Renormalization analysis of catalytic Wright-Fisher diffusions

Klaus Fleischmann
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstr. 39
D–10117 Berlin, Germany
e-mail: fleischm@wias-berlin.de

Jan M. Swart
Mathematical Institute
University Tübingen
Auf der Morgenstelle 10
D–72076 Tübingen, Germany
e-mail: jan.swart@uni-tuebingen.de

June 14, 2005

Abstract

Recently, several authors have studied maps where a function, describing the local diffusion matrix of a diffusion process with a linear drift towards an attraction point, is mapped into the average of that function with respect to the unique invariant measure of the diffusion process, as a function of the attraction point. Such mappings arise in the analysis of infinite systems of diffusions indexed by the hierarchical group, with a linear attractive interaction between the components. In this context, the mappings are called renormalization transformations. We consider such maps for catalytic Wright-Fisher diffusions. These are diffusions on the unit square where the first component (the catalyst) performs an autonomous Wright-Fisher diffusion, while the second component (the reactant) performs a Wright-Fisher diffusion with a rate depending on the first component through a catalyzing function. We determine the limit of rescaled iterates of renormalization transformations acting on the diffusion matrices of such catalytic Wright-Fisher diffusions.

Contents

Part I

1 Introduction and main result

Several authors [BCGdH95, BCGdH97, dHS98, Sch98, CDG04] have studied maps where a function, describing the local diffusion matrix of a diffusion process, is mapped into the average of that function with respect to the unique invariant measure of the diffusion process itself. Such mappings arise in the analysis of infinite systems of diffusion processes indexed by the hierarchical group, with a linear attractive interaction between the components [DG93a, DG96, DGV95]. In this context, the mappings are called renormalization transformations. We follow this terminology. For more on the relation between hierarchically interacting diffusions and renormalization transformations, see Appendix A.1.

Formally, such renormalization transformations can be defined as follows.

Definition 1.1 (Renormalization class and transformation) Let $D \subset \mathbb{R}^d$ be nonempty, convex, and open. Let $\mathcal{W}$ be a collection of continuous functions w from the closure $\overline{D}$ into the space M_+^d of symmetric non-negative definite $d \times d$ real matrices, such that $\lambda w \in \mathcal{W}$ for every $\lambda > 0$, $w \in \mathcal{W}$. We call $\mathcal{W}$ a *prerenormalization class* on $\overline{D}$ if the following three conditions are satisfied:

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

$$A_x^{c,w} f(y) := \sum_{i=1}^{d} c\,(x_i - y_i)\tfrac{\partial}{\partial y_i} f(y) + \sum_{i,j=1}^{d} w_{ij}(y)\tfrac{\partial^2}{\partial y_i \partial y_j} f(y) \qquad (y \in \overline{D}), \tag{1.1}$$

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

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

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

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

$$F_c w(x) := \int_{\overline{D}} \nu_x^{c,w}(dy) w(y) \qquad (x \in \overline{D}). \tag{1.2}$$

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

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

If $\mathcal{W}$ is a renormalization class and $c > 0$, then the map $F_c : \mathcal{W} \to \mathcal{W}$ defined by (1.2) is called the *renormalization transformation* on $\mathcal{W}$ with *migration constant* c. In (1.1), w is called the *diffusion matrix* and x the *attraction point.* $\diamond$

Remark 1.2 (Associated SDE) It is well-known that $\overline{D}$-valued (weak) solutions $\mathbf{y} = (\mathbf{y}^1, \ldots, \mathbf{y}^d)$ to the stochastic differential equation (SDE)

$$\mathrm{d}\mathbf{y}^i_t = c\,(x_i - \mathbf{y}^i_t)\mathrm{d}t + \sqrt{2}\sum_{j=1}^{n} \sigma_{ij}(\mathbf{y}_t)\mathrm{d}B^j_t \qquad (t \geq 0,\ i = 1, \ldots, d), \tag{1.3}$$

where $B = (B^1, \ldots, B^n)$ is n-dimensional (standard) Brownian motion ($n \geq 1$), solve the martingale problem for $A^{c,w}_x$ if the $d \times n$ matrix-valued function σ is continuous and satisfies $\sum_k \sigma_{ik}\sigma_{jk} = w_{ij}$. Conversely [EK86, Theorem 5.3.3], every solution to the martingale problem for $A^{c,w}_x$ can be represented as a solution to the SDE (1.3), where there is some freedom in the choice of the root σ of the diffusion matrix w. $\diamond$

In the present paper, we concern ourselves with the following renormalization class on $[0, 1]^2$.

Definition 1.3 (Renormalization class of catalytic Wright-Fisher diffusions) We set $\mathcal{W}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{H}\}$, where

$$w^{\alpha,p}(x) := \begin{pmatrix} \alpha x_1(1 - x_1) & 0 \\ 0 & p(x_1)x_2(1 - x_2) \end{pmatrix} \qquad (x = (x_1, x_2) \in [0, 1]^2), \tag{1.4}$$

and

$$\mathcal{H} := \{p : p \text{ a real function on } [0, 1],\ p \geq 0,\ p \text{ Lipschitz continuous}\}. \tag{1.5}$$

Moreover, we put

$$\mathcal{H}_{l,r} := \{p \in \mathcal{H} :\ 1_{\{p(0)>0\}} = l,\ 1_{\{p(1)>0\}} = r\} \qquad (l, r = 0, 1), \tag{1.6}$$

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

By Remark 1.2, solutions $\mathbf{y} = (\mathbf{y}^1, \mathbf{y}^2)$ to the martingale problem for $A^{c,w^{\alpha,p}}_x$ can be represented as solutions to the SDE

$$\begin{aligned} &\text{(i)} \quad \mathrm{d}\mathbf{y}^1_t = c\,(x_1 - \mathbf{y}^1_t)\mathrm{d}t + \sqrt{2\alpha\mathbf{y}^1_t(1 - \mathbf{y}^1_t)}\mathrm{d}B^1_t, \\ &\text{(ii)} \quad \mathrm{d}\mathbf{y}^2_t = c\,(x_2 - \mathbf{y}^2_t)\mathrm{d}t + \sqrt{2p(\mathbf{y}^1_t)\mathbf{y}^2_t(1 - \mathbf{y}^2_t)}\mathrm{d}B^2_t. \end{aligned} \tag{1.7}$$

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

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

$$F^{(n+1)}w := F_{c_n}(F^{(n)}w) \quad (n \geq 0) \quad \text{with} \quad F^{(0)}w := w \qquad (w \in \mathcal{W}_{\rm cat}). \tag{1.8}$$

We set $s_0 := 0$ and

$$s_n := \sum_{k=0}^{n-1} \frac{1}{c_k} \qquad (1 \leq n \leq \infty). \tag{1.9}$$

Here is our main result:

Theorem 1.4 (Main result)
(a) *The set* $\mathcal{W}_{\rm cat}$ *is a renormalization class on* $[0,1]^2$ *and* $F_c(\mathcal{W}^{l,r}_{\rm cat}) \subset \mathcal{W}^{l,r}_{\rm cat}$ $(c>0,\ l,r=0,1)$.

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

$$\text{(i)}\quad s_n \underset{n\to\infty}{\longrightarrow} \infty \qquad \textit{and} \qquad \text{(ii)}\quad \frac{s_{n+1}}{s_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^* \tag{1.10}$$

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

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*, \tag{1.11}$$

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

$$\begin{array}{llll} \text{(i)} & (1+\gamma^*)F_{1/\gamma^*}w^* = w^* & & \textit{if}\ \ \gamma^*>0, \\ \text{(ii)} & \frac{1}{2}\sum_{i,j=1}^{2} w^*_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}w^*(x) + w^*(x) = 0 & (x\in[0,1]^2) & \textit{if}\ \ \gamma^*=0. \end{array} \tag{1.12}$$

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

$$p^*_{0,0,\gamma^*} \equiv 0 \quad \textit{and} \quad p^*_{1,1,\gamma^*} \equiv 1 \qquad \textit{for all}\ \gamma^* \geq 0. \tag{1.13}$$

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

Conditions (1.10) (i) and (ii) are satisfied, for example, for $c_k = (1+\gamma^*)^{-k}$. Note that the functions $p^*_{0,0,\gamma^*}$ and $p^*_{1,1,\gamma^*}$ are independent of $\gamma^* \geq 0$. We believe that on the other hand, $p^*_{0,1,\gamma^*}$ is not constant as a function of γ^*, but we have not proved this.

