, r, s, m)$-resem-process started in $\mathcal{X}_0 = \phi$ with $|\phi| < \infty$, then by (1.11),

$$\begin{aligned} &\mathbb{E}\big[(1-(1+\alpha)\phi)^{\mathrm{Thin}_{\frac{1+\beta}{1+\alpha}}(\overline{X}_t)}\big] = \mathbb{E}\big[(1-(1+\beta)\phi)^{\overline{X}_t}\big] = \mathbb{E}\big[(1-(1+\beta)\mathcal{X}^\dagger_t)^{\overline{X}_0}\big] \\ &= \mathbb{E}\big[(1-(1+\alpha)\mathcal{X}^\dagger_t)^{\mathrm{Thin}_{\frac{1+\beta}{1+\alpha}}(\overline{X}_0)}\big] = \mathbb{E}\big[(1-(1+\alpha)\mathcal{X}^\dagger_t)^{X_0}\big] = \mathbb{E}\big[(1-(1+\alpha)\phi)^{X_t}\big] \end{aligned} \tag{1.16}$$

$(t \geq 0)$, where we have used that by [1, Lemma 20] one has $|\mathcal{X}^\dagger_t| < \infty$ a.s. for each $t \geq 0$, which guarantees that the infinite products are a.s. well-defined. Since (1.16) holds for all $\phi \in [0,1]^\Lambda$ with $|\phi| < \infty$, (1.15) follows. □

As an immediate corollary of formula (1.15), we have:

Corollary 1.3. (Thinnings of processes without annihilation) *Let* $a, b, c, d \geq 0$ *and* $a + c > 0$. *Let* X *be the* (q, a, b, c, d)*-branco-process,* $\alpha := \frac{a}{a+c}$, *and let* $\overline{X}$ *be the* $(q, 0, (1+\alpha)b, a+c, \alpha b + d)$*-branco-process. Then*

$$\mathbb{P}[X_0 \in \cdot\,] = \mathbb{P}[\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}_0) \in \cdot\,] \quad \textit{implies} \quad \mathbb{P}[X_t \in \cdot\,] = \mathbb{P}[\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}_t) \in \cdot\,] \qquad (t \geq 0). \tag{1.17}$$

In particular, each branco-process with a positive annihilation rate can be obtained as a thinning of a process with zero annihilation rate.

1.4 Main results

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 $\overline{\mathbb{N}}^\Lambda$.

Our first main result shows that it is possible to start a (q, a, b, c, d)-branco-process with infinitely many particles at each site. We call this the (q, a, b, c, d)-branco process started at infinity. This result generalizes [1, Theorem 2]. For branching-coalescing particle systems on $\mathbb{Z}^d$ with more general branching and coalescing mechanisms, but without annihilation, a similar result has been proved in [8].

Theorem 1.4. (The maximal process) *Assume that* $a + c > 0$. *Then there exists an* $\mathcal{E}_\gamma(\Lambda)$*-valued process* $X^{(\infty)} = (X^{(\infty)}_t)_{t>0}$ *with the following properties:*

(a) *For each* $\varepsilon > 0$, $(X^{(\infty)}_t)_{t \geq \varepsilon}$ *is the* (q, a, b, c, d)*-branco-process starting in* $X^{(\infty)}_\varepsilon$.

(b) *Set $r := a + b + c - d$. Then*

$$\mathbb{E}[X_t^{(\infty)}(i)] \leq \begin{cases} \dfrac{r}{(2a+c)(1-e^{-rt})} & \text{if } r \neq 0, \\ \dfrac{1}{(2a+c)t} & \text{if } r = 0 \end{cases} \qquad (i \in \Lambda). \tag{1.18}$$

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

$$x^{(n)}(i) \uparrow \infty \quad \text{as} \quad n \uparrow \infty \qquad (i \in \Lambda), \tag{1.19}$$

then

$$\mathcal{L}(X_t^{(n)}) \underset{n\to\infty}{\Longrightarrow} \mathcal{L}(X_t^{(\infty)}) \qquad (t > 0). \tag{1.20}$$

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

$$\mathcal{L}(X_t^{(\infty)}) \underset{t\to\infty}{\Longrightarrow} \overline{\nu}. \tag{1.21}$$

(e) *The measure $\overline{\nu}$ is uniquely characterised by*

$$\int \overline{\nu}(\mathrm{d}x)(1-(1+\alpha)\phi)^x = \mathbb{P}^\phi[\exists t \geq 0 \text{ such that } \mathcal{X}_t^\dagger = 0] \qquad (\phi \in [0,1]^\Lambda,\ |\phi| < \infty), \tag{1.22}$$

where $\alpha := a/(a+c)$ and $\mathcal{X}^\dagger$ denotes the $(q^\dagger, a+c, (1+\alpha)b, \alpha b+d))$-resem-process started in ϕ.

(f) *If r, s, m, α, β are as in Proposition 1.2 and $X^{(\infty)}$ and $\overline{X}^{(\infty)}$ are the corresponding branco-processes started at infinity, then*

$$\mathbb{P}[X_t^{(\infty)} \in \cdot\,] = \mathbb{P}[\mathrm{Thin}_{\frac{1+\alpha}{1+\beta}}(\overline{X}_t^{(\infty)}) \in \cdot\,] \qquad (t \geq 0). \tag{1.23}$$

A similar thinning relation holds between their long-time limit laws.

If $a = 0$, then it has been shown in [1, Theorem 2 (e)] that $\overline{\nu}$ dominates any other invariant measure in the stochastic order, hence $\overline{\nu}$ can righteously be called the *upper invariant measure* of the process. In the general case, when we have annihilation, we do not know how to compare $\overline{\nu}$ with other invariant measures in the stochastic order, and we only work with the characterization of $\overline{\nu}$ in (1.22).

To formulate our final result, we need some definitions. Let (Λ, q) be our lattice with jump kernel of the underlying motion, as before. By definition, an *automorphism* of (Λ, q) is a bijection $g : \Lambda \to \Lambda$ such that $q(gi, gj) = q(i,j)$ for all $i,j \in \Lambda$. We denote the group of all automorphisms of (Λ, q) by $\mathrm{Aut}(\Lambda, q)$. We say that a subgroup $G \subset \mathrm{Aut}(\Lambda, q)$ is *transitive* if for each $i, j \in \Lambda$ there exists a $g \in G$ such that $gi = j$. We say that (Λ, q) is *homogeneous* if $\mathrm{Aut}(\Lambda, q)$ is transitive. We define shift operators $T_g : \mathbb{N}^\Lambda \to \mathbb{N}^\Lambda$ by

$$T_g x(j) := x(g^{-1}j) \qquad (i \in \Lambda,\ x \in \mathbb{N}^\Lambda,\ g \in \mathrm{Aut}(\Lambda, q)). \tag{1.24}$$

If G is a subgroup of $\mathrm{Aut}(\Lambda, q)$, 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 $q(i,j) = 1_{\{|i-j|=1\}}$ (nearest-neighbor random walk), then the group G of translations $i \mapsto i + j$ ($j \in \Lambda$) is a transitive subgroup of $\mathrm{Aut}(\Lambda, q)$ and the G-homogeneous probability measures are the translation invariant probability measures.

The next theorem, which generalizes [1, Theorem 4 (a)], is our main result.

Theorem 1.5. **(Convergence to the upper invariant measure)** *Assume that (Λ, q) is infinite and homogeneous, G is a transitive subgroup of $\mathrm{Aut}(\Lambda, q)$, and $a + c > 0$. Let X be the (q,a,b,c,d)-branco 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 measure in (1.22).*

1.5 Discussion and outline

The dualities in Proposition 1.1 and [1, Theorem 1 (b)], as well as the Poissonization and thinning relations in Proposition 1.2 play a central role in the present paper. These relations, whose discovery was the starting point of the present work, are similar to duality and thinning relations between general nearest-neighbor interacting particle systems discovered by Lloyd and Sudbury in [23, 24, 27]. In fact, as has been demonstrated in [29, Prop. 6 and Lemma 7] (see also the more detailed preprint of the same paper, [30, Prop 4.2 and Lemma 4.3]), our relations can (at least formally) be obtained as 'local mean field' limits of (a special case of) the relations of Lloyd and Sudbury. In [24], Lloyd and Sudbury observed that quite generally, if two interacting particle systems have the same dual (whith a special sort of duality relation as described in that article), then one is a thinning of the other. This general principle is also responsible for the Poissonization and thinning relations of our Proposition 1.2.

The thinning relation in Corollary 1.3 is especially noteworthy, since it allows us to compare non-monotone systems (which are generally hard to study) with monotone systems. Also, the thinning relation (1.23) allows us to prove that the unique nontrivial homogeneous invariant measures of $(q, \alpha r, \frac{1}{1+\alpha}s, (1-\alpha)r, m - \frac{\alpha}{1+\alpha}s)$-branco-processes are monotone in α (w.r.t. to the stochastic order). Such sort of comparison results between non-monotone systems are rarely available. In fact, these thinning relations suggest that the ergodic behavior of $(q, \alpha r, \frac{1}{1+\alpha}s, (1-\alpha)r, m - \frac{\alpha}{1+\alpha}s)$-branco-processes (with r, s, m fixed but arbitrary α) and the (r, s, m)-resem process should all be 'basically the same'.

It does not seem straightforward to make this claim rigorous, however. The reason is that Poissonization or thinning can only produce certain initial laws. Thus, an ergodic result for resampling-selection processes, as has been proved in [25], only implies an ergodic result for branching-annihilating particle systems started in initial laws that are Poisson with random intensity, and likewise, the ergodic result for branching-annihilating particle systems in [1] implies our Theorem 1.5 only for special initial laws, that are thinnings of other laws.

Our main tool for proving the statement for general initial laws is, like in our previous paper, duality. In this respect, our methods differ from those in [8], which are based on entropy calculations, but are similar to those used in, for example, [25, 9, 1, 22]. The papers [25, 1] are particularly close in spirit. The sort of cancellative systems type duality that we have to use in the present paper is somewhat harder to work with than the additive systems type duality in [25, 1]. Earlier applications of this sort of 'cancellative' duality can be found in [9, 22].

The remainder of this paper is devoted to proofs. Proposition 1.1 and Theorems 1.4 and 1.5 are proved in Sections 2.5, 2.6 and 2.8, respectively.

2 Proofs

2.1 Construction and approximation

2.2 Finite systems

We denote the set of finite particle configurations by $\mathcal{N}(\Lambda) := \{x \in \mathbb{N}^\Lambda : |x| < \infty\}$ and let

$$\mathcal{S}(\mathcal{N}(\Lambda)) := \{f : \mathcal{N}(\Lambda) \to \mathbb{R} : |f(x)| \leq K|x|^k + M \text{ for some } K, M, k \geq 0\} \tag{2.1}$$

denote the space of real functions on $\mathcal{N}(\Lambda)$ satisfying a polynomial growth condition. Recall the definition of the operator G from (1.2). Generalizing [1, Prop. 8], we have the following result. Below and in what follows, we let $\mathbb{P}^x$ denote the law of the (q, a, b, c, d)-branco-process started in x and we let $\mathbb{E}^x$ denote expectation with respect to $\mathbb{P}^x$.

Proposition 2.1. (Finite branco-processes) *Let X be the (q,a,b,c,d)-branco-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*

$$\mathbb{E}^x\big[|X_t|^{\langle k\rangle}\big] \le |x|^{\langle k\rangle} e^{kbt} \qquad (k=1,2,\ldots,\ t\ge 0). \tag{2.2}$$

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. The proof of [1, Prop. 8] carries over without a change. □

We equip $\mathbb{N}^\Lambda$ with the componentwise order, i.e., for two states $x,\tilde{x}\in\mathbb{N}^\Lambda$, we write $x\le\tilde{x}$ if $x(i)\le\tilde{x}(i)$ for all $i\in\Lambda$. In [1], we made extensive use of monotonicity of branching-coalescing particle systems. For systems with annihilation, most of these arguments do no longer work. In fact, we can only prove the following fact.

Lemma 2.2. (Comparison of branco-processes) *Let X and $\tilde{X}$ be the (q,a,b,c,d)-branco-process and the $(q,0,\tilde{b},\tilde{c},\tilde{d})$-branco-process started in finite initial states x and $\tilde{x}$, respectively. Assume that*

$$x\le\tilde{x},\quad b\le\tilde{b},\quad a+c\ge\tilde{c},\quad d\ge\tilde{d}. \tag{2.3}$$

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

$$X_t\le\tilde{X}_t \qquad (t\ge 0). \tag{2.4}$$

Proof. This can be proved in the same way as [1, Lemma 9], by constructing 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, with dynamics as described there, except that each pair of black particles, present at the same site, is replaced with rate $2(1-\theta)c$ by one black and one white particle, with rate $2(1-\theta)a$ by two white particles, with rate $2\theta c$ by one black particle, and with rate $2\theta a$ by one white particle, where $\theta := \tilde{c}/(a+c)$. □

We will often need to compare two (q,a,b,c,d)-branco-processes with the same parameters but different initial states. A convenient way to do this is to use coupling. Let (Y^{01},Y^{11},Y^{10}) be a trivariate process, in which particles jump, die and give birth to particles of their own type, and pairs of particles of the same type annihilate and coalesce in the usual way of a (q,a,b,c,d)-branco-processes, and in addition, pairs of particles of different types coalesce to one new particle with a type that depends on its parents, according to the following rates:

$$\begin{aligned} 01+10 &\mapsto \quad 11 \quad \text{at rate } r,\\ 01+11 &\mapsto \begin{cases} 10 & \text{at rate } 2a,\\ 11 & \text{at rate } 2c,\end{cases}\end{aligned} \tag{2.5}$$

and similarly $10+11\mapsto 01$ or 11 at rate $2a$ resp. $2c$. Then it is easy to see that, for any choice of the parameter $r\ge 0$, both $X := Y^{01}+Y^{11}$ and $X' := Y^{10}+Y^{11}$ are (q,a,b,c,d)-branco-processes. We will call this the *standard coupling* with parameter r. Note that if $a=0$, then $X_0\le X'_0$ implies $X_t\le X'_t$ for all $t\ge 0$ but the same conclusion cannot be drawn if $a>0$ because of the transition $01+11\mapsto 10$.

Let X be the (q,a,b,c,d)-branco-process. It follows from Proposition 2.1 that the semigroup $(S_t)_{t\ge 0}$ defined by

$$S_tf(x) := \mathbb{E}^x[f(X_t)] \qquad (t\ge 0,\ x\in\mathcal{N}(\Lambda),\ f\in\mathcal{S}(\mathcal{N}(\Lambda))) \tag{2.6}$$

maps $\mathcal{S}(\mathcal{N}(\Lambda))$ into itself. The semigroup gives first moments of functions of our process. We will also need a covariance formula for functions of our process, the general form of which is well-known. Below, for any measure μ and function f, we write $\mu f := \int f \mathrm{d}\mu$ whenever the integral is well-defined, and we let $\mathrm{Cov}_\mu(f,g) := \mu(fg) - (\mu f)(\mu g)$ denote the covariance of functions f, g under μ. Note that if μ is a probability measure on $\mathcal{N}(\Lambda)$, then $\mu S_t f = \int \mu(\mathrm{d}x)\mathbb{E}^x(f(X_t))$, i.e., μS_t is the law at time t of the (q,a,b,c,d)-branco-processes started in the initial law μ.

Lemma 2.3. **(Covariance formula)** *Let $(S_t)_{t\geq 0}$ be the semigroup defined in (2.6) and let μ be a probability measure on $\mathcal{N}(\Lambda)$ such that $\int \mu(\mathrm{d}x)|x|^k < \infty$ for all $k \geq 1$. Then, for each $t \geq 0$ and $f, g \in \mathcal{S}(\mathcal{N}(\Lambda))$, one has*

$$\mathrm{Cov}_{\mu S_t}(f,g) = \mathrm{Cov}_\mu(S_t f, S_t g) + 2\int_0^t \mu S_{t-s}\Gamma(S_s f, S_s g)\mathrm{d}s, \tag{2.7}$$

where $\Gamma(f,g) := \frac{1}{2}\big(G(fg) - (Gf)g - f(Gg)\big)$ is given by

$$\begin{aligned}
2\Gamma(f,g)(x) = &\sum_{ij} q(i,j)x(i)\big(f(x+\delta_j-\delta_i) - f(x)\big)\big(g(x+\delta_j-\delta_i) - g(x)\big)\\
&+a\sum_i x(i)(x(i)-1)\big(f(x-2\delta_i) - f(x)\big)\big(g(x-2\delta_i) - g(x)\big)\\
&+b\sum_i x(i)\big(f(x+\delta_i) - f(x)\big)\big(g(x+\delta_i) - g(x)\big)\\
&+c\sum_i x(i)(x(i)-1)\big(f(x-\delta_i) - f(x)\big)\big(g(x-\delta_i) - g(x)\big)\\
&+d\sum_i x(i)\big(f(x-\delta_i) - f(x)\big)\big(g(x-\delta_i) - g(x)\big).
\end{aligned} \tag{2.8}$$

Proof. Formula (2.7) is standard, but the details of the proof vary depending on the Markov process under consideration. In the present case, we can copy the proof of [31, Prop. 2.2] almost without a change. We start by noting that $fg \in \mathcal{S}(\mathcal{N}(\Lambda))$ for all $f, g \in \mathcal{S}(\mathcal{N}(\Lambda))$, hence $\Gamma(f,g) := \frac{1}{2}\big(G(fg) - (Gf)g - f(Gg)\big)$ is well-defined for all $f, g \in \mathcal{S}(\mathcal{N}(\Lambda))$. It is a straightforward excercise to check that $\Gamma(f,g)$ is given by (2.8). Now (2.7) will follow from a standard argument (such as given in [31, Prop. 2.2]) provided we show that

$$\begin{aligned}
\tfrac{\partial}{\partial s} S_s\big((S_t f)(S_u g)\big) &= S_s G\big((S_t f)(S_u g)\big),\\
\tfrac{\partial}{\partial t} S_s\big((S_t f)(S_u g)\big) &= S_s\big((G S_t f)(S_u g)\big),\\
\tfrac{\partial}{\partial u} S_s\big((S_t f)(S_u g)\big) &= S_s\big((S_t f)(G S_u g)\big)
\end{aligned} \tag{2.9}$$

for all $0 \leq s, t, u$ and $f, g \in \mathcal{S}(\mathcal{N}(\Lambda))$. Let us say that a sequence of functions $f_n \in \mathcal{S}(\mathcal{N}(\Lambda))$ converges 'nicely' to a limit $f \in \mathcal{S}(\mathcal{N}(\Lambda))$ if $f_n \to f$ pointwise and there exist constants $K, M, k \geq 0$ such that $\sup_n |f_n(x)| \leq K|x|^k + M$. Then (2.2) and dominated convergence show that $f_n \to f$ 'nicely' implies $S_t f_n \to S_t f$ 'nicely'. Note also that if $f_n, f, g \in \mathcal{S}(\mathcal{N}(\Lambda))$ and $f_n \to f$ 'nicely', then $f_n g \to fg$ 'nicely'. It is easy to check that $Gf \in \mathcal{S}(\mathcal{N}(\Lambda))$ for all $f \in \mathcal{S}(\mathcal{N}(\Lambda))$. Since the (q,a,b,c,d)-branco-process X^x started in a deterministic initial state $X_0^x = x \in \mathcal{N}(\Lambda)$ solves the martingale problem for G, we have

$$t^{-1}\big(S_t f(x) - f(x)\big) = t^{-1}\int_0^t \mathbb{E}\big[Gf(X_s^x)\big]\mathrm{d}s \underset{t\downarrow 0}{\longrightarrow} Gf(x) \qquad \big(x \in \mathcal{N}(\Lambda)\big), \tag{2.10}$$

which by (2.2) and the fact that $Gf \in \mathcal{S}(\mathcal{N}(\Lambda))$ implies that $t^{-1}(S_t f - f) \to Gf$ 'nicely' as $t \downarrow 0$. Combining three facts, we see that

$$\begin{aligned}
\tfrac{\partial}{\partial s} S_s\big((S_t f)(S_u g)\big) &= \lim_{\varepsilon\downarrow 0} S_s(P_\varepsilon - 1)\big((S_t f)(S_u g)\big) = S_s G\big((S_t f)(S_u g)\big),\\
\tfrac{\partial}{\partial t} S_s\big((S_t f)(S_u g)\big) &= \lim_{\varepsilon\downarrow 0} S_s\big(((P_\varepsilon - 1)S_t f)(S_u g)\big) = S_s\big((G S_t f)(S_u g)\big),
\end{aligned} \tag{2.11}$$

and similarly for the derivative w.r.t. u, where we are using that if the right-hand derivative of a continuous real function exists in each point and depends continuously on t, then the function is continuously differentiable (see, e.g., [16, Excersise 17.24]). $\square$

2.3 Infinite systems

Recall the definition of the Liggett-Spitzer space $\mathcal{E}_\gamma(\Lambda)$ from (1.3). We let $\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ denote the class of Lipschitz functions on $\mathcal{E}_\gamma(\Lambda)$, i.e., $f : \mathcal{E}_\gamma(\Lambda) \to \mathbb{R}$ such that $|f(x) - f(y)| \leq L\|x - y\|_\gamma$ for some $L < \infty$.

The main result of this section is the following generalization of [1, Prop. 11].

Proposition 2.4. (Construction of branco-processes) *Let $(S_t)_{t\geq 0}$ be the semigroup defined in (2.6). For each $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ and $t \geq 0$, the function $S_t f$ defined in (2.6) 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*

$$\mathbb{E}^x[f(X_t)] = S_t f(x) \qquad (f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda)),\ x \in \mathcal{E}_\gamma(\Lambda),\ t \geq 0). \tag{2.12}$$

To prepare for the proof of Proposition 2.4, we start with the following lemma, which generalizes [1, Lemma 12].

Lemma 2.5. (Action of the semigroup on Lipschitz functions) *Let $(S_t)_{t\geq 0}$ be the semigroup of the (q, a, b, c, d)-branco-process, defined in (2.6). If $f : \mathcal{N}(\Lambda) \to \mathbb{R}$ is Lipschitz continuous in the norm $\|\cdot\|_\gamma$ from (1.4), with Lipschitz constant L, then*

$$|S_t f(x) - S_t f(x')| \leq L e^{(K+b-d)t}\|x - x'\|_\gamma \qquad (x, x' \in \mathcal{N}(\Lambda),\ t \geq 0), \tag{2.13}$$

where K is the constant from (1.4).

Proof. Let $X = Y^{01} + Y^{11}$ and $X' = Y^{10} + Y^{11}$ be (q, a, b, c, d)-branco-processes started in $X_0 = x$ and $X'_0 = x'$, coupled using the standard coupling from (2.5), in such a way that $(Y^{01}_0, Y^{11}_0, Y^{10}_0) = ((x - x')_+, x \wedge x', (x' - x)_+)$. Then

$$\begin{aligned} |S_t f(x) - S_t f(x')| &= \big|\mathbb{E}[f(X_t)] - \mathbb{E}[f(X'_t)]\big| \leq \mathbb{E}\big[|f(X_t) - f(X'_t)|\big] \\ &\leq L\mathbb{E}\big[\|X_t - X'_t\|_\gamma\big] = L\mathbb{E}\big[\|Y^{01}_t + Y^{10}_t\|_\gamma\big]. \end{aligned} \tag{2.14}$$

Let us choose the parameter r in the standard coupling as $r := 2(a + c)$. Then it is easy to see that (Y^{01}, Y^{10}) can be coupled to a $(q, 0, b, a + c, d)$-branco-process Z started in $Z_0 = |x - x'|$ in such a way that $Y^{01}_t + Y^{10}_t \leq Z_t$ for all $t \geq 0$. Therefore, by [1, formula (3.13)], we can further estimate the quantity in (2.14) as

$$|S_t f(x) - S_t f(x')| \leq L\mathbb{E}\big[\|Z_t\|_\gamma\big] \leq L e^{(K+b-d)t}\|x - x'\|_\gamma. \tag{2.15}$$

$\square$

Proof of Proposition 2.4. Since $\mathcal{N}(\Lambda)$ is a dense subset of $\mathcal{E}_\gamma(\Lambda)$, Lemma 2.5 implies that for each $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$ and $t \geq 0$, the function $S_t f$ defined in (2.6) can be extended to a unique Lipschitz function on $\mathcal{E}_\gamma(\Lambda)$. The proof of Lemma 2.5 moreover shows that two (q, a, b, c, d)-branco-processes X, X' started in finite initial states x, x' can be coupled such that

$$\mathbb{E}\big[\|X_t - X'_t\|_\gamma\big] \leq e^{(K+b-d)t}\|x - x'\|_\gamma \qquad (t \geq 0). \tag{2.16}$$

It is not hard to see that for each $x \in \mathcal{E}_\gamma(\Lambda)$ we can choose $x_n \in \mathcal{N}(\Lambda)$ such that $\|x_n - x\| \to 0$ and

$$\sum_{n=1}^{\infty} \|x_n - x_{n-1}\|_\gamma < \infty. \tag{2.17}$$

(For example, any $x_n \uparrow x$ has these properties.) Let X^n be the process started in $X^n_0 = x_n$. By (2.16), we can inductively couple the processes $X^0, X^1, X^2, \ldots$ in such a way that

$$\mathbb{E}\big[\|X^n_t - X^{n-1}_t\|_\gamma\big] \leq e^{(K+b-d)t}\|x_n - x_{n-1}\|_\gamma \qquad (n \geq 1,\ t \geq 0). \tag{2.18}$$

It follows that for each (deterministic) $t \geq 0$, the sequence $(X^n_t)^{n\geq 0}$ is a.s. a Cauchy sequence in the complete metric space $\mathcal{E}_\gamma(\Lambda)$, hence for each $t \geq 0$ there a.s. exists an $\mathcal{E}_\gamma(\Lambda)$-valued random variable X_t such that $\|X^n_t - X_t\|_\gamma \to 0$. By Fatou,

$$\mathbb{E}\big[\|X^n_t - X_t\|_\gamma\big] \leq \liminf_{m\to\infty} \mathbb{E}\big[\|X^n_t - X^m_t\|_\gamma\big] \leq e^{(K+b-d)t} \sum_{m=n}^{\infty} \|x_{m+1} - x_m\|_\gamma \underset{n\to\infty}{\longrightarrow} 0. \tag{2.19}$$

Just as in [1, Lemma 13], it is now straightforward to check that $(X_t)_{t\geq 0}$ is a Markov process with semigroup $(S_t)_{t\geq 0}$. Note, however, that in the arguments so far we have only constructed $X = (X_t)_{t\geq 0}$ at deterministic times. To show that X has a version with cadlag sample paths (where the limits from the left and right are defined w.r.t. the norm $\|\cdot\|_\gamma$), we adapt arguments from the proof of [1, Prop. 11]. It suffices to prove X has cadlag sample paths on the time interval $[0, 1]$.

Let V be the process with generator

$$G_V f(x) := \sum_{ij} q(i,j)x(i)\{f(x+\delta_j) - f(x)\} + b\sum_i x(i)\{f(x+\delta_i) - f(x)\}, \tag{2.20}$$

which describes a branching process in which particles don't move or die, and each particle at i gives with rate $q(i,j)$ birth to a particle at j and with rate b to a particle at i. We claim that a (q,a,b,c,d)-branco-process X, started in a finite initial state $X_0 = x$, can be coupled to the process V started in $V_0 = x$ in such a way that $X_t \leq V_t$ for all $t \geq 0$. To see this, let (B, W) be a bivariate process, say of black and white particles, started in $(B_0, W_0) = (x, 0)$, such that the black particles evolve as a (q,a,b,c,d)-branco-process, the white particles evolve according to the generator in (2.20), and each time a black particle disappears from a site i due to jumps, annihilation or coalescence, a white particle is created at i. Then it is easy to see that $X = B$ and $V = B + W$. By [1, formula (3.25)],

$$\mathbb{E}\big[\|V_t\|_\gamma\big] \leq e^{(K+b)t}\|x\|_\gamma, \tag{2.21}$$

where K is the constant from (1.4). Since V is nondecreasing in t, since $V_t(i)$ increases by one each time $X_t(i)$ does, and since X cannot become negative, it follows that

$$\big|\big\{t \in [0,1] : X_{t-}(i) \neq X_t(i)\big\}\big| \leq x(i) + 2V_1(i). \tag{2.22}$$

Applying this to the process X^n, multiplying with γ_i and summing over i, we see that

$$\sum_i \gamma_i \mathbb{E}\big[\big|\{t \in [0,1] : X^n_{t-}(i) \neq X^n_t(i)\}\big|\big] \leq (1 + 2e^{K+b})\|x_n\|_\gamma, \tag{2.23}$$

which by the convergence of $\|x_n\|_\gamma$ gives us a uniform bound on the number of jumps made by X^n.

We wish to show that for large n, the processes X^n and X^{n+1} make mostly the same jumps. To this aim, let $X^n = Y^{01} + Y^{11}$ and $X^{n+1} = Y^{10} + Y^{11}$ be two (q,a,b,c,d)-branco-processes, coupled using the standard coupling from (2.5), with $r = 2(a+c)$ and $(Y^{01}_0, Y^{11}_0, Y^{10}_0) = ((x_n - x_{n+1})_+, x_n \wedge x_{n+1}, (x_{n+1} - x)_+)$. Then, just as in the proof of Lemma 2.5, the process (Y^{01}, Y^{10}) can be coupled to a $(q, 0, b, a+c, d)$-branco-process Z started in $Z_0 = |x_n - x_{n+1}|$ in such a way that $Y^{01}_t + Y^{10}_t \leq Z_t$ for all $t \geq 0$. Likewise, it is not hard to see that we can couple (Y^{01}, Y^{10}) to a process V with dynamics as in (2.20) started in $V_0 = |x_n - x_{n+1}|$, in such a way that $Y^{01}_t + Y^{10}_t \leq Z_t$ for all $t \geq 0$ and moreover, whenever $Y^{01}(i)$ or $Y^{01}(i)$ increases, the process $V(i)$ increases by the same amount. Let

$$J_n(i) := \big\{t \in [0,1] : X^n_{t-}(i) \neq X^n_t(i)\big\} \tag{2.24}$$

be the set of jump times up to time one of the process $X^n(i)$ and let

$$I(i) := \big\{t \in [0,1] : Y^{01}_{t-}(i) \neq Y^{01}_t(i)\big\} \cup \big\{t \in [0,1] : Y^{10}_{t-}(i) \neq Y^{10}_t(i)\big\}. \tag{2.25}$$

Then the symmetric difference $J_n(i) \vartriangle J_{n+1}(i) = (J_n(i)\backslash J_{n+1}(i)) \cup (J_{n+1}(i)\backslash J_n(i))$ of $J_n(i)$ and $J_{n+1}(i)$ is contained in $I(i)$ and, by the arguments leading up to (2.22), $|I(i)| \leq |x_n(i) - x_{n+1}(i)| + 2V_1(i)$. Thus, in analogy with (2.23), we find that

$$\sum_i \gamma_i \mathbb{E}\big[|J_n(i) \vartriangle J_{n+1}(i)|\big] \leq (1 + 2e^{K+b})\|x_n - x_{n+1}\|_\gamma. \tag{2.26}$$

By (2.17), it follows that the sets $J_n(i)$ converge as $n \to \infty$, i.e., for each $i \in \Lambda$ there is a (random) n such that $J_n(i) = J_{n+1}(i) = J_{n+2}(i) = \cdots$. Taking into account also (2.22), it follows that the limit process $(X(i))_{t\geq 0}$ has cadlag sample paths for each $i \in \Lambda$ and the set of jump times of $X^n(i)$ converges to the set of jump times of $X(i)$. The fact that the sample path of $(X)_{t\geq 0}$ are also cadlag in the norm $\|\cdot\|_\gamma$ can be proved in the same way as [1, formula (3.31)]. □

The proof of Proposition 2.4 yields a useful side result.