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

$$\tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p(x) + p(x)(1-p(x)) = 0 \qquad (x\in[0,1]) \tag{1.14}$$

with boundary conditions $p(0)=0$ and $p(1)>0$. This function occurred before in the work of Greven, Klenke, and Wakolbinger [GKW01, formulas (1.10)–(1.11)]. In Section 4.1 we discuss the relation between their work and ours.

Outline In Part I of the paper (Sections 1–4) we present our results and our main techniques for proving them. Part II (Sections 5–9) contains detailed proofs. Since the motivation for studying renormalization classes comes from the study of linearly interacting diffusions on the hierarchical group, we explain this connection in Appendix A.

Outline of Part I In the next section, we place our main result in a broader context. We give a more thorough introduction to the theory of renormalization classes on compact sets and discuss earlier results on this topic. In Section 3, we discuss special properties of the renormalization class $\mathcal{W}_{\rm cat}$ from Definition 1.3. In particular, we show how techniques from the theory of spatial branching processes can be used to prove Theorem 1.4. In Section 4 we discuss the relation of our work with that in [GKW01] and mention some open problems.

Notation If E is a separable, locally compact, metrizable space, then $\mathcal{C}(E)$ denotes the space of continuous real functions on E. If E is compact then we equip $\mathcal{C}(E)$ with the supremumnorm

$\|\cdot\|_\infty$. We let $B(E)$ denote the space of all bounded Borel measurable real functions on E. We write $\mathcal{C}_+(E)$ and $\mathcal{C}_{[0,1]}(E)$ for the spaces of all $f \in \mathcal{C}(E)$ with $f \geq 0$ and $0 \leq f \leq 1$, respectively, and define $B_+(E)$ and $B_{[0,1]}(E)$ analogously. We let $\mathcal{M}(E)$ denote the space of all finite measures on E, equipped with the topology of weak convergence. The subspaces of probability measures is denoted by $\mathcal{M}_1(E)$. We write $\mathcal{N}(E)$ for the space of finite counting measures, i.e., measures of the form $\nu = \sum_{i=1}^m \delta_{x_i}$ with $x_1, \ldots, x_m \in E$ $(m \geq 0)$. We interpret ν as a collection of particles, situated at positions $x_1, \ldots, x_m$. For $\mu \in \mathcal{M}(E)$ and $f \in B(E)$ we use the notation $\langle \mu, f \rangle := \int_E f \, \mathrm{d}\mu$ and $|\mu| := \mu(E)$. By definition, $\mathcal{D}_E[0,\infty)$ is the space of cadlag functions $w : [0,\infty) \to E$, equipped with the Skorohod topology. We denote the law of a random variable y by $\mathcal{L}(y)$. If $\mathbf{y} = (\mathbf{y}_t)_{t\geq 0}$ is a Markov process in E and $x \in E$, then P^x denotes the law of $\mathbf{y}$ started in $\mathbf{y}_0 = x$. If μ is a probability law on E then P^μ denotes the law of $\mathbf{y}$ started with initial law $\mathcal{L}(\mathbf{y}_0) = \mu$. For time-inhomogeneous processes, we use the notation $P^{t,x}$ or $P^{t,\mu}$ to denote the law of the process started at time t with initial state $\mathbf{y}_t = x$ or initial law $\mathcal{L}(\mathbf{y}_t) = \mu$, respectively. We let $E^x, E^\mu, \ldots$ etc. denote expectation with respect to $P^x, P^\mu, \ldots$, respectively.

2 Renormalization classes on compact sets

2.1 Some general facts and heuristics

In this section, we explain that our main result is a special case of a type of theorem that we believe holds for many more renormalization classes on compact sets in $\mathbb{R}^d$. Moreover, we describe some elementary properties that hold generally for such renormalization classes. The proofs of Lemmas 2.1–2.8 can be found in Section 5.1 below.

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

Lemma 2.1 (Continuity in parameters)

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

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

In particular, $x \mapsto \nu_x^{c,w}$ is a continuous probability kernel on $\overline{D}$, and $F_c w \in \mathcal{C}(\overline{D}, M_+^d)$ for all $c > 0$ and $w \in \mathcal{W}$. Recall from Definition 1.1 that $\lambda w \in \mathcal{W}$ for all $w \in \mathcal{W}$ and $\lambda > 0$. The reason why we have included this assumption is that it is convenient to have the next scaling lemma around, which is a consequence of time scaling.

Lemma 2.2 (Scaling property of renormalization transformations) *One has*

$$\left.\begin{array}{ll} \text{(i)} & \nu_x^{\lambda c, \lambda w} = \nu_x^{c,w} \\ \text{(ii)} & F_{\lambda c}(\lambda w) = \lambda F_c w \end{array}\right\} \quad (\lambda, c > 0,\ w \in \mathcal{W},\ x \in \overline{D}). \tag{2.1}$$

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

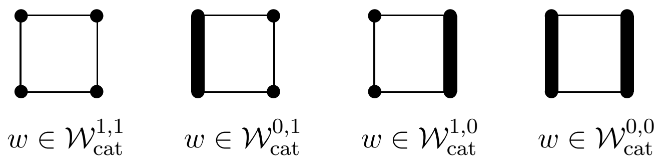


Figure 1: Effective boundaries for $w \in \mathcal{W}_{\text{cat}}$.

Lemma 2.3 (Mean and covariance matrix) *For all $x \in \overline{D}$ and $i, j = 1, \ldots, d$, the mean and covariances of $\nu_x^{c,w}$ are given by*

$$\begin{aligned} &\text{(i)} & \int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)(y_i - x_i) &= 0, \\ &\text{(ii)} & \int_{\overline{D}} \nu_x^{c,w}(\mathrm{d}y)(y_i - x_i)(y_j - x_j) &= \tfrac{1}{c} F_c w_{ij}(x). \end{aligned} \tag{2.2}$$

For any $w \in \mathcal{C}(\overline{D}, M_+^d)$, we call

$$\partial_w D := \{x \in \overline{D} : w_{ij}(x) = 0 \ \forall i, j = 1, \ldots, d\} \tag{2.3}$$

the *effective boundary* of D (associated with w). If $\mathbf{y}$ is a solution to the martingale problem for the operator $\sum_{i,j=1}^d w_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j}$ (i.e., the operator in (1.1) without the drift), then, by martingale convergence, $\mathbf{y}_t$ converges a.s. to a limit $\mathbf{y}_\infty$; it is not hard to see that $\mathbf{y}_\infty \in \partial_w D$ a.s. The next lemma says that the effective boundary is invariant under renormalization.

Lemma 2.4 (Invariance of effective boundary) *One has $\partial_{F_c w} D = \partial_w D$ for all $w \in \mathcal{W}$, $c > 0$.*

For example, for diffusion matrices w from the renormalization class $\mathcal{W} = \mathcal{W}_{\text{cat}}$, there occur four different effective boundaries, depending on whether $w \in \mathcal{W}_{\text{cat}}^{1,1}$, $\mathcal{W}_{\text{cat}}^{0,1}$, $\mathcal{W}_{\text{cat}}^{1,0}$, or $\mathcal{W}_{\text{cat}}^{0,0}$. These effective boundaries are depicted in Figure 1. The statement from Theorem 1.4 (a) that $F_c(\mathcal{W}_{\text{cat}}^{l,r}) \subset \mathcal{W}_{\text{cat}}^{l,r}$ is just the translation of Lemma 2.4 to the special set-up there.

From now on, let $\mathcal{W}$ be a renormalization class, i.e., $\mathcal{W}$ satisfies also condition (iv) from Definition 1.1. Fix a sequence of (positive) migration constants $(c_k)_{k \geq 0}$. By definition, the *iterated probability kernels* $K^{w,(n)}$ associated with a diffusion matrix $w \in \mathcal{W}$ (and the constants $(c_k)_{k \geq 0}$) are the probability kernels on $\overline{D}$ defined inductively by

$$K_x^{w,(n+1)}(\mathrm{d}z) := \int_{\overline{D}} \nu_x^{c_n, F^{(n)} w}(\mathrm{d}y) K_y^{w,(n)}(\mathrm{d}z) \quad (n \geq 0) \quad \text{with} \quad K_x^{w,(0)}(\mathrm{d}y) := \delta_x(\mathrm{d}y), \tag{2.4}$$

with $F^{(n)}$ as in (1.8). Note that

$$F^{(n)} w(x) = \int_{\overline{D}} K_x^{w,(n)}(\mathrm{d}y) w(y) \qquad (x \in \overline{D},\ n \geq 0). \tag{2.5}$$

The next lemma follows by iteration from Lemmas 2.1 and 2.3. It their essence, this lemma and Lemma 2.6 below go back to [BCGdH95].

Lemma 2.5 (Basic properties of iterated probability kernels) *For each $w \in \mathcal{W}$, the $K^{w,(n)}$ are continuous probability kernels on $\overline{D}$. Moreover, for all $x \in \overline{D}$, $i,j = 1,\ldots,d$, and $n \geq 0$, the mean and covariance matrix of $K^{w,(n)}_x$ are given by*

$$\begin{aligned} &\text{(i)} \qquad \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)(y_i - x_i) = 0, \\ &\text{(ii)} \quad \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)(y_i - x_i)(y_j - x_j) = s_n F^{(n)} w_{ij}(x). \end{aligned} \tag{2.6}$$

We equip the space $\mathcal{C}(\overline{D}, \mathcal{M}_1(\overline{D}))$ of continuous probability kernels on $\overline{D}$ with the topology of uniform convergence (since $\mathcal{M}_1(\overline{D})$ is compact, there is a unique uniform structure on $\mathcal{M}_1(\overline{D})$ generating the topology). For 'nice' renormalization classes, it seems reasonable to conjecture that the kernels $K^{w,(n)}$ converge as $n \to \infty$ to some limit $K^{w,*}$ in $\mathcal{C}(\overline{D}, \mathcal{M}_1(\overline{D}))$. If this happens, then formula (2.6) (ii) tells us that the rescaled renormalized diffusion matrices $s_n F^{(n)} w$ converge uniformly on $\overline{D}$ to the covariance matrix of $K^{w,*}$. This gives a heuristic explanation why we need to rescale the iterates $F^{(n)} w$ with the scaling constants s_n from (1.9) to get a nontrivial limit in (1.11).

We now explain the relevance of the conditions (1.10) (i) and (ii) in the present more general context. If the iterated kernels converge to a limit $K^{w,*}$, then condition (1.10) (i) guarantees that this limit is concentrated on the effective boundary:

Lemma 2.6 (Concentration on the effective boundary) *If $s_n \underset{n\to\infty}{\longrightarrow} \infty$, then for any $f \in \mathcal{C}(\overline{D})$ such that $f = 0$ on $\partial_w D$:*

$$\lim_{n\to\infty} \sup_{x\in\overline{D}} \Big| \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y) f(y) \Big| = 0. \tag{2.7}$$

In fact, the condition $s_n \to \infty$ guarantees that the corresponding system of hierarchically interacting diffusions with migration constants $(c_k)_{k\geq 0}$ *clusters in the local mean field limit*, see [DG93a, Theorem 3] or Appendix A.1 below.

To explain also the relevance of condition (1.10) (ii), we observe that using Lemma 2.2, we can convert the rescaled iterates $s_n F^{(n)}$ into (usual, not rescaled) iterates of another transformation. For this purpose, it will be convenient to modify the definition of our scaling constants s_n a little bit. Fix some $\beta > 0$ and put

$$\overline{s}_n := \beta + s_n \qquad (n \geq 0). \tag{2.8}$$

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

$$\overline{F}_\gamma w := (1+\gamma) F_{1/\gamma} w \qquad (\gamma > 0,\ w \in \mathcal{W}). \tag{2.9}$$

Using (2.1) (ii), one easily deduces that

$$\overline{s}_n F^{(n)} w = \overline{F}_{\gamma_{n-1}} \circ \cdots \circ \overline{F}_{\gamma_0}(\beta w) \qquad (w \in \mathcal{W},\ n \geq 1), \tag{2.10}$$

where

$$\gamma_n := \frac{1}{\overline{s}_n c_n} \qquad (n \geq 0). \tag{2.11}$$

We can reformulate the conditions (1.10) (i) and (ii) in terms of the constants $(\gamma_n)_{n\geq 0}$. Indeed, it is not hard to check[1] that equivalent formulations of condition (1.10) (i) are:

$$\text{(i)}\quad s_n \underset{n\to\infty}{\longrightarrow} \infty, \qquad \text{(ii)}\quad \overline{s}_n \underset{n\to\infty}{\longrightarrow} \infty, \qquad \text{(iii)}\quad \sum_n \gamma_n = \infty. \tag{2.12}$$

Since $\overline{s}_{n+1}/\overline{s}_n = 1 + \gamma_n$ we see moreover that, for any $\gamma^* \in [0,\infty]$, equivalent formulations of condition (1.10) (ii) are:

$$\text{(i)}\quad \frac{s_{n+1}}{s_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^*, \qquad \text{(ii)}\quad \frac{\overline{s}_{n+1}}{\overline{s}_n} \underset{n\to\infty}{\longrightarrow} 1+\gamma^*, \qquad \text{(iii)}\quad \gamma_n \underset{n\to\infty}{\longrightarrow} \gamma^*. \tag{2.13}$$

If $0 < \gamma^* < \infty$, then, in the light of (2.10), we expect $\overline{s}_n F^{(n)} w$ to converge to a fixed point of the transformation $\overline{F}_{\gamma^*}$. If $\gamma^* = 0$, the situation is more complex. In this case, we expect the orbit $\overline{s}_n F^{(n)} w \mapsto \overline{s}_{n+1} F^{(n+1)} w \mapsto \cdots$, for large n, to approximate a continuous flow, the generator of which is

$$\lim_{\gamma\to 0} \gamma^{-1}\Big(\overline{F}_\gamma w - w\Big)(x) = \tfrac{1}{2}\sum_{i,j=1}^{d} w_{ij}(x)\tfrac{\partial^2}{\partial x_i \partial x_j} w(x) + w(x) \qquad (x \in \overline{D}). \tag{2.14}$$

To see that the right-hand side of this equation equals the left-hand side if w is twice continuously differentiable, one needs a Taylor expansion of w together with the moment formulas (2.2) for $\nu_x^{1/\gamma,w}$. Under condition condition (2.12) (iii), we expect this continuous flow to reach equilibrium.

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

Conjecture 2.7 (Limits of rescaled renormalized diffusion matrices) *Assume that $s_n \to \infty$ and $s_{n+1}/s_n \to 1 + \gamma^*$ for some $\gamma^* \in [0,\infty]$. Then, for any $w \in \mathcal{W}$,*

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*, \tag{2.15}$$

where w^* satisfies

$$\begin{array}{llll} \text{(i)} & \overline{F}_{\gamma^*} w^* = w^* & & \text{if } 0 < \gamma^* < \infty, \\ \text{(ii)} & \tfrac{1}{2}\sum_{i,j=1}^{d} w^*_{ij}(x)\tfrac{\partial^2}{\partial x_i \partial x_j} w^*(x) + w^*(x) = 0 & (x \in \overline{D}) & \text{if } \gamma^* = 0, \\ \text{(iii)} & \lim_{\gamma\to\infty} \overline{F}_\gamma w^* = w^* & & \text{if } \gamma^* = \infty. \end{array} \tag{2.16}$$

We call (2.16) (ii), which is in some sense the $\gamma^* \to 0$ limit of the fixed point equation (2.16) (i), the *asymptotic fixed point equation.* A version of formula (2.16) (ii) occured in [Swa99, formula (1.3.5)] (a minus sign is missing there).