Corollary 2.6. (Approximation with finite systems) *Let $x \in \mathcal{E}_\gamma(\Lambda)$ and $x_n \in \mathcal{N}(\Lambda)$ satisfy $\|x_n - x\|_\gamma \to 0$ and $\sum_{n\geq 1} \|x_n - x_{n-1}\|_\gamma < \infty$. Then the (q,a,b,c,d)-branco-processes X^n, X started in $X^n_0 = x_n$ and $X_0 = x$ can be coupled in such a way that $\|X^n_t - X_t\|_\gamma \to 0$ a.s. for each $t \geq 0$.*

2.4 Covariance estimates

In this section, we give an upper estimate on the covariance of two functions of a (q,a,b,c,d)-branco-process, which shows in particular that events that are sufficiently far apart are almost independent.

For any continuous $f : \mathcal{E}_\gamma(\Lambda) \to \mathbb{R}$, we define $\delta f : \Lambda \to [0,\infty]$ by

$$\delta f(i) := \sup_{x \in \mathcal{E}_\gamma(\Lambda)} \big|f(x + \delta_i) - f(x)\big| \qquad (i \in \Lambda). \tag{2.27}$$

It is easy to see that for each continuous $f : \mathcal{E}_\gamma(\Lambda) \to \mathbb{R}$,

$$\big|f(x) - f(y)\big| \leq \sum_i \delta f(i)|x(i) - y(i)| \qquad \big(x, y \in \mathcal{E}_\gamma(\Lambda)\big). \tag{2.28}$$

Lemma 2.7. (Lipschitz functions) *A continuous function $f : \mathcal{E}_\gamma(\Lambda) \to \mathbb{R}$ is Lipschitz with respect to the norm $\|\cdot\|_\gamma$ if and only if there exists a constant $L < \infty$ such that $\delta f(i) \leq L\gamma_i$ $(i \in \Lambda)$.*

Proof. If $f \in \mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$, we have $|f(x+\delta_i) - f(x)| \leq L\|(x+\delta_i) - x\|_\gamma = L\gamma_i$, where L is the Lipschitz constant of f, hence $\delta f(i) \leq L\gamma_i$ $(i \in \Lambda)$. Conversely, if the latter condition holds, then by (2.28)

$$\big|f(x) - f(y)\big| \leq L\sum_i \gamma_i|x(i) - y(i)| = L\|x - y\|_\gamma \qquad \big(x, y \in \mathcal{E}_\gamma(\Lambda)\big). \tag{2.29}$$

□

We let $B_\gamma(\Lambda)$ denote the space of all functions $\phi:\Lambda\to\mathbb{R}$ such that

$$\sup_i \gamma_i^{-1}|\phi(i)|<\infty. \tag{2.30}$$

Note that by Lemma 2.7, $\delta f\in B_\gamma(\Lambda)$ for each $f\in\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))$.

Let $P_t(i,j)$ denote the probability that the random walk on Λ that jumps from k to l with rate $q(k,l)$, started in i, is a time t located at the position j. For any $\phi\in B_\gamma(\Lambda)$, we write

$$P_t\phi(i):=\sum_j P_t(i,j)\phi(j)\qquad(t\geq 0,\ i\in\Lambda). \tag{2.31}$$

It is not hard to check that P_t is well-defined on $B_\gamma(\Lambda)$ and maps this space into itself.

Recall that $(S_t)_{t\geq 0}$ denotes the semigroup of the (q,a,b,c,d)-branco-process, defined in (2.6).

Lemma 2.8. (Variation estimate) *For any (q,a,b,c,d)-branco-process, one has*

$$\delta S_t f\leq e^{(b-d)t}P_t\delta f\qquad\big(t\geq 0,\ f\in\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))\big). \tag{2.32}$$

Proof. Fix $i\in\Lambda$ and let $X=Y^{01}+Y^{11}$ and $X'=Y^{10}+Y^{11}$ be (q,a,b,c,d)-branco-processes started in $X_0=x$ and $X'_0=x+\delta_i$, coupled using the standard coupling from (2.5), in such a way that $(Y^{01}_0,Y^{11}_0,Y^{10}_0)=(0,x,\delta_i)$. Then

$$\begin{aligned}
&|S_tf(x)-S_tf(x+\delta_i)|=\big|\mathbb{E}[f(X_t)]-\mathbb{E}[f(X'_t)]\big|\\
&\quad\leq\mathbb{E}\big[|f(X_t)-f(X'_t)|\big]\leq\mathbb{E}\big[\sum_j\delta f(j)|X_t(j)-X'_t(j)|\big]\\
&\quad=\sum_j\delta f(j)\mathbb{E}\big[Y^{01}_t(j)+Y^{10}_t(j)\big]\leq\sum_j\delta f(j)e^{(b-d)t}P_t(i,j),
\end{aligned} \tag{2.33}$$

where in the last step we have used that $Y^{01}+Y^{10}$ can be estimated from above by a $(q,0,b,0,d)$-branco-process. ☐

Proposition 2.9. (Covariance estimate) *Let $X=(X_t)_{t\geq 0}$ be a (q,a,b,c,d)-branco-processes started in $X_0=x\in\mathcal{E}_\gamma(\Lambda)$. Then, for each $t\geq 0$, there exist functions $K_t:\Lambda\times\Lambda^2\to[0,\infty)$ and $L_t:\Lambda^2\times\Lambda^2\to[0,\infty)$ satisfying*

$$\left.\begin{array}{l}K_t(gi;gk,gl)=K_t(i;k,l)\\ L_t(gi,gj;gk,gl)=L_t(i,j;k,l)\end{array}\right\}\quad\big(i,j,k,l\in\Lambda,\ g\in{\rm Aut}(\Lambda,q)\big), \tag{2.34}$$

and $$\sup_{t\in[0,T]}\sum_{i,k}K_t(i;k,0)<\infty,\quad\textit{and}\quad\sup_{t\in[0,T]}\sum_{i,j,k}L_t(i,j;k,0)<\infty\quad(T<\infty), \tag{2.35}$$

such that

$$\begin{aligned}\big|{\rm Cov}_x\big(f(X_t),g(X_t)\big)\big|\leq&\sum_{i,k,l}x(i)K_t(i;k,l)\delta f(k)\delta g(l)\\
&+\sum_{i,j,k.l}x(i)x(j)L_t(i,j;k,l)\delta f(k)\delta g(l).\end{aligned} \tag{2.36}$$

for all bounded functions $f,g\in\mathcal{C}_{\rm Lip}(\mathcal{E}_\gamma(\Lambda))\big)$.

Proof. It suffices to prove the claim for finite initial states $x\in\mathcal{N}(\Lambda)$. For once the proposition is proved for finite systems, for arbitrary $x\in\mathcal{E}_\gamma(\Lambda)$ we can find $\mathcal{N}(\Lambda)\ni x_n\uparrow x$. Then by Corollary 2.6, the processes X^n,X started in x_n,x can be coupled such that $\|X^n_t-X_t\|_\gamma\to 0$ for each $t\geq 0$, hence by bounded pointwise convergence, the left-hand side of (2.36) for X^n converges to the same formula for X, while the right-hand side is obviously continuous under monotone limits.

We will show that for finite systems, the estimate (2.36) holds even without the boundednes assumption on f, g. We apply Lemma 2.3. A little calculation based on (2.8) shows that

$$\begin{aligned} 2|\Gamma(f,g)(x)| \leq & \sum_{ij} q(i,j) x(i) \big(\delta f(i) + \delta f(j)\big)\big(\delta g(i) + \delta g(j)\big) \\ & + (2a+c) \sum_i x(i)(x(i)-1)\delta f(i) \delta g(i) \\ & + (b+d) \sum_i x(i) \delta f(i) \delta g(i). \end{aligned} \tag{2.37}$$

In view of Lemma 2.8, we define $\tilde{P}_t := e^{(b-d)t} P_t$. Then (2.7), (2.37) and Lemma 2.8 show that for processes started in a deterministic initial state,

$$\begin{aligned} & \big|\mathrm{Cov}_x\big(f(X_t), g(X_t)\big)\big| \\ & \leq \int_0^t \sum_{ij} q(i,j) \big(\tilde{P}_s \delta f(i) + \tilde{P}_s \delta f(j)\big)\big(\tilde{P}_s \delta g(i) + \tilde{P}_s \delta g(j)\big) \mathbb{E}^x[X_{t-s}(i)] \mathrm{d}s \\ & \quad + (2a+c) \int_0^t \sum_i \tilde{P}_s \delta f(i) \tilde{P}_s \delta g(i) \mathbb{E}^x\big[X_{t-s}(i)(X_{t-s}(i)-1)\big] \mathrm{d}s \\ & \quad + (b+d) \int_0^t \sum_i \tilde{P}_s \delta f(i) \tilde{P}_s \delta g(i) \mathbb{E}^x[X_{t-s}(i)] \mathrm{d}s, \end{aligned} \tag{2.38}$$

Let $Y = (Y_t)_{t \geq 0}$ be the $(q, 0, b, 0, d)$-branco-process started in $Y_0 = x$. By Lemma 2.2, we can couple X and Y such that $X_t \leq Y_t$ for all $t \geq 0$. We estimate

$$\begin{aligned} &\text{(i)} \qquad \mathbb{E}^x[X_t(i)] \leq \mathbb{E}^x[Y_t(i)] = \sum_j x(j) \tilde{P}_t(j,i), \\ &\text{(ii)} \quad \mathbb{E}^x[X_t(i)(X_t(i)-1)] \leq \mathbb{E}^x[Y_t(i)^2] = \mathbb{E}^x[Y_t(i)]^2 + \mathrm{Var}_x(Y_t(i)). \end{aligned} \tag{2.39}$$

To estimate $\mathrm{Var}_x(Y_t(i))$, we apply (2.38) to the process Y and $f = g := f_i$ where $f_i(x) := x(i)$. Since the annihilation and coalescence rates of Y are zero, this yields

$$\begin{aligned} & \mathrm{Var}_x(Y_t(i)) \\ & \leq \int_0^t \sum_{jk} q(j,k) \big(\tilde{P}_s \delta f_i(j) + \tilde{P}_s \delta f_i(k)\big)\big(\tilde{P}_s \delta f_i(j) + \tilde{P}_s \delta f_i(k)\big) \mathbb{E}^x[Y_{t-s}(j)] \mathrm{d}s \\ & \quad + (b+d) \int_0^t \sum_j \tilde{P}_s \delta f_i(j) \tilde{P}_s \delta f_i(j) \mathbb{E}^x[Y_{t-s}(j)] \mathrm{d}s. \end{aligned} \tag{2.40}$$

Define

$$\begin{aligned} A_t(i;k,l) &:= \sum_j q(i,j) \big(\tilde{P}_t(i,k) + \tilde{P}_t(j,k)\big)\big(\tilde{P}_t(i,l) + \tilde{P}_t(j,l)\big), \\ B_t(i;k,l) &:= \tilde{P}_t(i,k) \tilde{P}_t(i,l). \end{aligned} \tag{2.41}$$

Then (2.38) can be rewritten as

$$\begin{aligned} \big|\mathrm{Cov}_x\big(f(X_t), g(X_t)\big)\big| \leq & \int_0^t \sum_{ikl} \mathbb{E}^x[X_{t-s}(i)] \big(A_s(i;k,l) + (b+d) B_s(i;k,l)\big) \delta f(k) \delta_g(l) \mathrm{d}s \\ & + (2a+c) \int_0^t \sum_{ikl} \mathbb{E}^x\big[X_{t-s}(i)(X_{t-s}(i)-1)\big] B_s(i;k,l) \delta f(k) \delta_g(l) \mathrm{d}s, \end{aligned} \tag{2.42}$$

while (2.40) can be rewritten as

$$\mathrm{Var}_x(Y_t(i)) \leq \int_0^t \sum_j \mathbb{E}^x[Y_{t-s}(j)] \big(A_s(j;i,i) + (b+d) B_s(j;i,i)\big) \mathrm{d}s, \tag{2.43}$$

where we have used that $\delta f_i(j) = 1_{\{i=j\}}$. Setting

$$C_t(i;k,l) := A_s(i;k,l) + (b+d)B_s(i;k,l), \tag{2.44}$$

and inserting (2.39) and (2.43) into (2.42), we obtain

$$\begin{aligned}
&\big|\mathrm{Cov}_x\big(f(X_t),g(X_t)\big)\big| \\
&\le \int_0^t \sum_{ijkl} x(i)\tilde P_{t-s}(i,j)C_s(j;k,l)\delta f(k)\delta_g(l)\mathrm{d}s \\
&\quad + (2a+c)\int_0^t \sum_{ijklm} x(i)x(j)\tilde P_{t-s}(i,m)\tilde P_{t-s}(j,m)B_s(m;k,l)\delta f(k)\delta_g(l)\mathrm{d}s \\
&\quad + (2a+c)\int_0^t \mathrm{d}s \sum_{ijklm}\int_0^{t-s}\mathrm{d}u\, x(i)\tilde P_u(i,j)C_u(j;m,m)B_s(m;k,l)\delta f(k)\delta_g(l).
\end{aligned} \tag{2.45}$$

Recalling the definition of $B_t(i;j,k)$, this shows that (2.36) is satisfied with

$$\begin{aligned}
K_t(i;k,l) &:= \int_0^t \mathrm{d}s \sum_j \tilde P_{t-s}(i,j)C_s(j;k,l) \\
&\quad +(2a+c)\int_0^t \mathrm{d}s \sum_{jm}\int_0^{t-s}\mathrm{d}u\, \tilde P_u(i,j)C_u(j;m,m)\tilde P_s(m,k)\tilde P_s(m,l),
\end{aligned} \tag{2.46}$$
$$L_t(i,j;k,l) := (2a+c)\int_0^t \mathrm{d}s \sum_m \tilde P_{t-s}(i,m)\tilde P_{t-s}(j,m)\tilde P_s(m,k)\tilde P_s(m,l).$$