In particular, one may hope that for a given effective boundary, the equations in (2.16) have a unique solution. Our main result (Theorem 1.4) confirms this conjecture for the renormalization class $\mathcal{W}_{\rm cat}$ and for $\gamma^* < \infty$. In the next section, we discuss numerical evidence

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

that supports Conjecture 2.7 in the case $\gamma^* = 0$ for other renormalization classes on compacta as well.

In previous work on renormalization classes, fixed shapes have played an important role. By definition, for any prerenormalization class $\mathcal{W}$, a *fixed shape* is a subclass $\hat{\mathcal{W}} \subset \mathcal{W}$ of the form $\hat{\mathcal{W}} = \{\lambda w : \lambda > 0\}$ with $0 \neq w \in \mathcal{W}$, such that $F_c(\hat{\mathcal{W}}) \subset \hat{\mathcal{W}}$ for all $c > 0$. The next lemma describes how fixed shapes for renormalization classes on compact sets typically arise.

Lemma 2.8 (Fixed shapes) *Assume that for each $0 < \gamma^* < \infty$, there is a $0 \neq w^* = w^*_{\gamma^*} \in \mathcal{W}$ such that $s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*_{\gamma^*}$ whenever $w \in \mathcal{W}$, $s_n \to \infty$, and $s_{n+1}/s_n \to 1 + \gamma^*$. Then:*

(a) *$w^*_{\gamma^*}$ is the unique solution in $\mathcal{W}$ of equation (2.16) (i).*

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

$$F_c(\lambda w^*) = (\tfrac{1}{\lambda} + \tfrac{1}{c})^{-1} w^* \qquad (\lambda, c > 0). \tag{2.17}$$

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

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

Note that by Theorem 1.4, $\mathcal{W}^{0,1}_{\rm cat}$ is a renormalization class satisfying the general assumptions of Lemma 2.8. The unique solution of (2.16) (i) in $\mathcal{W}^{0,1}_{\rm cat}$ is of the form $w^* = w^{1,p^*}$ where $p^* = p^*_{0,1,\gamma^*}$. We conjecture that the $p^*_{0,1,\gamma^*}$ for different values of γ^* are not constant multiples of each other, and, as a consequence, that $\mathcal{W}^{0,1}_{\rm cat}$ contains no fixed shapes.

Many facts and conjectures that we have discussed can be generalized to renormalization classes on unbounded D, but in this case, the second moments of the iterated kernels $K^{w,(n)}$ may diverge as $n \to \infty$. As a result, because of formula (2.6) (ii), the s_n may no longer be the right scaling factors to find a nontrivial limit of the renormalized diffusion matrices; see, for example, [BCGdH97].

2.2 Numerical solutions to the asymptotic fixed point equation

Let $t \mapsto w(t, \cdot\,)$ be a solution to the continuous flow with the generator in (2.14), i.e., w is an M^d_+-valued solution to the nonlinear partial differential equation

$$\tfrac{\partial}{\partial t} w(t,x) = \tfrac{1}{2} \sum_{i,j=1}^{d} w_{ij}(t,x) \tfrac{\partial^2}{\partial x_i \partial x_j} w(t,x) + w(t,x) \qquad (t \geq 0,\ x \in \overline{D}). \tag{2.18}$$

Solutions to (2.18) are quite easy to simulate on a computer. We have simulated solutions for all kind of diffusion matrices (including nondiagonal ones) on the unit square $[0,1]^2$, with the effective boundaries 1–6 depicted in Figure 2. For all initial diffusion matrices $w(0, \cdot\,)$ we tried, the solution converged as $t \to \infty$ to a fixed point w^*. In all cases except case 6, the fixed point was unique. The fixed points are listed in Figure 2. The functions $p^*_{0,1,0}$ and q^* from Figure 2 are plotted in Figure 3. Here $p^*_{0,1,0}$ is the function from Theorem 1.4 (c).

The fixed points for the effective boundaries in cases 1,2, and 4 are the unique solutions of equation (1.12) (ii) from Theorem 1.4 in the classes $\mathcal{W}^{1,1}_{\rm cat}$, $\mathcal{W}^{0,1}_{\rm cat}$, and $\mathcal{W}^{0,0}_{\rm cat}$, respectively. The simulations suggest that the domain of attraction of these fixed points (within the class of "all" diffusion matrices on $[0,1]^2$) is actually a lot larger than the classes $\mathcal{W}^{1,1}_{\rm cat}$, $\mathcal{W}^{0,1}_{\rm cat}$, and $\mathcal{W}^{0,0}_{\rm cat}$.

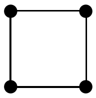
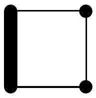
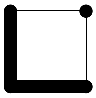
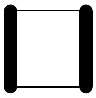
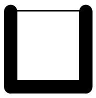
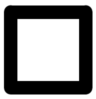
case	effective boundary	fixed points w^* of (2.18)
1		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & x_2(1-x_2) \end{pmatrix}$
2		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & p^*_{0,1,0}(x_1)x_2(1-x_2) \end{pmatrix}$
3		$\begin{pmatrix} q^*(x_1,x_2) & 0 \\ 0 & q^*(x_2,x_1) \end{pmatrix}$
4		$\begin{pmatrix} x_1(1-x_1) & 0 \\ 0 & 0 \end{pmatrix}$
5		$\begin{pmatrix} x_1(1-x_1)1_{\{x_2>0\}} & 0 \\ 0 & 0 \end{pmatrix}$
6		$g^*(x_1,x_2)\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$

Figure 2: Fixed points of the flow (2.18).

The function q^* from case 3 satisfies $q^*(x_1,1) = x_1(1-x_1)$ and is zero on the other parts of the boundary. In contrast to what one might perhaps guess in view of case 2, q^* is *not* of the form $q^*(x_1,x_2) = f(x_2)x_1(1-x_1)$ for some function f.

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

The only case where the fixed point is not unique is case 6. Here, m can be any positive definite matrix, while g^*, depending on m, is the unique solution on $(0,1)^2$ of the equation $1 + \frac{1}{2}\sum_{i,j=1}^{2} m_{ij}\frac{\partial^2}{\partial x_i \partial x_i} g^*(x) = 0$, with zero boundary conditions.

2.3 Previous rigorous results

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

Theorem 2.9 [BCGdH95, DGV95] (Universality class of Wright-Fisher models) *Let $D := \{x \in \mathbb{R}^d : x_i > 0\ \forall i,\ \sum_{i=1}^d x_i < 1\}$, and let $\{e_0,\ldots,e_d\}$, with $e_0 := (0,\ldots,0)$ and $e_1 := (1,0,\ldots,0),\ldots,\ e_d := (0,\ldots,0,1)$ be the extremal points of $\overline{D}$. Let $w^*_{ij}(x) := x_i(\delta_{ij} - x_j)$ ($x \in \overline{D}$, $i,j = 1,\ldots,d$) denote the standard Wright-Fisher diffusion matrix, and assume that $\mathcal{W}$ is a renormalization class on $\overline{D}$ such that $w^* \in \mathcal{W}$ and $\partial_w \overline{D} = \{e_0,\ldots,e_d\}$ for all $w \in \mathcal{W}$. Let $(c_k)_{k\geq 0}$ be migration constants such that $s_n \to \infty$ as $n \to \infty$. Then, for all $w \in \mathcal{W}$, uniformly on $\overline{D}$,*

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*. \tag{2.19}$$

The convergence in (2.19) is a consequence of Lemmas 2.5 and 2.6: The first moment formula (2.6) (i) and (2.7) show that $K_x^{w,(n)}$ converges to the unique distribution on $\{e_0,\ldots,e_d\}$ with

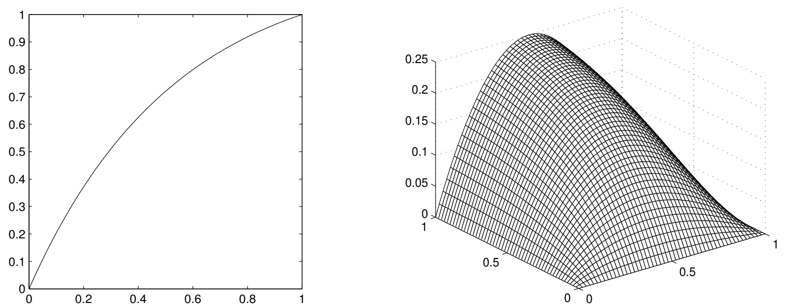


Figure 3: The functions $p^*_{0,1,0}$ and q^* from cases 2 and 3 of Figure 2.

mean x, and by the second moment formula (2.6) (ii) this implies the convergence of $s_n F^{(n)} w$.