The invariance of K_t and L_t under automorphisms of (Λ,q) is obvious from the analogue property of $\tilde P_t$, but the summability condition (2.35) needs proof. Since $P_t(i,\cdot\,)$ is a probability distribution and since the counting measure on Λ is an invariant law for P_t by assumption (1.1) (iii),

$$\sum_j P_t(i,j) = 1 = \sum_j P_t(j,i) \qquad (t\ge 0,\ i\in\Lambda). \tag{2.47}$$

Setting $|q| := \sum_j q(i,j) = \sum_j q(j,i)$, we see that

$$\begin{aligned}
\sum_{ik} A_t(i;k,l) &= \sum_{ijk} q(i,j)\big(\tilde P_t(i,k)+\tilde P_t(j,k)\big)\big(\tilde P_t(i,l)+\tilde P_t(j,l)\big) \\
&= \sum_{ij} q(i,j)2e^{(b-d)t}\big(\tilde P_t(i,l)+\tilde P_t(j,l)\big) = 4|q|e^{2(b-d)t} \qquad (l\in\Lambda),
\end{aligned} \tag{2.48}$$

and therefore, by a similar calculation for $B_t(i;k,l)$,

$$\sum_j C_t(j;m,m) \le \sum_{jk} C_t(j;k,l) = \big(4|q|+b+d\big)e^{2(b-d)t} \qquad (l,m\in\Lambda), \tag{2.49}$$

which by (2.46) implies that

$$\begin{aligned}
\sum_{ik} K_t(i;k,l) &\le \big(4|q|+b+d\big)\int_0^t \mathrm{d}s\, e^{(b-d)(t-s)}e^{2(b-d)s} \\
&\quad +(2a+c)\big(4|q|+b+d\big)\int_0^t \mathrm{d}s\int_0^{t-s}\mathrm{d}u\, e^{(b-d)u}e^{2(b-d)u}e^{2(b-d)s} < \infty, \\
\sum_{ijk} L_t(i,j;k,l) &\le (2a+c)\int_0^t \mathrm{d}s\, e^{2(b-d)(t-s)}e^{2(b-d)s} < \infty \qquad (t\ge 0,\ l\in\Lambda).
\end{aligned} \tag{2.50}$$

□

Corollary 2.10. (Exponential functions) *Let $X = (X_t)_{t\geq 0}$ be a (q,a,b,c,d)-branco-processes started in $X_0 = x \in \mathcal{E}_\gamma(\Lambda)$, and let $\mu : \Lambda \to [0,\infty)$ satisfy $\sum_i \mu(i) < \infty$. Then*

$$\begin{aligned}&\Big|\mathbb{E}^x\big[e^{-\sum_i \mu(i)X_t(i)}\big] - \prod_i \mathbb{E}^x\big[e^{-\mu(i)X_t(i)}\big]\Big|\\ &\quad\leq \tfrac{1}{2}\sum_{\substack{i,k,l\\ k\neq l}} x(i)K_t(i;k,l)\mu(k)\mu(l) + \tfrac{1}{2}\sum_{\substack{i,j,k,l\\ k\neq l}} x(i)x(j)L_t(i,j;k,l)\mu(k)\mu(l),\end{aligned} \tag{2.51}$$

where K_t, L_t are as in Proposition 2.9.

Proof. We first prove the statement if μ is finitely supported. Assume that $\mathrm{support}(\mu) = \{k_1,\ldots,k_m\}$ and set

$$f_\alpha(x) := e^{-\mu(k_\alpha)x(k_\alpha)} \quad \text{and} \quad g_\beta(x) := \prod_{\alpha=1}^{\beta} f_\alpha(x). \tag{2.52}$$

Then

$$\begin{aligned}\mathbb{E}^x\big[g_m(X_t)\big] &= \mathbb{E}^x\big[g_{m-1}(X_t)\big]\mathbb{E}^x\big[f_m(X_t)\big] + \mathrm{Cov}_x\big(g_{m-1}(X_t), f_m(X_t)\big)\\ &= \mathbb{E}^x\big[g_{m-2}(X_t)\big]\mathbb{E}^x\big[f_{m-1}(X_t)\big] + \mathrm{Cov}_x\big(g_{m-2}(X_t), f_{m-1}(X_t)\big)\\ &\qquad\qquad + \mathrm{Cov}_x\big(g_{m-1}(X_t), f_m(X_t)\big)\\ &= \ldots\\ &= \prod_{\alpha=1}^{m}\mathbb{E}[f_\alpha(X_t)] + \sum_{\alpha=2}^{m}\mathrm{Cov}_x\big(g_{\alpha-1}(X_t), f_\alpha(X_t)\big).\end{aligned} \tag{2.53}$$

Therefore, since

$$\delta g_\alpha(k) = \begin{cases}\mu(k) & \text{if } k\in\{k_1,\ldots,k_\alpha\},\\ 0 & \text{otherwise,}\end{cases} \qquad \delta f_\alpha(k) = \begin{cases}\mu(k) & \text{if } k = k_\alpha,\\ 0 & \text{otherwise,}\end{cases} \tag{2.54}$$

Proposition 2.9 tells us that

$$\begin{aligned}&\Big|\mathbb{E}^x\big[g_m(X_t)\big] - \prod_{\alpha=1}^{m}\mathbb{E}[f_\alpha(X_t)]\Big|\\ &\leq \sum_{\alpha=2}^{m}\sum_{\beta=1}^{\alpha-1}\Big(\sum_i x(i)K_t(i;k_\beta,k_\alpha)\mu(k_\beta)\mu(k_\alpha) + \sum_{i,j} x(i)x(j)L_t(i,j;k_\beta,k_\alpha)\mu(k_\beta)\mu(k_\alpha)\Big).\end{aligned} \tag{2.55}$$

To generalize the statement to the case that $\sum_i \mu(i) < \infty$ but μ is not finitely supported, it suffices to choose finitely supported $\mu_n \uparrow \mu$ and to observe that all terms in (2.51) are continuous in μ w.r.t. increasing limits. □

2.5 Duality and subduality

Recall the definition of ϕ^x from (1.7).

Lemma 2.11. (Infinite products) *Let $0 \leq \alpha \leq 1$, $\phi \in [0,1]^\Lambda$, $x \in \mathbb{N}^\Lambda$.*

(a) *Assume that one or more of the following conditions are satisfied:*

$$\text{(i) } \alpha < 1, \quad \text{(ii) } |x| < \infty, \quad \text{(iii) } |\phi| < \infty. \tag{2.56}$$

Then $\big(1-(1+\alpha)\phi\big)^x$ is well-defined.

(b) *Assume that ϕ is supported on a finite set and $x_n \in \mathbb{N}^\Lambda$ converge pointwise to x. Then $\big(1-(1+\alpha)\phi\big)^{x_n} \to \big(1-(1+\alpha)\phi\big)^x$ as $n\to\infty$.*

(c) *Assume that $|\phi|<\infty$ and $\mathbb{N}^\Lambda \ni x_n \uparrow x$. Then $\big(1-(1+\alpha)\phi\big)^{x_n} \to \big(1-(1+\alpha)\phi\big)^x$ as $n\to\infty$.*

(d) *Assume that either $\alpha<1$ or $|\phi|<\infty$, and let $[0,1]^\Lambda \ni \phi_n \uparrow \phi$. Then $\big(1-(1+\alpha)\phi_n\big)^x \to \big(1-(1+\alpha)\phi\big)^x$ as $n\to\infty$.*

Proof. Since $\big(1-(1+\alpha)\phi\big)^x := \prod_i \big(1-(1+\alpha)\phi(i)\big)^{x(i)}$, where $-1 \le \big(1-(1+\alpha)\phi(i)\big)^{x(i)} \le 1$, the only way in which the infinite product can be ill-defined is that $\prod_i \big|1-(1+\alpha)\phi(i)\big|^{x(i)} > 0$ while $\big(1-(1+\alpha)\phi(i)\big)^{x(i)} < 0$ for infinitely many i. If $\alpha<1$, then $-1 < -\alpha \le 1-(1+\alpha)\phi(i)$, so if $\big(1-(1+\alpha)\phi(i)\big)^{x(i)} < 0$ for infinitely many i, then $\prod_i \big|1-(1+\alpha)\phi(i)\big|^{x(i)} = 0$ and the infinite product is always well-defined. If $|x|<\infty$, then $\big(1-(1+\alpha)\phi(i)\big)^{x(i)} = 1$ for all but finitely many i, hence the infinite product is certainly well-defined. If $|\phi|<\infty$, finally, then $\phi(i) > \frac{1}{2}$ for finitely many i, hence $\big(1-(1+\alpha)\phi(i)\big)^{x(i)} < 0$ for finitely many i and the infinite product is again well-defined. This completes the proof of part (a).

Part (b) is trivial since all but finitely many factors in the infinite product defining $\big(1-(1+\alpha)\phi\big)^x$ are one if ϕ is finitely supported.

To prove part (c), we split the product $\prod_i \big(1-(1+\alpha)\phi(i)\big)^{x(i)}$ in finitely many factors where $\phi(i) > \frac{1}{2}$ and the remaining factors where $\phi(i) \le \frac{1}{2}$ and hence $\big(1-(1+\alpha)\phi(i)\big) \ge 0$. Then the finite part of the product converges as in part (b) while the infinite part converges in a monotone way.

For the proof of part (d) set $I := \{i\in\Lambda : x(i)\neq 0\}$ and let I_-, I_0, I_+ be the subsets of I where $1-(1+\alpha)\phi(i) < 0$, $=0$ and >0, respectively. If $I_0 \neq \emptyset$ then it is easy to see that $\big(1-(1+\alpha)\phi_n\big)^x \to 0 = \big(1-(1+\alpha)\phi\big)^x$, so from now on we may assume that $I_0=\emptyset$. Note that $1-(1+\alpha)\phi_n(i) \ge 1-(1+\alpha)\phi(i) > 0$ for all $i\in I_+$. Therefore, if I_- is finite, as must be the case when $|\phi|<\infty$, then $\prod_{i\in I_-} \big(1-(1+\alpha)\phi_n(i)\big)^{x(i)}$ converges since I_- is finite while $\prod_{i\in I_+} \big(1-(1+\alpha)\phi_n(i)\big)^{x(i)} \downarrow \prod_{i\in I_+} \big(1-(1+\alpha)\phi(i)\big)^{x(i)}$. If I_- is infinite and $\alpha<1$, then the fact that $\big|1-(1+\alpha)\phi_n(i)\big|^{x(i)} \to \big((1+\alpha)\phi(i)-1\big)^{x(i)} \le \alpha$ for each $i\in I_-$ implies that $\big(1-(1+\alpha)\phi_n\big)^x \to 0 = \big(1-(1+\alpha)\phi\big)^x$. □

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

$$\phi \mapsto \big(\tfrac{\partial}{\partial\phi(i)}f(\phi)\big)_{i\in\Lambda} \quad\text{and}\quad \phi \mapsto \big(\tfrac{\partial^2}{\partial\phi(i)\partial\phi(j)}f(\phi)\big)_{i,j\in\Lambda} \tag{2.57}$$

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

$$\begin{aligned}\mathcal{G}f(\phi) :=& \sum_{ij} q(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi) + s\sum_i \phi(i)(1-\phi(i))\tfrac{\partial}{\partial\phi(i)}f(\phi)\\ &+r\sum_i \phi(i)(1-\phi(i))\tfrac{\partial^2}{\partial\phi(i)^2}f(\phi) - m\sum_i \phi(i)\tfrac{\partial}{\partial\phi(i)}f(\phi) \qquad (\phi\in[0,1]^\Lambda).\end{aligned} \tag{2.58}$$

One can check that for $f\in\mathcal{C}^2_{\rm sum}([0,1]^\Lambda)$, the infinite sums in (2.58) converge in the supremumnorm and the result does not depend on the summation order [28, Lemma

3.4.4]. It has been shown in [1, Section 3.4] that solutions to the SDE (1.5) solve the martingale problem for the operator $\mathcal{G}$. In view of this, we loosely refer to $\mathcal{G}$ as the *generator* of the (q,r,s,m)-resem-process.

Proof of Proposition 1.1. Since by Proposition 2.1 (resp [1, Lemma 20]), $|X_0|<\infty$ (resp. $|\mathcal{X}_0|<\infty$) implies $|X_t|<\infty$ (resp. $|\mathcal{X}_t|<\infty$) for all $t\geq 0$, by Lemma 2.11, each of the conditions (1.10) (i)–(iii) guarantees that both sides of equation (1.11) are well-defined.

It suffices to prove (1.11) for deterministic initial states, i.e., we want to prove that either $\alpha<1$, $|x|<\infty$, or $|\phi|<\infty$ imply that

$$\mathbb{E}^x\big[(1-(1+\alpha)\phi)^{X_t}\big]=\mathbb{E}^\phi\big[(1-(1+\alpha)\mathcal{X}^\dagger_t)^x\big]\qquad(t\geq 0),\tag{2.59}$$

where $\mathbb{E}^x$ and $\mathbb{E}^\phi$ denote expectation w.r.t. the law of the process X started in $X_0=x$ and the process $\mathcal{X}$ started in $\mathcal{X}_0=\phi$, respectively. We start by proving (2.59) if $|x|<\infty$. We wish to apply [1, Thm 7]. Unfortunately, the original formulation of this theorem contains an error, so we have to use the corrected version in [3, Corollary 2] (see also [2, Corollary 2]). We apply this to the duality function

$$\Psi(x,\phi):=\big(1-(1+\alpha)\phi\big)^x\qquad\big(x\in\mathcal{N}(\Lambda),\ \phi\in[0,1]^\Lambda\big).\tag{2.60}$$

Since $|\Psi(x,\phi)|\leq 1$, we obviously have $\Psi(\,\cdot\,,\phi)\in\mathcal{S}(\mathcal{N}(\Lambda))$ for each $\phi\in[0,1]^\Lambda$. Since for each $x\in\mathcal{N}(\Lambda)$, the function $\Psi(x,\,\cdot\,)$ depends only on finitely many coordinates, we moreover have $\Psi(x,\,\cdot\,)\in\mathcal{C}^2_{\rm sum}([0,1]^\Lambda)$ for each such x. Let G be the generator of the (q,a,b,c,d)-branco-process and let $\mathcal{G}^\dagger$ denote the generator of the $(q^\dagger,r,s,m)$-resem-process. In order to apply [3, Corollary 2], we need to check that

$$\Phi(x,\phi):=G\Psi(\,\cdot\,,\phi)(x)=\mathcal{G}^\dagger\Psi(x,\,\cdot\,)(\phi)\qquad\big(x\in\mathcal{N}(\Lambda),\ \phi\in[0,1]^\Lambda\big)\tag{2.61}$$

and

$$\int_0^T\mathrm{d}s\int_0^T\mathrm{d}t\,\mathbb{E}\big[\big|\Phi(X_s,\mathcal{X}_t)\big|\big]<\infty\qquad(T\geq 0).\tag{2.62}$$

To check (2.61), we calculate:

$$\begin{aligned}
&G\Psi(\,\cdot\,,\phi)(x)\\
&=\sum_{ij}q(i,j)x(i)(1-(1+\alpha)\phi)^{x-\delta_i}((1-(1+\alpha)\phi)^{\delta_j}-(1-(1+\alpha)\phi)^{\delta_i})\\
&+a\sum_i x(i)(x(i)-1)(1-(1+\alpha)\phi)^{x-2\delta_i}(1-(1-(1+\alpha)\phi)^{2\delta_i})\\
&+b\sum_i x(i)(1-(1+\alpha)\phi)^{x-\delta_i}\big((1-(1+\alpha)\phi)^{2\delta_i}-(1-(1+\alpha)\phi)^{\delta_i}\big)\\
&+c\sum_i x(i)(x(i)-1)(1-(1+\alpha)\phi)^{x-2\delta_i}\big((1-(1+\alpha)\phi)^{\delta_i}-(1-(1+\alpha)\phi)^{2\delta_i}\big)\\
&+d\sum_i x(i)(1-(1+\alpha)\phi)^{x-\delta_i}(1-(1-(1+\alpha)\phi)^{\delta_i}).
\end{aligned}\tag{2.63}$$

Since

$$\begin{aligned}
\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x&=-(1+\alpha)x(i)(1-(1+\alpha)\phi)^{x-\delta_i},\\
\tfrac{\partial^2}{\partial\phi(i)^2}(1-(1+\alpha)\phi)^x&=(1+\alpha)^2x(i)(x(i)-1)(1-(1+\alpha)\phi)^{x-2\delta_i}
\end{aligned}\tag{2.64}$$

and

$$\begin{aligned}
(1-(1+\alpha)\phi)^{\delta_i}&=1-(1+\alpha)\phi(i),\\
(1-(1+\alpha)\phi)^{2\delta_i}&=\big(1-(1+\alpha)\phi(i)\big)^2,\\
(1-(1+\alpha)\phi)^{\delta_i}-(1-(1+\alpha)\phi)^{2\delta_i}&=(1+\alpha)\phi(i)\big(1-(1+\alpha)\phi(i)\big),
\end{aligned}\tag{2.65}$$

we can rewrite the expression in (2.63) as

$$
\begin{aligned}
G\Psi(\,\cdot\,,\phi)(x) = &\sum_{ij} q(i,j)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
&+\frac{a}{(1+\alpha)^2}\big(2(1+\alpha)\phi(i)-(1+\alpha)^2\phi(i)^2\big)\tfrac{\partial^2}{\partial\phi(i)^2}(1-(1+\alpha)\phi)^x \\
&+\frac{b}{1+\alpha}(1+\alpha)\phi(i)\big(1-(1+\alpha)\phi(i)\big)\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
&+\frac{c}{(1+\alpha)^2}(1+\alpha)\phi(i)\big(1-(1+\alpha)\phi(i)\big)\tfrac{\partial^2}{\partial\phi(i)^2}(1-(1+\alpha)\phi)^x \\
&-\frac{d}{1+\alpha}(1+\alpha)\phi(i)\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x.
\end{aligned}
\tag{2.66}
$$

Reordering terms gives

$$
\begin{aligned}
G\Psi(\,\cdot\,,\phi)(x) = &\sum_{ij} q(i,j)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
&+\Big(\frac{2a+c}{1+\alpha}\phi(i)-(a+c)\phi(i)^2\Big)\tfrac{\partial^2}{\partial\phi(i)^2}(1-(1+\alpha)\phi)^x \\
&+\Big((b-d)\phi(i)-b(1+\alpha)\phi(i)^2\Big)\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
= &\sum_{ij} q^\dagger(j,i)(\phi(j)-\phi(i))\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
&+(a+c)\phi(i)(1-\phi(i))\tfrac{\partial^2}{\partial\phi(i)^2}(1-(1+\alpha)\phi)^x \\
&+(1+\alpha)b\phi(i)(1-\phi(i))\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \\
&-(\alpha b+d)\phi(i)\tfrac{\partial}{\partial\phi(i)}(1-(1+\alpha)\phi)^x \qquad = \mathcal{G}^\dagger\Psi(x,\,\cdot\,)(\phi),
\end{aligned}
\tag{2.67}
$$

where we have used (1.8), which implies in particular that

$$
\frac{2a+c}{1+\alpha} = \frac{2a+c}{1+a/(a+c)} = \frac{(2a+c)(a+c)}{(a+c)+a} = a+c. \tag{2.68}
$$

It is easy to see from (2.63) that there exists a constant K such that

$$
|\Phi(x,\phi)| \le K\big(1+|x|^2\big) \qquad \big(\phi\in[0,1]^\Lambda,\ x\in\mathcal{N}(\Lambda)\big), \tag{2.69}
$$

hence (2.62) follows from Proposition 2.1. This completes the proof of (2.59) in case $|x|<\infty$.

We next claim that (2.59) holds if $x\in\mathcal{E}_\gamma(\Lambda)$ and ϕ is supported on a finite set. Choose $\mathcal{N}(\Lambda)\ni x_n\uparrow x$ and let X^n denote the (q,a,b,c,d)-branco-process started in $X^n_0=x_n$. Then Corollary 2.6 implies that the X^n can be coupled such that $X^n_t(i)\to X_t(i)$ a.s. for each $i\in\Lambda$. Therefore, taking the limit in (2.59), using the fact that the integrands on the left- and right-hand sides converge in a bounded pointwise way by Lemma 2.11 (b) and (c), respectively, our claim follows.

To see that (2.59) holds more generally if $\alpha<1$ or $|\phi|<\infty$, we choose finitely supported $\phi_n\uparrow\phi$ and let $\mathcal{X}^n$ denote the (q,r,s,m)-resem-process started in $\mathcal{X}^n_0=\phi_n$. Then [1, Lemma 22] implies that the $\mathcal{X}^n$ can be coupled such that $\mathcal{X}^n_t(i)\uparrow\mathcal{X}_t(i)$ a.s. for each $i\in\Lambda$. The statement then follows by letting $n\to\infty$ and applying Lemma 2.11 (d). □

Fix constants $\beta\in\mathbb{R}$, $\gamma\ge 0$. 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

$$
\mathrm{d}\mathcal{Y}_t(i) = \sum_j a(j,i)(\mathcal{Y}_t(j)-\mathcal{Y}_t(i))\,\mathrm{d}t + \beta\mathcal{Y}_t(i)\,\mathrm{d}t + \sqrt{2\gamma\mathcal{Y}_t(i)}\,\mathrm{d}B_t(i) \tag{2.70}
$$

$(t \geq 0,\ i \in \Lambda)$. Then $\mathcal{Y}$ is the well-known super random walk with underlying motion a, growth parameter β and activity γ. One has [7, Section 4.2]

$$\mathbb{E}^{\phi}\big[e^{-\langle \mathcal{Y}_t, \psi\rangle}\big] = e^{-\langle \phi, \mathcal{U}_t\psi\rangle} \tag{2.71}$$

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

$$\tfrac{\partial}{\partial t}u_t(i) = \sum_j a(j,i)(u_t(j) - u_t(i)) + \beta u_t(i) - \gamma u_t(i)^2 \qquad (i \in \Lambda,\ t \geq 0) \tag{2.72}$$

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 $\mathcal{Y}$.

It has been shown in [1, Prop. 23] that the (q,a,b,c,d)-branco-process and the super random walk with underlying motion $q^{\dagger}$, growth parameter $b - d + c$ and activity c are related by a 'subduality', i.e., a duality formula with a nonnegative error term. The next proposition generalizes this to branco-processes with positive annihilation rate.

Proposition 2.12. (Subduality with a branching process) *Let X be the (q,a,b,c,d)-branco-process and let $\mathcal{Y}$ be the super random walk with underlying motion $q^{\dagger}$, growth parameter $2a + b - d + c$ and activity $2a + c$. Then*

$$\mathbb{E}^{x}\big[e^{-\langle \phi, X_t\rangle}\big] \geq \mathbb{E}^{\phi}\big[e^{-\langle \mathcal{Y}_t, x\rangle}\big] \qquad (x \in \mathcal{E}_\gamma(\Lambda),\ \phi \in [0,\infty)^{\Lambda},\ |\phi| < \infty). \tag{2.73}$$

Proof. We first prove the statement if $|x| < \infty$ and $|\phi| < \infty$. This goes exactly in the same way as in the proof of [1, Prop. 23]. Let $\mathcal{H}$ denote the generator of $\mathcal{Y}$, defined in [1, formula (4.14)], let G be the generator in (1.2), and let Ψ be the duality function $\Psi(x,\phi) := e^{-\langle \phi, x\rangle}$. Then one has

$$\begin{aligned}
&G\Psi(\cdot,\phi)(x) - \mathcal{H}\Psi(x,\cdot)(\phi) = \Big\{ \sum_{ij} q(i,j)x(i)\big(e^{\phi(i)-\phi(j)} - 1 - (\phi(i) - \phi(j))\big)\\
&\quad + a\sum_i x(i)(x(i)-1)\big(e^{2\phi(i)} - 1 - 2\phi(i)\big) + b\sum_i x(i)\big(e^{-\phi(i)} - 1 + \phi(i)\big)\\
&\quad + c\sum_i x(i)(x(i)-1)\big(e^{\phi(i)} - 1 - \phi(i)\big) + d\sum_i x(i)\big(e^{\phi(i)} - 1 - \phi(i)\big)\Big\} e^{-\langle \phi, x\rangle} \geq 0.
\end{aligned} \tag{2.74}$$

This is just [1, formula (4.19)], where the extra terms with the prefactor a obtain their $e^{2\phi(i)} - 1$ part from the generator G and the remaining $-2\phi(i)$ from $\mathcal{H}$. Using Proposition 2.1 to guarantee integrability we may apply [3, Corollary 2] to deduce (2.73).

To generalize (2.73) to $x \in \mathcal{E}_\gamma(\Lambda)$ and $\phi \in [0,\infty)^{\Lambda}$ supported on a finite set, we choose $\mathcal{N}(\Lambda) \ni x_n \uparrow x$ and let X^n denote the (q,a,b,c,d)-branco-process started in $X^n_0 = x_n$. Then Corollary 2.6 implies that the X^n can be coupled such that $X^n_t(i) \to X_t(i)$ a.s. for each $i \in \Lambda$. It follows that $e^{-\langle \phi, X^n_t\rangle} \to e^{-\langle \phi, X_t\rangle}$ a.s. and $e^{-\langle \mathcal{Y}_t, x_n\rangle} \downarrow e^{-\langle \mathcal{Y}_t, x_n\rangle}$ a.s., so taking the limit in (2.73) we obtain the statement for $x \in \mathcal{E}_\gamma(\Lambda)$ and ϕ finitely supported. To generalize this to $|\phi| < \infty$ we choose $\phi_n \uparrow \phi$ and let $\mathcal{Y}^n$ denote the super random walk started in $\mathcal{Y}^n_0 = \phi_n$. Then it is well-known (and can be proved in the same way as [1, Lemma 22]) that the $\mathcal{Y}^n$ can be coupled in such a way that $\mathcal{Y}^n_t \uparrow \mathcal{Y}_t$ for each $t \geq 0$. Therefore, taking the monotone limit in (2.73) our claim follows. □

2.6 The process started at infinity

In view of what follows, we recall the following projective limit theorem. Let E and $(E_i)_{i\in\mathbb{N}}$ be Polish spaces. Assume that $\pi_i : E \to E_i$ are continuous surjective maps that separate points, i.e., for all $x, y \in E$ with $x \neq y$, there exists an $i \in \mathbb{N}$ with $\pi_i(x) \neq \pi_i(y)$.

For each $i \leq j$, let $\pi_{ij} : E_j \to E_i$ be continuous maps satisfying $\pi_{ij} \circ \pi_j = \pi_i$. Assume moreover that for each sequence $(x_i)_{i\in\mathbb{N}}$ with $x_i \in E_i$ $(i \in \mathbb{N})$ that is consistent in the sense that $\pi_{ij}(x_j) = x_i$ $(i \leq j)$, there exists an $y \in E$ such that $\pi_i(y) = x_i$ $(i \in \mathbb{N})$. Let $(\mu_i)_{i\in\mathbb{N}}$ be probability measures on the E_i's, respectively (equipped with their Borel-σ-fields), that are consistent in the sense that $\mu_i = \mu_j \circ \pi_{ij}^{-1}$ for all $i \leq j$. Then there exists a unique probability measure μ on E such that $\mu \circ \pi_i^{-1} = \mu_i$ for all $i \in \mathbb{N}$.

This may be proved by invoking Kolmogorov's extension theorem to construct a probability measure μ' on the product space $\prod_i E_i$ whose marginals are the μ_i and that is moreover concentrated on the set $E' \subset \prod_i E_i$ consisting of all $(x_i)_{i\in\mathbb{N}}$ satisfying $\pi_{ij}(x_j) = x_i$ for all $i \leq j$. Now $\vec{\pi}(y) := (\pi_i(y))_{i\in\mathbb{N}}$ defines a bijection $\vec{\pi} : E \to E'$, so there exists a unique measure μ on the σ-algebra generated by the $(\pi_i(x))_{i\in\mathbb{N}}$ whose image under $\vec{\pi}$ equals μ'. By [21, Lemma II.18], this σ-algebra coincides with the Borel-σ-algebra on E.

Proof of Theorem 1.4. In the case without annihilation, parts (a)–(e) were proved in [1, Thm 2]. The proof there made essential use of monotonicity, which is not available in case $a > 0$. Instead of trying to adapt these arguments, replacing monotone convergence by some other form of convergence wherever necessary, we will make use of Corollary 1.3, which will simplify our life considerably.

In view of this, set $\alpha := a/(a + c)$ and let $\overline{X}^{(\infty)}$ be the $(q, 0, (1 + \alpha)b, a + c, \alpha b + d)$-branco-process started at infinity, as defined in [1, Thm 2]. Fix $\varepsilon > 0$ and let $(X^\varepsilon_t)_{t\geq\varepsilon}$ be a (q, a, b, c, d)-branco-process started at time ε in $X^\varepsilon_\varepsilon = \mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}^{(\infty)}_\varepsilon)$. It has been proved in [1, Thm 2] that $\overline{X}^{(\infty)}_t \in \mathcal{E}_\gamma(\Lambda)$ for all $t \geq 0$ a.s., hence $\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}^{(\infty)}_\varepsilon) \in \mathcal{E}_\gamma(\Lambda)$ and $(X^\varepsilon_t)_{t\geq\varepsilon}$ is well-defined by Proposition 2.4. By Corollary 1.3,

$$\mathbb{P}[X^\varepsilon_t \in \cdot\,] = \mathbb{P}\big[\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}^{(\infty)}_t) \in \cdot\,\big] \qquad (t \geq \varepsilon). \tag{2.75}$$