In order for the iterates in (2.19) to be well-defined, Theorem 2.9 *assumes* that a renormalization class $\mathcal{W}$ of diffusion matrices w on $\overline{D}$ with effective boundary $\{e_0, \ldots, e_d\}$ is given. The problem of finding a nontrivial example of such a renormalization class is open in dimensions greater than one. In the one-dimensional case, however, the following result is known.

Lemma 2.10 [DG93b] (Renormalization class on the unit interval) *The set*

$$\mathcal{W}_{\mathrm{DG}} := \{w \in \mathcal{C}[0,1] : w = 0 \text{ on } \{0,1\},\ w > 0 \text{ on } (0,1),\ w \text{ Lipschitz}\} \tag{2.20}$$

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

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

Theorem 2.11 [dHS98] (Universality class of isotropic models) *Let $D \subset \mathbb{R}^d$ be open, bounded, and convex and let $m \in M^d_+$ be fixed and (strictly) positive definite. Set $w^*_{ij}(x) := m_{ij} g^*(x)$, where g^* is the unique solution of $1 + \frac{1}{2}\sum_{ij} m_{ij} \frac{\partial^2}{\partial x_i \partial x_j} g^*(x) = 0$ for $x \in D$ and $g^*(x) = 0$ for $x \in \partial D$. Assume that $\mathcal{W}$ is a renormalization class on $\overline{D}$ such that $w^* \in \mathcal{W}$ and such that each $w \in \mathcal{W}$ is of the form*

$$w_{ij}(x) = m_{ij} g(x) \qquad (x \in \overline{D},\ i,j = 1, \ldots, d), \tag{2.21}$$

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

$$s_n F^{(n)} w \underset{n\to\infty}{\longrightarrow} w^*. \tag{2.22}$$

The proof of Theorem 2.11 follows the same lines as the proof of Theorem 2.9, with the difference that in this case one needs to generalize the first moment formula (2.6) (i) in the sense that $\int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y) h(y) = h(x)$ for any m-harmonic function h, i.e., $h \in \mathcal{C}(\overline{D})$ satisfying $\sum_{ij} m_{ij} \frac{\partial^2}{\partial x_i \partial x_j} h(x) = 0$ for $x \in D$. The kernel $K^{w,(n)}_x$ now converges to the m-harmonic measure on ∂D with mean x, and this implies (2.22).

Again, in dimensions $d \geq 2$, the problem of finding a 'reasonable' class $\mathcal{W}$ satisfying the assumptions of Theorem 2.11 is so far unresolved. The problem with verifying conditions (i)–(iv) from Definition 1.1 in an explicit set-up is that (i) and (ii) usually require some smoothness of w, while (iv) requires that one can prove the same smoothness for $F_c w$, which is difficult.

The proofs of Theorems 2.9 and 2.11 are based on the same principle. For any diffusion matrix w, let H_w denote the class of w-harmonic functions, i.e., functions $h \in \mathcal{C}(\overline{D})$ satisfying $\sum_{ij} w_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}h(x) = 0$ on D. If w belongs to one of the renormalization classes in Theorems 2.9 and 2.11, then H_w has the property that $T^c_{x,t}h(H_w) \subset H_w$ for all $c > 0$, $x \in \overline{D}$, and $t \geq 0$, where $T^c_{x,t}h(y) := h(x + (y-x)e^{-ct})$ is the semigroup with generator $\sum_{i=1}^d c(x_i - y_i)\frac{\partial}{\partial y_i}$, i.e., the operator in (1.1) without the diffusion part. In this case we say that w has *invariant harmonics*; see [Swa00]. As a consequence, one can prove that the iterated kernels satisfy $\int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)h(y) = h(x)$ for all $h \in H_w$ and $x \in \overline{D}$. If $s_n \to \infty$, then this implies that $K^{w,(n)}_x$ converges to the unique H_w-harmonic measure on $\partial_w D$ with mean x. Diffusion matrices from $\mathcal{W}_{\rm cat}$ do not in general have invariant harmonics. Therefore, to prove Theorem 1.4, we need new techniques.

Note that in the renormalization classes from Theorems 2.9 and 2.11, the unique attraction point w^* does not depend on γ^*. Therefore, by Lemma 2.8, these renormalization classes contain a unique fixed shape, which is given by $\{\lambda w^* : \lambda > 0\}$.

3 Connection with branching theory

From now on, we focuss on the renormalization class $\mathcal{W}_{\rm cat}$. We will show that for this renormalization class, the rescaled renormalization transformations $\overline{F}_\gamma$ from (2.9) can be expressed in terms of the log-Laplace operators of a discrete time branching process on $[0,1]$. This will allow us to use techniques from the theory of spatial branching processes to verify Conjecture 2.7 for the renormalization class $\mathcal{W}_{\rm cat}$ in the case $\gamma^* < \infty$.

3.1 Poisson-cluster branching processes

We first need some concepts and facts from branching theory. Finite measure-valued branching processes (on $\mathbb{R}$) in discrete time have been introduced by Jiřina [Jir64]. We need to consider only a special class. Let E be a separable, locally compact, and metrizable space. We call a continuous map $\mathcal{Q}$ from E into $\mathcal{M}_1(\mathcal{M}(E))$ a *continuous cluster mechanism*. By definition, an $\mathcal{M}(E)$-valued random variable $\mathcal{X}$ is a *Poisson cluster measure* on E with locally finite *intensity measure* μ and continuous cluster mechanism $\mathcal{Q}$, if its log-Laplace transform satisfies

$$-\log E\big[e^{-\langle \mathcal{X}, f\rangle}\big] = \int_E \mu(\mathrm{d}x)\Big(1 - \int_{\mathcal{M}(E)} \mathcal{Q}(x,\mathrm{d}\chi)e^{-\langle \chi, f\rangle}\Big) \quad (f \in B_+(E)). \tag{3.1}$$

For given μ and $\mathcal{Q}$, such a Poisson cluster measure exists, and is unique in distribution, provided that the right-hand side of (3.1) is finite for $f = 1$. It may be constructed as $\mathcal{X} = \sum_i \chi_{x_i}$, where $\sum_i \delta_{x_i}$ is a (possibly infinite) Poisson point measure with intensity μ, and given $x_1, x_2, \ldots$, the $\chi_{x_1}, \chi_{x_2}, \ldots$ are independent random variables with laws $\mathcal{Q}(x_1, \cdot\,), \mathcal{Q}(x_2, \cdot\,), \ldots$, respectively.