In particular, this implies that if we construct two processes $X^\varepsilon, X^{\varepsilon'}$ for two values $0 < \varepsilon < \varepsilon'$, then these are consistent in the sense that $(X^\varepsilon_t)_{t\geq\varepsilon'}$ is equally distributed with $(X^{\varepsilon'}_t)_{t\geq\varepsilon'}$. By applying the projective limit theorem sketched above, using the spaces of componentwise cadlag functions from (ε, ∞) to $\mathbb{N}^\Lambda$, we may construct a process $(X^{(\infty)}_t)_{t>0}$ such that $X^{(\infty)}_\varepsilon$ is equally distributed with $\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}^{(\infty)}_\varepsilon)$ for all $\varepsilon > 0$ and $(X^{(\infty)}_t)_{t>0}$ evolves as a (q, a, b, c, d)-branco-process. Let $\mathcal{X}^\dagger$ denote the (q, r, s, m)-resem process with r, s, m as in (1.8). Then $\mathcal{X}^\dagger$ is dual to both the (q, a, b, c, d)-branco-process (with parameter $\alpha = a/(a+c)$ in the duality function) and to the $(q, 0, (1+\alpha)b, a+c, \alpha b+d)$-branco-process (with duality function $\Psi(x, \phi) = (1 - \phi)^x$). We have

$$\begin{aligned}&\mathbb{E}\big[(1 - (1 + \alpha)\phi)^{X^{(\infty)}_t}\big] = \mathbb{E}\big[(1 - (1 + \alpha)\phi)^{\mathrm{Thin}_{\frac{1}{1+\alpha}}(\overline{X}^{(\infty)}_t)}\big] = \mathbb{E}\big[(1 - \phi)^{\overline{X}^{(\infty)}_t}\big]\\ &\quad= \mathbb{P}^\phi\big[\mathcal{X}^\dagger_t = 0\big] \qquad (t \geq 0,\ \phi \in [0, 1]^\Lambda,\ |\phi| < \infty),\end{aligned} \tag{2.76}$$

where the last equality follows from [1, formula (5.5)] and we assume $|\phi| < \infty$ to make sure the infinite products are well-defined. It has been shown in [1, Thm 2 (d)] that the law of $\overline{X}^{(\infty)}_t$ converges as $t \to \infty$ to an invariant law of the $(q, 0, (1 + \alpha)b, a + c, \alpha b + d)$-branco-process. By thinning, it follows that the law of $X^{(\infty)}_t$ converges as $t \to \infty$ to an invariant law $\overline{\nu}$ of the (q, a, b, c, d)-branco-process. Taking the limit $t \to \infty$ in (2.76) we arrive at (1.22). Setting

$$r := (1 + \alpha)b + (a + c) - (\alpha b + d) = a + b + c - d, \tag{2.77}$$

we obtain from [1, Thm 2 (b)] and the fact that $X^{(\infty)}$ is a $1/(1+\alpha)$-thinning of $\overline{X}^{(\infty)}$, that

$$\mathbb{E}[X_t^{(\infty)}(i)] \leq \begin{cases} \frac{1}{1+\alpha}\frac{r}{(a+c)(1-e^{-rt})} & \text{if } r \neq 0, \\ \frac{1}{1+\alpha}\frac{1}{(a+c)t} & \text{if } r = 0 \end{cases} \qquad (i \in \Lambda), \tag{2.78}$$

which by the fact that $1/(1+\alpha) = (a+c)/(2a+c)$ yields (1.18). Formula (1.23) is a simple consequence of the way we have defined $X^{(\infty)}$ as a thinning of $\overline{X}^{(\infty)}$. This completes the proof of parts (a), (b), and (d)–(f) of the theorem.

To prove also part (c), by formula (2.76) and duality, it suffices to show that for each $t > 0$

$$\mathbb{E}\big[\big(1-(1+\alpha)\phi\big)^{X_t^{(n)}}\big] = \mathbb{E}^{\phi}\big[\big(1-(1+\alpha)\mathcal{X}_t^{\dagger}\big)^{x^{(n)}}\big] \underset{n\to\infty}{\longrightarrow} \mathbb{P}\big[\mathcal{X}_t^{\dagger} = 0\big] \qquad (\phi \in [0,1]^{\Lambda},\ |\phi| < \infty). \tag{2.79}$$

By Lemma 2.16 (i) below, $\mathcal{X}_t^{\dagger}(i) < 1$ a.s. for all $i \in \Lambda$, hence a.s. on the event $\mathcal{X}_t^{\dagger} \neq 0$ there exists some $i \in \Lambda$ such that $0 < \mathcal{X}_t^{\dagger}(i) < 1$. It follows that $|1-(1+\alpha)\mathcal{X}_t^{\dagger}|^{x^{(n)}} \to 0$ as $n \to \infty$ a.s. on the event that $\mathcal{X}_t^{\dagger} \neq 0$, hence (2.79) follows from bounded pointwise convergence. □

Remark Let $X^{(n)}$ be as in Theorem 1.4 (c). Then, using Proposition 2.12, copying the proof of [1, Thm 2 (b)], we obtain the uniform estimate

$$\mathbb{E}[X_t^{(\infty)}(i)] \leq \begin{cases} \dfrac{r'}{(2a+c)(1-e^{-r't})} & \text{if } r' \neq 0, \\ \dfrac{1}{(2a+c)t} & \text{if } r' = 0 \end{cases} \qquad (i \in \Lambda), \tag{2.80}$$

where $r' := 2a + b + c - d$. It is easy to see that this estimate is always worse than the estimate (1.18) that we obtained with the help of thinning (Corollary 1.3).

2.7 Particles everywhere

The aim of this section is to prove Lemma 2.15 below, which, roughly speaking, says that if we start a (q,a,b,c,d)-branco-process in a nontrivial spatially homogeneous initial law, then for each $t > 0$, if we look at sufficiently many sites, then we are sure to find a particle somewhere. For zero annihilation rate, this has been proved in [1, Lemma 6]. Results of this type are well-known, see e.g. the proof of [18, Thm III.5.18]. It seems the main idea of the proof, and in particular the use of Hölder's inequality in (2.88) below or in [18, (III.5.30)] goes back to Harris [15]. Another essential ingredient of the proof is some form of almost independence for events that are sufficiently far apart. For systems where the number of particle per site is bounded from above, such asymptotic independence follows from [18, Thm I.4.6], but for branco-processes, the uniform estimate given there is not available. In [1], we solved this problem by using monotonicity, which is also not available in the presence of annihilation. Instead, we will base our proof on the covariance estimate from Proposition 2.9 above.

Lemma 2.13. (Particles at the origin) *Let G be a transitive subgroup of $\mathrm{Aut}(\Lambda)$ and let μ be a G-homogeneous probability measure on $\mathcal{E}_\gamma(\Lambda)$. Assume that $b > 0$. Then, for a.e. x w.r.t. μ, the (q,a,b,c,d)-branco-process started in $X_0 = x$ satisfies*

$$\mathbb{P}^x[X_t(0) > 0] > 0 \qquad (t > 0). \tag{2.81}$$

Proof. Although the statement is intuitively obvious, some work is needed to make this rigorous. If $a = 0$, then by monotonicity (see Lemma 2.2, which extends to infinite initial states by Corollary 2.6), it suffices to prove that for a.e. x w.r.t. μ, there exists some $i \in \Lambda$ with $x(i) > 0$ such that there is a positive probability that a random walk with jump rates q, started in i, is at time t in the origin. Since we are only assuming a weak form of irreducibility (see (1.1) (ii)), this is not entirely obvious, but it is nevertheless true as has been proved in [1, Lemma 31].

If $a > 0$, then, to avoid problems stemming from the non-monotonicity of X, we use duality. Let α, r, s, m be as in (1.8) and observe that $m > 0$ by our assumptions that $a, b > 0$. Define $\delta_0 \in [0,1]^\Lambda$ by $\delta_0(i) := 1_{\{i=0\}}$. Then, by duality (Proposition 1.1), letting $\mathcal{X}$ denote the $(q^\dagger, r, s, m)$-resem-process started in $\mathcal{X}_0 = \delta_0$, we have

$$\mathbb{E}^x\big[(1-(1+\alpha))^{X_t(0)}\big] = \mathbb{E}^{\delta_0}\big[(1-(1+\alpha)\mathcal{X}_t^\dagger)^x\big], \tag{2.82}$$

and our claim will follow once we show that for all $t > 0$, this quantity is strictly less than one for a.e. x w.r.t. μ. Thus, it suffices to show that $\mathbb{P}^{\delta_0}[0 < \mathcal{X}_t(i) < 1] > 0$ for some $i \in \Lambda$ such that $x(i) > 0$. By the fact that $m > 0$ and Lemma 2.16 (i) below, this can be relaxed to showing that $\mathbb{P}^{\delta_0}[\mathcal{X}_t(i) > 0] > 0$ for some $i \in \Lambda$ such that $x(i) > 0$. Letting $\tilde{X}$ denote the $(q, 0, s, r, m)$-branco-process, using duality again (this time with $\alpha = 0$), it suffices to show that

$$1 > \mathbb{E}^{\delta_0}\big[(1-\mathcal{X}_t^\dagger)^x\big] = \mathbb{E}^x\big[0^{\tilde{X}_t(0)}\big] = \mathbb{P}^x\big[\tilde{X}_t(0) = 0\big]. \tag{2.83}$$

Thus, the statement for systems with annihilation rate $a > 0$ follows from the statement for systems with $a = 0$. □

Lemma 2.14. (Finiteness of moments) *Let X be a (q, a, b, c, d)-branco-process started in an arbitrary initial law on $\mathcal{E}_\gamma(\Lambda)$. Assume that (Λ, q) is homogeneous and that $a+c > 0$. Then*

$$\mathbb{E}[X_t(i)^m] < \infty \qquad (m \geq 1,\ i \in \Lambda,\ t > 0). \tag{2.84}$$

Proof. By Lemma 2.2 and Corollary 2.6, for each $t > 0$ we can couple a (q, a, b, c, d)-branco-process X started in an arbitrary initial law on $\mathcal{E}_\gamma(\Lambda)$ to the $(q, 0, b, a+c, d)$-branco-process X' started in the same initial law, in such a way that $X_t \leq X'_t$ a.s. In view of this, it suffices to prove the statement for the system X' with zero annihilation rate and annihilation rate $c' := a+c$. Let $X^{'(n)}$ be the $(q, 0, b, c', d)$-branco-process started in $X^{'(n)}(i) = X'_0(i) \vee n$ $(i \in \Lambda)$. Then, by [1, Theorem 2 (c)], for each $t > 0$ the process $X_t^{'(n)}$ can be coupled to the process started at infinity, denoted by $X^{(\infty)}$, in such a way that $X_t^{'(n)} \uparrow X_t^{(\infty)}$ a.s. In view of this, it suffices to prove that for the process without annihiation started at infinity

$$\mathbb{E}[X_t^{(\infty)}(i)^m] < \infty \qquad (m \geq 1,\ i \in \Lambda,\ t > 0). \tag{2.85}$$

Let $X^{(n)}$ denote the $(q, 0, b, c', d)$-branco-process started in the constant initial state $X_0^{(n)}(i) = n$ $(i \in \Lambda)$. Again by [1, Theorem 2 (c)], it suffices to find upper bounds on $\mathbb{E}[X_t^{(n)}(i)^m]$ that are uniform in n. Such upper bounds have been derived in [8, Lemma (2.13)] for branching-coalescing particle systems on $\mathbb{Z}^d$ with more general branching mechanisms than considered in the present paper. In particular, their result includes $(q, 0, b, c', d)$-branco-processes on $\mathbb{Z}^d$ with $c' > 0$. Their arguments are not restricted to $\mathbb{Z}^d$ and apply more generally to underlying lattices Λ and jump kernels q as considered in the present paper, as long as (Λ, q) is homogeneous. □

Remark It seems likely that the assumption in Lemma 2.14 that (Λ, q) is homogeneous is not needed. The proof of [8, Lemma (2.13)], which we apply here, uses translation invariance in an essential way, however. Since we do not need Lemma 2.14 in the inhomogeneous case, we will be satisfied with the present statement. It does not seem easy to adapt the proof of formula (1.18) (which holds without a homogeneity assumption) to obtain estimates for higher moments.

Lemma 2.15. (Systems with particles everywhere) *Assume that (Λ, q) is infinite and homogeneous, G is a transitive subgroup of $\mathrm{Aut}(\Lambda, q)$, and $a + c > 0$, $b > 0$. Let X be a (q, a, b, c, d)-branco-process started in a G-homogeneous nontrivial initial law on $\mathcal{E}_\gamma(\Lambda)$. Then, for any $t > 0$,*

$$\lim_{n\to\infty} \mathbb{P}[\mathrm{Thin}_{\phi_n}(X_t) = 0] = 0 \tag{2.86}$$

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

Proof. By Lemma 2.14, restarting the process at some small positive time if necessary, we can without loss of generality assume that $\mathbb{E}[X_0(0)^2] < \infty$. Set $\pi_n := \phi_n/|\phi_n|$ and let $\mathbb{P}^x$ denote the law of the process started in a deterministic initial state x. Then, for each $r < \infty$ and $t > 0$, we can choose n sufficiently large such that $r \le |\phi_n|$. Then a $r\pi_n$-thinning is stochastically less than a ϕ_n-thinning and therefore

$$\begin{aligned}
&\mathbb{P}^x[\mathrm{Thin}_{\phi_n}(X_t) = 0] \le \mathbb{P}^x[\mathrm{Thin}_{r\pi_n}(X_t) = 0]\\
&\quad = \mathbb{E}^x\big[\prod_i (1 - r\pi_n(i))^{X_t(i)}\big] \le \mathbb{E}^x\big[\prod_i e^{-r\sum_i \pi_n(i)X_t(i)}\big]\\
&\quad =: \prod_{i\in A_n} \mathbb{E}^x\big[e^{-r\pi_n(i)X_t(i)}\big] + R_n(x) \le \prod_{i\in A_n} \mathbb{E}^x\big[e^{-X_t(i)}\big]^{r\pi_n(i)} + R_n(x),
\end{aligned} \tag{2.87}$$

where in the last step we have applied Jensen's inequality to the concave function $z \mapsto z^{r\pi_n(i)}$. For the process started in a nontrivial homogeneous initial law, we obtain, using Hölder's inequality, for all n sufficiently large such that $r \le |\phi_n|$,

$$\begin{aligned}
\mathbb{P}[\mathrm{Thin}_{\phi_n}(X_t) = 0] &= \int \mathbb{P}[X_0 \in \mathrm{d}x]\mathbb{P}^x[\mathrm{Thin}_{\phi_n}(X_t) = 0]\\
&\le \int \mathbb{P}[X_0 \in \mathrm{d}x]\Big[\prod_{i\in A_n} \mathbb{E}^x\big[e^{-X_t(i)}\big]^{r\pi_n(i)} + R_n(x)\Big]\\
&\le \prod_{i\in A_n} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}x]\mathbb{E}^x\big[e^{-X_t(i)}\big]^r\Big)^{\pi_n(i)} + \mathbb{E}[R_n(X_0)]\\
&= \prod_{i\in A_n} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}x]\mathbb{E}^x\big[e^{-X_t(0)}\big]^r\Big)^{\pi_n(i)} + \mathbb{E}[R_n(X_0)]\\
&= \int \mathbb{P}[X_0 \in \mathrm{d}x]\mathbb{E}^x\big[e^{-X_t(0)}\big]^r + \mathbb{E}[R_n(X_0)],
\end{aligned} \tag{2.88}$$

where we have used spatial homogeneity in the last step but one.