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

$$\mathcal{U}_k f(x) := q_k(x)\Big(1 - \int_{\mathcal{M}(E)} \mathcal{Q}_k(x,\mathrm{d}\chi)e^{-\langle \chi, f\rangle}\Big) \qquad (x \in E,\ f \in B_+(E),\ k = 1, \ldots, n), \tag{3.2}$$

and assume that

$$\sup_{x \in E} \mathcal{U}_k 1(x) < \infty \qquad (k = 1, \ldots, n). \tag{3.3}$$

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

$$E^\mu\big[e^{-\langle\mathcal{X}_n,f\rangle}\big]=e^{-\langle\mu,\mathcal{U}_1\circ\cdots\circ\mathcal{U}_nf\rangle}\qquad(\mu\in\mathcal{M}(E),\ f\in B_+(E)).\tag{3.4}$$

We call $\mathcal{X}=(\mathcal{X}_0,\ldots,\mathcal{X}_n)$ the *Poisson-cluster branching process* on E with *weight functions* $q_1,\ldots,q_n$ and cluster mechanisms $\mathcal{Q}_1,\ldots,\mathcal{Q}_n$. The operator $\mathcal{U}_k$ is called the *log-Laplace operator* of the transition law from $\mathcal{X}_{k-1}$ to $\mathcal{X}_k$. Note that we can write (3.4) in the suggestive form

$$P^\mu\big[\mathrm{Pois}(f\mathcal{X}_n)=0\big]=P\big[\mathrm{Pois}\big((\mathcal{U}_1\circ\cdots\circ\mathcal{U}_nf)\mu\big)=0\big].\tag{3.5}$$

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

3.2 The renormalization branching process

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

By Lemma 5.4 below, for each $\gamma>0$ and $x\in[0,1]$, the SDE

$$\mathrm{d}\mathbf{y}(t)=\tfrac{1}{\gamma}\,(x-\mathbf{y}(t))\mathrm{d}t+\sqrt{2\mathbf{y}(t)(1-\mathbf{y}(t))}\mathrm{d}B(t),\tag{3.6}$$

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

$$\mathcal{Z}^\gamma_x:=\int_0^{\tau_\gamma}\delta_{\mathbf{y}^\gamma_x(-t/2)}\mathrm{d}t\qquad(\gamma>0,\ x\in[0,1]).\tag{3.7}$$

Define constants q_γ and continuous (by Corollary 5.10 below) cluster mechanisms $\mathcal{Q}_\gamma$ by

$$q_\gamma:=\tfrac{1}{\gamma}+1\qquad\text{and}\qquad\mathcal{Q}_\gamma(x,\,\cdot\,):=\mathcal{L}(\mathcal{Z}^\gamma_x)\qquad(\gamma>0,\ x\in[0,1]),\tag{3.8}$$

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

$$\mathcal{U}_\gamma f(x):=q_\gamma\Big(1-\int_{\mathcal{M}([0,1])}\mathcal{Q}_\gamma(x,\mathrm{d}\chi)e^{-\langle\chi,f\rangle}\Big)\qquad(x\in[0,1],\ f\in B_+[0,1],\ \gamma>0).\tag{3.9}$$

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

Proposition 3.1 (Identification of the renormalization transformation) *Let $\overline{F}_\gamma$ be the rescaled renormalization transformation on $\mathcal{W}_{\rm cat}$ defined in (2.9). Then*

$$\overline{F}_\gamma w^{1,p}=w^{1,\mathcal{U}_\gamma p}\qquad(p\in\mathcal{H},\ \gamma>0).\tag{3.10}$$

Fix a diffusion matrix $w^{\alpha,p} \in \mathcal{W}_{\rm cat}$ and migration constants $(c_k)_{k\geq 0}$. Define constants $\overline{s}_n$ and γ_n as in (2.8) and (2.11), respectively, where $\beta := 1/\alpha$. Then Proposition 3.1 and formula (2.10) show that

$$\overline{s}_n F^{(n)} w^{\alpha,p} = w^{1,\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}(\frac{p}{\alpha})}. \tag{3.11}$$

Here $\mathcal{U}_{\gamma_{n-1}}, \ldots, \mathcal{U}_{\gamma_0}$ are the log-Laplace operators of the Poisson-cluster branching process $\mathcal{X} = (\mathcal{X}_{-n}, \ldots, \mathcal{X}_0)$ with weight functions $q_{\gamma_{n-1}}, \ldots, q_{\gamma_0}$ and cluster mechanisms $\mathcal{Q}_{\gamma_{n-1}}, \ldots, \mathcal{Q}_{\gamma_0}$. We call $\mathcal{X}$ (started at some time $-n$ in an initial law $\mathcal{L}(\mathcal{X}_{-n})$) the *renormalization branching process*. By formulas (3.4) and (3.11), the study of the limiting behavior of rescaled iterated renormalization transformations on $\mathcal{W}_{\rm cat}$ reduces to the study of the renormalization branching process $\mathcal{X}$ in the limit $n \to \infty$.

3.3 Convergence to a time-homogeneous process

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

Theorem 3.2 (Convergence to a time-homogenous branching process) *Assume that* $\mathcal{L}(\mathcal{X}_{-n}) \underset{n\to\infty}{\Longrightarrow} \mu$ *for some probability law* μ *on* $\mathcal{M}([0,1])$.

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

$$\mathcal{L}(\mathcal{X}_{-n}, \mathcal{X}_{-n+1}, \ldots) \underset{n\to\infty}{\Longrightarrow} \mathcal{L}(\mathcal{Y}_0^{\gamma^*}, \mathcal{Y}_1^{\gamma^*}, \ldots), \tag{3.12}$$

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

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

$$\mathcal{L}\Big(\big(\mathcal{X}_{-k_n(t)}\big)_{t\geq 0}\Big) \underset{n\to\infty}{\Longrightarrow} \mathcal{L}\Big(\big(\mathcal{Y}_t^0\big)_{t\geq 0}\Big), \tag{3.13}$$

where $\Rightarrow$ *denotes weak convergence of laws on path space,* $k_n(t) := \min\{k : 0 \leq k \leq n, \sum_{l=k}^{n-1} \gamma_l \leq t\}$, *and* $\mathcal{Y}^0$ *is the super-Wright-Fisher diffusion with activity and growth parameter both identically* 1 *and initial law* $\mathcal{L}(\mathcal{Y}_0^0) = \mu$.

The super-Wright-Fisher diffusion was studied in [FS03]. By definition, $\mathcal{Y}^0$ is the time-homogeneous Markov process in $\mathcal{M}[0,1]$ with continuous sample paths, whose Laplace functionals are given by

$$E^\mu\big[e^{-\langle \mathcal{Y}_t^0, f\rangle}\big] = e^{-\langle \mu, \mathcal{U}_t^0 f\rangle} \qquad (\mu \in \mathcal{M}[0,1],\ f \in B_+[0,1],\ t \geq 0). \tag{3.14}$$

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

$$\begin{cases} \frac{\partial}{\partial t} u_t(x) = \frac{1}{2} x(1-x) \frac{\partial^2}{\partial x^2} u_t(x) + u_t(x)(1 - u_t(x)) & (t \geq 0,\ x \in [0,1]), \\ u_0 = f. \end{cases} \tag{3.15}$$

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

3.4 Weighted and Poissonized branching processes

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

Let E again be separable, locally compact, and metrizable. For $\nu \in \mathcal{N}(E)$ and $f \in B_{[0,1]}(E)$, we adopt the notation

$$f^0 := 1 \quad \text{and} \quad f^\nu := \prod_{i=1}^m f(x_i) \quad \text{when} \quad \nu = \sum_{i=1}^m \delta_{x_i} \quad (m \geq 1). \tag{3.16}$$

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

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

$$E^\nu\big[(1-f)^{X_n}\big] = (1 - U_1 \circ \cdots \circ U_n f)^\nu \qquad (\nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)), \tag{3.17}$$

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

$$U_k f(x) := 1 - \int_{\mathcal{N}(E)} Q^k(x, \mathrm{d}\nu)(1-f)^\nu \qquad (1 \leq k \leq n,\ x \in E,\ f \in B_{[0,1]}(E)). \tag{3.18}$$

We call U_k the *generating operator* of the transition law from X_{k-1} to X_k, and we call $X = (X_0, \ldots, X_n)$ the *branching particle system* on E with generating operators $U_1, \ldots, U_n$. It is often useful to write (3.17) in the suggestive form

$$P^\nu\big[\mathrm{Thin}_f(X_n) = 0\big] = P\big[\mathrm{Thin}_{U_1 \circ \cdots \circ U_n f}(\nu) = 0\big] \qquad (\nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)). \tag{3.19}$$

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

$$\mathrm{Thin}_f(\mathrm{Thin}_g(\nu)) \stackrel{\mathcal{D}}{=} \mathrm{Thin}_{fg}(\nu) \quad \text{and} \quad \mathrm{Thin}_f(\mathrm{Pois}(\mu)) \stackrel{\mathcal{D}}{=} \mathrm{Pois}(f\mu), \tag{3.20}$$

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

We are now ready to describe weighted and Poissonized branching processes. Let $\mathcal{X} = (\mathcal{X}_0, \ldots, \mathcal{X}_n)$ be a Poisson-cluster branching process on E, with continuous weight functions $q_1, \ldots, q_n$, continuous cluster mechanisms $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$, and log-Laplace operators $\mathcal{U}_1, \ldots, \mathcal{U}_n$ given by (3.2) and satisfying (3.3). Let $\mathcal{Z}_x^k$ denote an $\mathcal{M}(E)$-valued random variable with law $\mathcal{Q}_k(x, \cdot\,)$. Let $h \in \mathcal{C}_+(E)$ be bounded, $h \neq 0$, and put $E^h := \{x \in E : h(x) > 0\}$. For $f \in B_+(E^h)$, define $hf \in B_+(E)$ by $hf(x) := h(x)f(x)$ if $x \in E^h$ and $hf(x) := 0$ otherwise.

Proposition 3.3 (Weighting of Poisson-cluster branching processes) *Assume that there exists a constant $K < \infty$ such that $\mathcal{U}_k h \leq Kh$ for all $k = 1, \ldots, n$. Then there exists a Poisson-cluster branching process $\mathcal{X}^h = (\mathcal{X}_0^h, \ldots, \mathcal{X}_n^h)$ on E^h with weight functions $(q_1^h, \ldots, q_n^h)$ given by $q_k^h := q_k/h$, continuous cluster mechanisms $\mathcal{Q}_1^h, \ldots, \mathcal{Q}_n^h$ given by*

$$\mathcal{Q}_k^h(x, \cdot\,) := \mathcal{L}(h\mathcal{Z}_x^k) \qquad (x \in E^h), \tag{3.21}$$

and log-Laplace operators $\mathcal{U}_1^h,\ldots,\mathcal{U}_n^h$ *satisfying*

$$h\mathcal{U}_k^h f := \mathcal{U}_k(hf) \qquad (f\in B_+(E^h)). \tag{3.22}$$

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

$$\mathcal{L}(\mathcal{X}_0^h)=\mathcal{L}(h\mathcal{X}_0) \quad \textit{implies} \quad \mathcal{L}(\mathcal{X}_k^h)=\mathcal{L}(h\mathcal{X}_k) \qquad (0\leq k\leq n). \tag{3.23}$$

Proposition 3.4 (Poissonization of Poisson-cluster branching processes) *Assume that* $\mathcal{U}_k h\leq h$ *for all* $k=1,\ldots,n$. *Then there exists a branching particle system* $X^h=(X_0^h,\ldots,X_n^h)$ *on* E^h *with continuous offspring mechanisms* $Q_1^h,\ldots,Q_n^h$ *given by*

$$Q_k^h(x,\cdot\,):=\frac{q_k(x)}{h(x)}P\big[\mathrm{Pois}(h\mathcal{Z}_x^k)\in\cdot\,\big]+\Big(1-\frac{q_k(x)}{h(x)}\Big)\delta_0(\,\cdot\,) \qquad (x\in E^h), \tag{3.24}$$

and generating operators $U_1^h,\ldots,U_n^h$ *satisfying*

$$hU_k^h f := \mathcal{U}_k(hf) \qquad (f\in B_{[0,1]}(E^h)). \tag{3.25}$$

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

$$\mathcal{L}(X_0^h)=\mathcal{L}(\mathrm{Pois}(h\mathcal{X}_0)) \quad \textit{implies} \quad \mathcal{L}(X_k^h)=\mathcal{L}(\mathrm{Pois}(h\mathcal{X}_k)) \qquad (0\leq k\leq n). \tag{3.26}$$

Here, the right-hand side of (3.24) is always a probability measure, despite that it may happen that $q_k(x)/h(x)>1$. The (straightforward) proofs of Propositions 3.3 and 3.4 can be found in Section 7.1 below. If (3.23) holds then we say that $\mathcal{X}^h$ is obtained from $\mathcal{X}$ by *weighting* with density h. If (3.26) holds then we say that X^h is obtained from $\mathcal{X}$ by *Poissonization* with density h. Proposition 3.4 says that a Poisson-cluster branching process $\mathcal{X}$ contains, in a way, certain ‘embedded’ branching particle systems X^h. Poissonization relations for superprocesses and embedded particle systems have enjoyed considerable attention, see [FS04] and references therein.

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

3.5 Extinction versus unbounded growth for embedded particle systems

In this section we explain how embedded particle systems can be used to prove Theorem 1.4. Throughout this section $(\gamma_k)_{k\geq 0}$ are positive constants such that $\sum_n\gamma_n=\infty$ and $\gamma_n\to\gamma^*$ for some $\gamma^*\in[0,\infty)$, and $\mathcal{X}=(\mathcal{X}_{-n},\ldots,\mathcal{X}_0)$ is the renormalization branching process on $[0,1]$ defined in Section 3.2. We write

$$\mathcal{U}^{(n)}:=\mathcal{U}_{\gamma_{n-1}}\circ\cdots\circ\mathcal{U}_{\gamma_0}. \tag{3.27}$$

In view of formula (3.11), in order to prove Theorem 1.4, we need the following result.

Proposition 3.5 (Limits of iterated log-Laplace operators) *Uniformly on* $[0,1]$*,*

$$\begin{array}{lll} \text{(i)} & \lim_{n\to\infty}\mathcal{U}^{(n)}p=1 & (p\in\mathcal{H}_{1,1}),\\ \text{(ii)} & \lim_{n\to\infty}\mathcal{U}^{(n)}p=0 & (p\in\mathcal{H}_{0,0}),\\ \text{(iii)} & \lim_{n\to\infty}\mathcal{U}^{(n)}p=p^*_{0,1,\gamma^*} & (p\in\mathcal{H}_{0,1}), \end{array} \tag{3.28}$$

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

In our proof of Proposition 3.5, we will use embedded particle systems $X^h = (X^h_{-n}, \ldots, X^h_0)$ obtained from $\mathcal{X}$ by Poissonization with certain h taken from the classes $\mathcal{H}_{1,1}$, $\mathcal{H}_{0,0}$, and $\mathcal{H}_{0,1}$.

Lemma 3.6 (Embedded particle system with $h_{1,1}$) *The constant function $h_{1,1}(x) := 1$ is $\mathcal{U}_\gamma$-harmonic for each $\gamma > 0$. The corresponding embedded particle system $X^{h_{1,1}}$ on $[0,1]$ satisfies*

$$P^{-n,\delta_x}\big[|X_0^{h_{1,1}}| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \delta_\infty \tag{3.29}$$

uniformly[2] *for all $x \in [0,1]$.*

In (3.29) and similar formulas below, $\Rightarrow$ denotes weak convergence of probability measures on $[0,\infty]$. Thus, (3.29) says that for processes started with one particle on the position x at times $-n$, the number of particles at time zero converges to infinity as $n \to \infty$.

Lemma 3.7 (Embedded particle system with $h_{0,0}$) *The function $h_{0,0}(x) := x(1-x)$ ($x \in [0,1]$) is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. The corresponding embedded particle system $X^{h_{0,0}}$ on $(0,1)$ is critical and satisfies*

$$P^{-n,\delta_x}\big[|X_0^{h_{0,0}}| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \delta_0 \tag{3.30}$$

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

Here, a branching particle system X is called *critical* if each particle produces on average one offspring (in each time step and independent of its position). Formula (3.30) says that the embedded particle system $X^{h_{0,0}}$ gets extinct during the time interval $\{-n, \ldots, 0\}$ with probability tending to one as $n \to \infty$. We can summarize Lemmas 3.6 and 3.7 by saying that the embedded particle system associated with $h_{1,1}$ grows unboundedly while the embedded particle system associated with $h_{0,0}$ becomes extinct as $n \to \infty$.