By Corollary 2.10, the quantity $R_n(x)$ defined in (2.87) can be estimated as

$$|R_n(x)| \le \tfrac{1}{2}r^2 \sum_{\substack{k,l\\k\neq l}} \Big(\sum_i x(i)K_t(i;k,l) + \sum_{i,j} x(i)x(j)L_t(i,j;k,l)\Big)\pi_n(k)\pi_n(l). \tag{2.89}$$

It follows that

$$\begin{aligned}
&\mathbb{E}\big[\big|R_n(X_0(0))\big|\big]\\
&\leq \tfrac{1}{2}r^2\sum_{\substack{k,l\\k\neq l}}\Big(\sum_i \mathbb{E}[X_0(i)]K_t(i;k,l)+\sum_{i,j}\mathbb{E}[X_0(i)X_0(j)]L_t(i,j;k,l)\Big)\pi_n(k)\pi_n(l)\\
&\leq \tfrac{1}{2}r^2\sum_{\substack{k,l\\k\neq l}}\Big(\mathbb{E}[X_0(0)]\sum_i K_t(i;k,l)+\mathbb{E}[X_0(0)^2]\sum_{i,j}L_t(i,j;k,l)\Big)\pi_n(k)\pi_n(l)\\
&=: r^2\sum_{k,l}C(k,l)\pi_n(k)\pi_n(l),
\end{aligned} \tag{2.90}$$

where by definition $C(k,k):=0$ and we have used that by Cauchy-Schwartz and translation invariance:

$$\big|\mathbb{E}[X_0(i)X_0(j)]\big|\leq \mathbb{E}[X_0(i)^2]^{1/2}\mathbb{E}[X_0(j)^2]^{1/2}=\mathbb{E}[X_0(0)^2]. \tag{2.91}$$

We claim that

$$\sum_{k,l}C(k,l)\pi_n(k)\pi_n(l)\underset{n\to\infty}{\longrightarrow}0. \tag{2.92}$$

To see this, we observe that by (2.34), (2.35) and our assumption that $\mathbb{E}[X_0(0)^2]<\infty$,

$$C(gk,gl)=C(k,l)\quad (g\in G)\quad\text{and}\quad \sum_k C(k,0)<\infty. \tag{2.93}$$

Since G is transitive, for each $l\in\Lambda$ we can choose some $g_l\in G$ such that $g_l l=0$. In view of this, (2.93) shows in particular that for each $\varepsilon>0$, the quantity

$$\big|\{k\in\Lambda: C(k,l)\geq\varepsilon\}\big|=\big|\{g_lk\in\Lambda: C(g_lk,0)\geq\varepsilon\}\big|=\big|\{j\in\Lambda: C(j,0)\geq\varepsilon\}\big|=: K_\varepsilon \tag{2.94}$$

does not depend on $l\in\Lambda$ and is finite. It follows that

$$\sum_l\pi_n(l)\sum_k C(k,l)\pi_n(k)\leq\sum_l\pi_n(l)\Big(\sum_{k:\,C(k,l)\geq\varepsilon}\pi_n(k)+\sum_{k:\,C(k,l)<\varepsilon}\pi_n(k)\Big)\leq K_\varepsilon/|\phi_n|+\varepsilon. \tag{2.95}$$

Since $|\phi_n|\to\infty$ and $\varepsilon>0$ is arbitrary, this proves (2.92). By (2.88) and (2.90), we conclude that for each $r<\infty$,

$$\limsup_{n\to\infty}\mathbb{P}[\mathrm{Thin}_{\phi_n}(X_t)=0]\leq\int\mathbb{P}[X_0\in\mathrm{d}x]\mathbb{E}^x\big[e^{-X_t(0)}\big]^r. \tag{2.96}$$

Letting $r\to\infty$, using $b>0$ and Lemma 2.13, we arrive at (2.86). □

2.8 Long-time limit law

In this section, we prove Theorem 1.5. We first need some preparatory results.

Lemma 2.16. (Not exactly one) *Let $\mathcal{X}$ be a (q,r,s,m)-resem process started in a finite initial state $\phi\in[0,1]^\Lambda$, $|\phi|<\infty$. Assume that (Λ,q) is infinite and homogeneous and that $m>0$. Then*

(i) $\mathbb{P}^\phi[\mathcal{X}_t(i)=1]=0$ *for each* $t>0$, $i\in\Lambda$.

(ii) $\mathbb{P}^\phi[0<|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|<K]\to 0$ *as* $t\to\infty$ *for all* $K<\infty$.

Proof. Let $\mathcal{X}^+$ and $\mathcal{X}^-$ satisfy $\mathcal{X}^+_0=\mathcal{X}^-_0=\mathcal{X}_0=\phi$ and be given, for times $t>0$, by the solutions to the stochastic differential equations

$$\begin{aligned} \mathrm{d}\mathcal{X}^+_t(i) &= \big(|q|+s\big)(1-\mathcal{X}^+_t(i))\mathrm{d}t - m\mathcal{X}^+_t(i)\mathrm{d}t + \sqrt{2r\mathcal{X}^+_t(i)(1-\mathcal{X}^+_t(i))}\mathrm{d}B_t(i) \\ \mathrm{d}\mathcal{X}^-_t(i) &= -\big(|q|+m\big)\mathcal{X}^-_t(i)\mathrm{d}t + \sqrt{2r\mathcal{X}^-_t(i)(1-\mathcal{X}^-_t(i))}\mathrm{d}B_t(i) \qquad (t\geq 0,\ i\in\Lambda), \end{aligned} \tag{2.97}$$

where $|q|:=\sum_j q(i,j)$, which does not depend on $j\in\Lambda$ by the transitivity of $\mathrm{Aut}(\Lambda,q)$, and $(B(i))_{i\in\Lambda}$ is the same collection of independent Brownian motions as those driving $\mathcal{X}$. By the arguments used in the proof of [1, Lemma 18], solutions of (2.97) are pathwise unique and satisfy

$$\mathcal{X}^-_t \leq \mathcal{X}_t \leq \mathcal{X}^+_t \quad (t\geq 0) \quad \text{a.s.} \tag{2.98}$$

Moreover, since formula (2.97) contains no interaction terms, the $[0,1]^2$-valued processes $(\mathcal{X}^-_t(i),\mathcal{X}^+_t(i))_{t\geq 0}$ are independent for different values of $i\in\Lambda$. Since $\mathcal{X}^+(i)$ is a one-dimensional diffusion with (by grace of the fact that $m>0$) the drift on the boundary point 1 pointing inwards, it can be proved by standard methods that

$$\mathbb{P}[\mathcal{X}^+_t(i)=1]=0 \qquad (t>0,\ i\in\Lambda). \tag{2.99}$$

We defer a precise proof of this fact to Lemma A.1 in the appendix. Together with (2.98), formula (2.99) proves part (i) of the lemma.

To prove also part (ii), we observe that

$$\mathbb{E}[\mathcal{X}^-_t(i)] = e^{-(|q|+m)t}\phi(i) \qquad (t>0,\ i\in\Lambda). \tag{2.100}$$

With a bit of work, it is possible to show that there exists a $t_0>0$ and function $(0,t_0]\ni t\mapsto c_t>0$ such that

$$\mathbb{E}\big[\mathcal{X}^-_t(i)\wedge(1-\mathcal{X}^+_t(i))\big] \geq c_t\phi(i) \qquad (0<t\leq t_0). \tag{2.101}$$

A precise proof of this fact can be found in Lemma A.4 of the appendix. We note that for any $[0,1]$-valued random variable Z, one has $\mathrm{Var}(Z)=\mathbb{E}[(Z-\mathbb{E}[Z])^2]\leq\mathbb{E}[|Z-\mathbb{E}[Z]|]\leq\mathbb{E}[Z+\mathbb{E}[Z]]=2\mathbb{E}[Z]$. Applying this to $Z=\mathcal{X}^-_t(i)\wedge(1-\mathcal{X}^+_t(i))$, using (2.100), we see that

$$\mathrm{Var}\big(\mathcal{X}^-_t(i)\wedge(1-\mathcal{X}^+_t(i))\big) \leq 2e^{-(|q|+m)t}\phi(i) \qquad (t>0,\ i\in\Lambda). \tag{2.102}$$

Now (2.101) implies $\mathbb{E}[|\mathcal{X}^-_t\wedge(1-\mathcal{X}^+_t)|]\geq c_t|\phi|$, while by (2.102) and the independence of coordinates $i\in\Lambda$,

$$\mathrm{Var}\big(|\mathcal{X}^-_t\wedge(1-\mathcal{X}^+_t)|\big) \leq 2e^{-(|q|+m)t}|\phi| \qquad (0<t\leq t_0). \tag{2.103}$$

Since $\mathcal{X}_t(i)\wedge(1-\mathcal{X}_t(i))\geq\mathcal{X}^-_t(i)\wedge(1-\mathcal{X}^+_t(i))$, by Chebyshev, it follows that

$$\mathbb{P}^\phi\big[|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|\leq\tfrac12 c_t|\phi|\big] \leq \frac{2e^{-(|q|+m)t}|\phi|}{\frac14 c_t^2|\phi|^2} \qquad (0<t\leq t_0), \tag{2.104}$$

which tends to zero for $|\phi|\to\infty$. By [1, Lemma 5],

$$\mathbb{P}^\phi\big[0<|\mathcal{X}_t|<K\big] \underset{t\to\infty}{\longrightarrow} 0 \qquad (K<\infty). \tag{2.105}$$

It follows that we can choose $L_t\to\infty$ slow enough such that

$$\mathbb{P}^\phi\big[0<|\mathcal{X}_t|<L_t\big] \underset{t\to\infty}{\longrightarrow} 0. \tag{2.106}$$

By (2.104), we conclude that

$$\begin{aligned}
&\limsup_{t\to\infty} \mathbb{P}^\phi\big[0<|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|<K\big]\\
&\quad\le \limsup_{t\to\infty} \mathbb{P}^\phi\big[0<|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|<K\,\big|\,0<|\mathcal{X}_{t-t_0}|<L_{t-t_0}\big]\mathbb{P}^\phi\big[0<|\mathcal{X}_{t-t_0}|<L_{t-t_0}\big]\\
&\qquad+\limsup_{t\to\infty} \mathbb{P}^\phi\big[0<|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|<K\,\big|\,|\mathcal{X}_{t-t_0}|\ge L_{t-t_0}\big]\mathbb{P}^\phi\big[|\mathcal{X}_{t-t_0}|\ge L_{t-t_0}\big]\\
&\quad\le \limsup_{t\to\infty} \mathbb{P}^\phi\big[|\mathcal{X}_t\wedge(1-\mathcal{X}_t)|\le \tfrac{1}{2}c_{t_0}L_{t-t_0}\,\big|\,|\mathcal{X}_{t-t_0}|\ge L_{t-t_0}\big]\\
&\quad\le \limsup_{t\to\infty} \frac{2e^{-(|q|+m)t}L_{t-t_0}}{\frac{1}{4}c_{t_0}^2L_{t-t_0}^2}=0.
\end{aligned} \tag{2.107}$$

□

Remark It seems likely that the condition $m>0$ in Lemma 2.16 is not necessary, at least for part (i). Indeed, it seems likely that (q,r,s,m)-resem-processes have the 'noncompact support property'

$$\mathbb{P}\big[\mathcal{X}_t(i)>0,\ \mathcal{X}_t(j)=0\big]=0 \qquad \big(t>0,\ i,j\in\Lambda,\ q(i,j)>0\big), \tag{2.108}$$

similar to what is known for super random walks [12]. Since proving (2.108) is quite involved and we don't know a reference, we will be satisfied with proving Lemma 2.16 only for $m>0$, which is sufficient for our purposes.

Lemma 2.17. (Systems with particles everywhere) *Assume that (Λ,q) is infinite and homogeneous and that G is a transitive subgroup of $\mathrm{Aut}(\Lambda,q)$ and $a+c>0$, $b>0$. Let X be the (q,a,b,c,d)-branco process started in a G-homogeneous nontrivial initial law $\mathcal{L}(X_0)$. Then, for any $t>0$ and $0\le\alpha\le1$ and for any $\varepsilon>0$, there exists a $K<\infty$ such that*

$$|\phi|<\infty \text{ and } |\phi\wedge(1-\phi)|\ge K \text{ implies } \mathbb{E}\big[|1-(1+\alpha)\phi|^{X_t}\big]\le\varepsilon. \tag{2.109}$$

Proof. We start by proving that if $\phi_n\in[0,1]^\Lambda$ satisfy $|\phi_n|<\infty$ and $|\phi_n\wedge(1-\phi_n)|\to\infty$, then

$$\lim_{n\to\infty}\mathbb{E}\big[|1-(1+\alpha)\phi_n|^{X_t}\big]=0. \tag{2.110}$$

Set $\psi_n:=\phi_n\wedge(1-\phi_n)$. Then, for each $i\in\Lambda$, we have and $\psi_n(i)\le1-\phi_n(i)\le 2-(1+\alpha)\phi_n(i)$ and $\psi_n(i)\le\phi_n(i)\le(1+\alpha)\phi_n(i)$, from which we see that

$$\psi_n(i)-1\le 1-(1+\alpha)\phi_n(i)\le 1-\psi_n(i), \tag{2.111}$$

or, in other words, $|1-(1+\alpha)\phi_n(i)|\le1-\psi_n$. It follows that

$$\big|\mathbb{E}\big[1-(1+\alpha)\phi_n^{X_t}\big]\big|\le\mathbb{E}\big[|1-(1+\alpha)\phi_n|^{X_t}\big]\le\mathbb{E}\big[(1-\psi_n)^{X_t}\big]=\mathbb{P}\big[\mathrm{Thin}_{\psi_n}(X_t)=0\big], \tag{2.112}$$

which tends to zero by Lemma 2.15 and our assumption that $|\psi_n|\to\infty$.

Now imagine that the lemma does not hold. Then there exists some $\varepsilon>0$ such that for all $n\ge1$ we can choose ϕ_n with $|\phi_n|<\infty$ and $|\phi_n\wedge(1-\phi_n)|\ge n$ such that $\mathbb{E}\big[|1-(1+\alpha)\phi_n|^{X_t}\big]>\varepsilon$. Since this contradicts (2.110), we conclude that the lemma must hold. □

Proof of Theorem 1.5. For $a=0$ the statement has been proved in [1, Thm 4 (a)], so without loss of generality we may assume that $a>0$. By Theorem 1.4 (e), it suffices to show that

$$\mathbb{E}\big[\big(1-(1+\alpha)\phi\big)^{X_t}\big]\underset{t\to\infty}{\longrightarrow}\mathbb{P}^\phi\big[\exists t\ge0 \text{ such that } \mathcal{X}_t^\dagger=0\big] \qquad (\phi\in[0,1]^\Lambda,\ |\phi|<\infty), \tag{2.113}$$

where $\alpha := a/(a+c)$ and $\mathcal{X}^\dagger$ denotes the $(q^\dagger, a+c, (1+\alpha)b, \alpha b+d))$-resem-process started in ϕ. By duality (Proposition 1.1), for each $t \geq 1$,

$$\mathbb{E}\big[\big(1-(1+\alpha)\phi\big)^{X_t}\big] = \mathbb{E}\big[\big(1-(1+\alpha)\mathcal{X}^\dagger_{t-1}\big)^{X_1}\big], \tag{2.114}$$

where $\mathcal{X}^\dagger$ is independent of X and started in $\mathcal{X}^\dagger_0 = \phi$. For each $K < \infty$, we may write

$$\begin{aligned}
&\mathbb{E}\big[\big(1-(1+\alpha)\mathcal{X}^\dagger_{t-1}\big)^{X_1}\big] = \mathbb{P}[\mathcal{X}^\dagger_{t-1} = 0]\\
&+\mathbb{E}\big[\big(1-(1+\alpha)\mathcal{X}^\dagger_{t-1}\big)^{X_1}\,\big|\,0 < |\mathcal{X}^\dagger_{t-1} \wedge (1-\mathcal{X}^\dagger_{t-1})| < K\big]\mathbb{P}\big[0 < |\mathcal{X}^\dagger_{t-1} \wedge (1-\mathcal{X}^\dagger_{t-1})| < K\big]\\
&+\mathbb{E}\big[\big(1-(1+\alpha)\mathcal{X}^\dagger_{t-1}\big)^{X_1}\,\big|\,K \leq |\mathcal{X}^\dagger_{t-1} \wedge (1-\mathcal{X}^\dagger_{t-1})|\big]\mathbb{P}\big[K \leq |\mathcal{X}^\dagger_{t-1} \wedge (1-\mathcal{X}^\dagger_{t-1})|\big].
\end{aligned} \tag{2.115}$$

Here the first term converges, as $t \to \infty$, to $\mathbb{P}^\phi\big[\exists t \geq 0 \text{ such that } \mathcal{X}^\dagger_t = 0\big]$. Note that $\alpha > 0$ by our assumption that $a > 0$. Assume for the moment that also $b > 0$. Then Lemma 2.16 (ii) tells us that the second term on the right-hand side of (2.115) tends to zero. By Lemma 2.17, for each $\varepsilon > 0$ we can choose K large enough such that the third term is bounded in absolute value by ε. Putting these things together, we arrive at (2.113).

If $b = 0$, then Lemma 2.16 (ii) is not available, but in this case $|\mathcal{X}^\dagger_t|$ is a supermartingale, hence [1, Lemma 5] tells us that $\mathbb{P}^\phi\big[\exists t \geq 0 \text{ such that } \mathcal{X}^\dagger_t = 0\big] = 1$, and the proof proceeds as above. □

A Some facts about coupled Wright-Fisher diffusions

The aim of this appendix is to prove two simple facts about (coupled) Wright-Fisher diffusions. In particular, applying Lemmas A.1 and A.4 to $X = \mathcal{X}^+(i)$, $Y = \mathcal{X}^-(i)$, $a = |q| + s$, $b = m$ and $c = |q| + m$ yields formulas (2.99) and (2.101), respectively.

For $a, b \geq 0$ and $r > 0$, let X denote the pathwise unique (by [32]) $[0,1]$-valued solution to the stochastic differential equation

$$\mathrm{d}X_t = a(1-X_t)\mathrm{d}t - bX_t\mathrm{d}t + \sqrt{2rX(1-X)}\mathrm{d}B_t, \tag{A.1}$$

where B is standard Brownian motion.

Lemma A.1. (No mass on boundary) *If $b > 0$, then*

$$\mathbb{P}[X_t = 1] = 0 \qquad (t > 0), \tag{A.2}$$

regardless of the initial law.

Proof. If $a, b > 0$, then it is well known that X has a transition density (see Propositions 3 and 4 in [20] along with the discussion on page 1183 or [14, 13]). Consequently $\mathbb{P}[X_t = 1] = 0$ and hence the result follows. If $a = 0$ but $b > 0$, then by standard comparision results (see [4, Thm. 6.2] or [1, Lemma 18]), if $Z_0 = X_0$ and Z solves the SDE (A.1) with $a = b/2$ and b replaced by $b/2$, relative to the same Brownian motion, then $X_t \leq Z_t$ and hence $\mathbb{P}[X_t = 1] \leq \mathbb{P}[Z_t = 1] = 0$ for all $t > 0$. □

Lemma A.2. (Moment dual) *Let $K = (K_t)_{t\geq 0}$ be a Markov process with state space $\mathbb{N} \cup \{\infty\}$, where ∞ is a trap, and K jumps from states $k \in \mathbb{N}$ as*

$$\begin{aligned}
&k \mapsto k-1 &&\text{with rate } ak + rk(k-1),\\
&k \mapsto \infty &&\text{with rate } bk.
\end{aligned} \tag{A.3}$$

Then

$$\mathbb{E}^x[X_t^k] = \mathbb{E}^k[x^{K_t}] \qquad (t \geq 0,\ x \in [0,1],\ k \in \mathbb{N}), \tag{A.4}$$

where $x^0 := 1$ and $x^\infty := 0$ for all $x \in [0,1]$.

Proof. Let

$$\begin{aligned}\mathcal{G}f(x) &:= \big[a(1-x)\tfrac{\partial}{\partial x} - bx\tfrac{\partial}{\partial x} + rx(1-x)\tfrac{\partial^2}{\partial x^2}\big]f(x),\\ Gf(k) &:= \big[ak + rk(k-1)\big]\big\{f(k-1) - f(k)\big\} + bk\big\{f(\infty) - f(k)\big\}\end{aligned} \tag{A.5}$$

be the generators of the processes X and K, respectively, and let $\psi(x,k) := x^k$ be the duality function. Then

$$\mathcal{G}\psi(\,\cdot\,,k)(x) = ak(x^{k-1} - x^k) - bkx^k + rk(k-1)(x^{k-1} - x^k) = G\psi(x,,\,\cdot\,)(k), \tag{A.6}$$

where the term with $k(k-1)$ is zero for $k = 1$ and both sides of the equation are zero for $k = 0$. The claim now follows from [1, Thm 7] and [3] and the fact that the expression in (A.6) is bounded uniformly in x and k, which guarantees the required integrability.

Although this is not needed for the proof, this duality may be understood as follows. We can view X_t as the frequency of type-one organisms in a large population where pairs of organisms are resampled with rate $2r$ and organisms mutate to type 1 and 0, respectively, with rates a and b. Then $\mathbb{E}[X_t^k]$ is the probability that k organisms, sampled from the population at time t, are all of type one. We can view K_t as the ancestors of these organism at time zero, where we neglect organisms that due to mutation are sure to be of type one while on the other hand the state $K_t = \infty$ signifies that due to a mutatation event, at least one of these ancestors is of type zero. □

Now let X be as in (A.1), let $c \geq 0$, and let Y be given by the pathwise unique solution to the stochastic differential equation

$$\mathrm{d}Y_t = -cY_t\mathrm{d}t + \sqrt{2rY(1-Y)}\mathrm{d}B_t, \tag{A.7}$$

driven by the *same* Brownian motion as X.

Lemma A.3. (Feller property) *Let (X,Y) be given by the pathwise unique solutions of (A.1) and (A.7), and let $K_t((x,y),\,\cdot\,) := \mathbb{P}^{(x,y)}[(X_t,Y_t) \in \,\cdot\,]$ denote the transition probabilities of (X,Y). Then the map $(t,x,y) \mapsto K_t((x,y),\,\cdot\,)$ from $[0,\infty) \times [0,1]$ into the probability measures on $[0,1]^2$ is continuous w.r.t. weak convergence of probability measures.*

Proof. It follows from well-known results [11, Corollary 5.3.4 and Theorem 5.3.6] that pathwise uniqueness for a stochastic differential equation implies uniqueness of solutions to the martingale problem for the associated differential operator, which is in our case given by

$$A := a(1-x)\tfrac{\partial}{\partial x} - bx\tfrac{\partial}{\partial x} + rx(1-x)\tfrac{\partial^2}{\partial x^2} - cy\tfrac{\partial}{\partial y} + ry(1-y)\tfrac{\partial^2}{\partial y^2} + 2r\sqrt{x(1-x)y(1-y)}\,\tfrac{\partial^2}{\partial x\partial y}, \tag{A.8}$$

with domain $\mathcal{C}^2[0,1]^2$. Now if (X^n,Y^n) are solutions to this martingale problem with deterministic initial states $(X^n_0,Y^n_0) = (x_n,y_n)$ converging to some limit $(x,y) \in [0,1]^2$, and (X,Y) denotes the process started in (x,y), then [11, Lemma 4.5.1 and Remark 4.5.2] imply that

$$\mathbb{P}[(X^n_t,Y^n_t)_{t\geq 0} \in \cdot\,] \underset{n\to\infty}{\Longrightarrow} \mathbb{P}[(X_t,Y_t)_{t\geq 0} \in \cdot\,], \tag{A.9}$$

where $\Rightarrow$ denotes weak convergence of probability laws on the space $\mathcal{C}_{[0,1]^2}[0,\infty)$ of continuous functions from $[0,\infty)$ into $[0,1]^2$, equipped with the topology of locally uniform convergence. In particular, this implies the stated continuity of the transition probabilities. □

Lemma A.4. (Linear estimate) *Assume that $b > 0$. Then there exists a $t_0 > 0$ and function $(0,t_0] \ni t \mapsto \lambda_t > 0$ such that the process started in $(X_0,Y_0) = (z,z)$ satisfies*

$$\mathbb{E}^{(z,z)}\big[Y_t \wedge (1 - X_t)\big] \geq \lambda_t z \qquad (0 < t \leq t_0,\ 0 \leq z \leq 1). \tag{A.10}$$

Proof. We estimate

$$\begin{aligned}&\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]\geq\mathbb{E}^{(z,z)}\big[Y_t(1-X_t)\big]\geq\tfrac{1}{2}\mathbb{E}^{(z,z)}\big[Y_t1_{\{X_t\leq\frac{1}{2}\}}\big]\\&\quad=\tfrac{1}{2}\big(\mathbb{E}^z[Y_t]-\mathbb{E}^z[1_{\{X_t>\frac{1}{2}\}}]\big)\geq\tfrac{1}{2}\mathbb{E}^z[Y_t]-2\mathbb{E}^z[X_t^2],\end{aligned}\tag{A.11}$$

where the last step we have used that $1_{\{x>\frac{1}{2}\}}\leq 4x^2$. By Lemma A.2,

$$\mathbb{E}^z[Y_t]=e^{-ct}z\qquad(t\geq 0,\ z\in[0,1]),\tag{A.12}$$

while by the same lemma

$$\begin{aligned}&\mathbb{E}^z[X_t^2]=\mathbb{E}^2[z^{K_t}]\leq\mathbb{P}^2[K_t\leq 1]z+\mathbb{P}^2[K_t=2]z^2\\&\quad=(1-e^{-2(a+r)t})z+e^{-2(a+b+r)t}z^2\leq 2(a+r)tz+z^2.\end{aligned}\tag{A.13}$$

Combining this with (A.11) yields

$$\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]\geq\big(\tfrac{1}{2}e^{-ct}-4(a+r)t-2z\big)z.\tag{A.14}$$

Choosing $t_0>0$ and $z_0>0$ small enough, we find that

$$\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]\geq\tfrac{1}{4}z\qquad(0\leq t\leq t_0,\ 0\leq z\leq z_0).\tag{A.15}$$

To extend this to all $z\in[0,1]$, at the cost of assuming that $t>0$ and replacing the constant $1/4$ by a possibly worse, time-dependent constant λ_t, we observe that by Lemma A.3, the function $[0,1]\ni z\mapsto\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]$ is continuous. Since by Lemma A.1 and (A.12),

$$\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]>0\qquad(t>0,\ z\in(0,1]),\tag{A.16}$$

using continuity, we may estimate $\mathbb{E}^{(z,z)}\big[Y_t\wedge(1-X_t)\big]$ uniformly from below on $[z_0,1]$, which together with (A.15) yields (A.10). □

References

[1] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. *Prob. Theory Relat. Fields.* 131(3), 376–414, 2005. MR-2123250

[2] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. Erratum. *Prob. Theory Relat. Fields.* 145(3-4), 639–640, 2009. MR-2529442

[3] S.R. Athreya and J.M. Swart. Correction to: Branching-coalescing particle systems. arXiv:0904.2288v1

[4] R.F. Bass. *Diffusions and elliptic operators.* Probability and its Applications. Springer, New York, 1998. MR-1483890

[5] M. Bramson and L. Gray. The survival of branching annihilating random walk. *Z. Wahrsch. verw. Gebiete* 68: 447–460, 1985. MR-0772192

[6] M.F. Chen. Existence theorems for interacting particle systems with non-compact state space. *Sci. China Ser. A* 30, 148–156, 1987. MR-0892470

[7] 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. MR-1242575

[8] W. Ding, R. Durrett, and T.M. Liggett. Ergodicity of reversible reaction diffusion processes. *Probab. Theory Relat. Fields* 85(1), 13–26, 1990. MR-1044295

[9] M. Bramson, W. Ding, and R. Durrett. Annihilating branching processes. *Stochastic Process. Appl.* 37(1), 1–17, 1991. MR-1091690

[10] R. Durrett. A new method for proving the existence of phase transitions. *Spatial Stochastic Processes, Festschr. in Honor of Ted Harris 70th Birthday* (K.S. Alexander and J.C. Watkins eds.) 141–169. Prog. Probab. 19, Birkhäuser, Boston, 1991. MR-1144095

[11] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* John Wiley & Sons, New York, 1986. MR-0838085

[12] S.N. Evans and E. Perkins. Absolute continuity results for superprocesses with some applications. *Trans. Am. Math. Soc.* 325(2), 661–681, 1991. MR-1012522

[13] R.C. Griffiths. On the distribution of allele frequencies in a diffusion model. *Theoret. Population Biol.* 15, 140–158, 1979. MR-0528914

[14] R.C. Griffiths. A transition density expansion for a multi-allele diffusion model. *Adv. in Appl. Probab.* 11, 310–325, 1979. MR-0526415

[15] T.E. Harris. On a class of set-valued Markov processes. *Ann. Probab.* 4, 175–194, 1976. MR-0400468

[16] E. Hewitt and K. Stromberg. *Real and abstract analysis. A modern treatment of the theory of functions of a real variable.* 3rd printing. Springer, New York, 1975. MR-0367121

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

[18] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985. MR-0776231

[19] 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. MR-0621659

[20] Soumik Pal. Analysis of market weights under volatility-stabilized market models. *Anal. Appl. Probab.* 21(3), 1180–1213, 2011. MR-2830616

[21] L. Schwartz. *Radon Measures on Arbitrary Topological Spaces and Cylindical Measures.* Tata Institute, Oxford University Press, London, 1973. MR-0426084

[22] A. Sturm and J.M. Swart. Voter models with heterozygosity selection. *Ann. Appl. Probab.* 18(1), 59–99, 2008. MR-2380891

[23] A. Sudbury and P. Lloyd. Quantum operators in classical probability theory. II: The concept of duality in interacting particle systems. *Ann. Probab.* 23(4), 1816–1830, 1995. MR-1379169

[24] A. Sudbury and P. Lloyd. Quantum operators in classical probability theory. IV: Quasi-duality and thinnings of interacting particle systems *Ann. Probab.* 25(1), 96–114, 1997. MR-1428501

[25] 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. MR-0849066

[26] A. Sudbury. The branching annihilating process: an interacting particle system. *Ann. Probab.* 18: 581–601, 1990. MR-1055421

[27] A. Sudbury. Dual families of interacting particle systems on graphs. *J. Theor. Probab.* 13(3), 695–716, 2000. MR-1785526

[28] J.M. Swart. *Large Space-Time Scale Behavior of Linearly Interacting Diffusions.* PhD thesis, Katholieke Universiteit Nijmegen, 1999. `http://webdoc.ubn.kun.nl/mono/s/swart_j/largspscb.pdf`

[29] J.M. Swart. Duals and thinnings of some relatives of the contact process. *Preprint* (18 pages). arXiv:0604335

[30] J.M. Swart. Duals and thinnings of some relatives of the contact process. Pages 203–214 in: *Prague Stochastics 2006*, M. Hušková and M. Janžura (eds.), Matfyzpress, Prague, 2006.

[31] J.M. Swart. The contact process seen from a typical infected site. *J. Theoret. Probab.* 22(3), 711-740, 2009. MR-2530110

[32] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* 11, 155–167, 1971. MR-0278420

Acknowledgments. Work sponsored by GAČR grant: P201/10/0752. Part of this work was done when the authors were visiting the Universität Duisburg-Essen. We would like to thank Anita Winter and the staff for their hospitality; Soumik Pal for useful discussions; and an anonymous referee for a careful reading of the paper.