We will also consider an embedded particle systems $X^{h_{0,1}}$ for a certain $h_{0,1}$ taken from $\mathcal{H}_{0,1}$. It turns out that this system either gets extinct or grows unboundedly, each with a positive probability. In order to determine these probabilities, we need to consider embedded particle systems for the time-homogeneous processes $\mathcal{Y}^{\gamma^*}$ ($\gamma^* \in [0,\infty)$) from (3.12) and (3.13). If $h \in \mathcal{H}_{0,1}$ is $\mathcal{U}_{\gamma^*}$-superharmonic for some $\gamma^* > 0$, then Poissonizing the process $\mathcal{Y}^{\gamma^*}$ with h yields a branching particle system on $(0,1]$ which we denote by $Y^{\gamma^*,h} = (Y_0^{\gamma^*,h}, Y_1^{\gamma^*,h}, \ldots)$. Likewise, if $h \in \mathcal{H}_{0,1}$ is twice continuously differentiable and satisfies

$$\tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2}h(x) - h(x)(1-h(x)) \leq 0, \tag{3.31}$$

then Poissonizing the super-Wright-Fisher diffusion $\mathcal{Y}^0$ with h yields a continuous-time branching particle system on $(0,1]$, which we denote by $Y^{0,h} = (Y_t^{0,h})_{t\geq 0}$. For example, for $m \geq 4$, the function $h(x) := 1-(1-x)^m$ satisfies (3.31).

Lemma 3.8 (Embedded particle system with $h_{0,1}$) *The function $h_{0,1}(x) := 1-(1-x)^7$ is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. The corresponding embedded particle system $X^{h_{0,1}}$ on $(0,1]$ satisfies*

$$P^{-n,\delta_x}\big[|X_0^{h_{0,1}}| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \rho_{\gamma^*}(x)\delta_\infty + (1-\rho_{\gamma^*}(x))\delta_0, \tag{3.32}$$

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

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

$$\rho_{\gamma^*}(x) := \begin{cases} P^{\delta_x}[Y_k^{\gamma^*, h_{0,1}} \neq 0\ \forall k \geq 0] & (0 < \gamma^* < \infty), \\ P^{\delta_x}[Y_t^{0, h_{0,1}} \neq 0\ \forall t \geq 0] & (\gamma^* = 0). \end{cases} \tag{3.33}$$

We now explain how Lemmas 3.6–3.8 imply Proposition 3.5. In doing so, it will be more convenient to work with weighted branching processes than with Poissonized branching processes. A little argument (which can be found in Lemma 7.12 below) shows that Lemmas 3.6–3.8 are equivalent to the next proposition.

Proposition 3.9 (Extinction versus unbounded growth) *Let $h_{1,1}$, $h_{0,0}$, and $h_{0,1}$ be as in Lemmas 3.6–3.8. For $\gamma^* \in [0,\infty)$, put $p^*_{1,1,\gamma^*}(x) := 1$, $p^*_{0,0,\gamma^*}(x) := 0$ ($x \in [0,1]$), and*

$$p^*_{0,1,\gamma^*}(0) := 0 \quad \textit{and} \quad p^*_{0,1,\gamma^*}(x) := h_{0,1}(x)\rho_{\gamma^*}(x) \qquad (x \in (0,1]), \tag{3.34}$$

with ρ_{γ^*} as in (3.33). Then, for $(l,r) = (1,1), (0,0)$, and $(0,1)$,

$$P^{-n,\delta_x}\big[\langle \mathcal{X}_0, h_{l,r}\rangle \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} e^{-p^*_{l,r,\gamma^*}(x)}\delta_0 + \big(1 - e^{-p^*_{l,r,\gamma^*}(x)}\big)\delta_\infty, \tag{3.35}$$

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

Formula (3.35) says that the weighted branching process $\mathcal{X}^{h_{l,r}}$ exhibits a form of 'extinction versus unbounded growth'. More precisely, for large n the total mass of $h_{l,r}\mathcal{X}_0$ is close to 0 or ∞ with high probability.

Proof of Proposition 3.5 By (3.4),

$$\mathcal{U}^{(n)}p(x) = -\log E^{-n,\delta_x}\big[e^{-\langle \mathcal{X}_0, p\rangle}\big] \qquad (p \in B_+[0,1],\ x \in [0,1]). \tag{3.36}$$

We first prove formula (3.28) (ii). For $(l,r) = (0,0)$, formula (3.35) says that

$$P^{-n,\delta_x}[\langle \mathcal{X}_0, h_{0,0}\rangle \in \cdot\,] \underset{n\to\infty}{\Longrightarrow} \delta_0 \tag{3.37}$$

uniformly for all $x \in [0,1]$. If $p \in \mathcal{H}_{0,0}$, then we can find $r > 0$ such that $p \leq r h_{0,0}$. Therefore, (3.37) implies that for any $p \in \mathcal{H}_{0,0}$,

$$P^{-n,\delta_x}[\langle \mathcal{X}_0, p\rangle \in \cdot\,] \underset{n\to\infty}{\Longrightarrow} \delta_0. \tag{3.38}$$

By (3.36) it follows that

$$\mathcal{U}^{(n)}p(x) = -\log E^{-n,\delta_x}\big[e^{-\langle \mathcal{X}_0, p\rangle}\big] \underset{n\to\infty}{\longrightarrow} 0, \tag{3.39}$$

where the limits in (3.38) and (3.39) are uniform in $x \in [0,1]$. This proves formula (3.28) (ii). To prove formula (3.28) (iii), note that for any $p \in \mathcal{H}_{0,1}$ we can choose $0 < r_- < r_+$ such that $r_- h_{0,1} \leq p + h_{0,0} \leq r_+ h_{0,1}$. Therefore, (3.35) implies that

$$P^{-n,\delta_x}[\langle \mathcal{X}_0, p\rangle + \langle \mathcal{X}_0, h_{0,0}\rangle \in \cdot\,] \underset{n\to\infty}{\Longrightarrow} e^{-p^*_{0,1,\gamma^*}(x)}\delta_0 + \big(1 - e^{-p^*_{0,1,\gamma^*}(x)}\big)\delta_\infty. \tag{3.40}$$

Using moreover (3.37), we see that

$$P^{-n,\delta_x}[\langle\mathcal{X}_0,p\rangle\in\cdot\,]\underset{n\to\infty}{\Longrightarrow}e^{-p^*_{0,1,\gamma^*}(x)}\delta_0+\big(1-e^{-p^*_{0,1,\gamma^*}(x)}\big)\delta_\infty. \tag{3.41}$$

By (3.36), it follows that

$$\mathcal{U}^{(n)}p(x)=-\log E^{-n,\delta_x}\big[e^{-\langle\mathcal{X}_0,p\rangle}\big]\underset{n\to\infty}{\longrightarrow}p^*_{0,1,\gamma^*}(x) \tag{3.42}$$

where all limits are uniform in $x\in[0,1]$. This proves (3.28) (iii). The proof of (3.28) (i) is similar but easier. ∎

4 Discussion, open problems

4.1 Discussion

Consider a $([0,1]^2)^{\mathbb{Z}^2}$-valued process $\mathbf{x}=(\mathbf{x}_\xi)_{\xi\in\mathbb{Z}^2}=(\mathbf{x}^1_\xi,\mathbf{x}^2_\xi)_{\xi\in\mathbb{Z}^2}$, solving a system of SDE's of the form

$$\begin{aligned}\mathrm{d}\mathbf{x}^1_\xi(t)&=\sum_{\eta:\,|\eta-\xi|=1}\big(\mathbf{x}^1_\eta(t)-\mathbf{x}^1_\xi(t)\big)\,\mathrm{d}t+\sqrt{2\alpha\mathbf{x}^1_\xi(t)(1-\mathbf{x}^1_\xi(t))}\,\mathrm{d}B^1_\xi(t),\\ \mathrm{d}\mathbf{x}^2_\xi(t)&=\sum_{\eta:\,|\eta-\xi|=1}\big(\mathbf{x}^2_\eta(t)-\mathbf{x}^2_\xi(t)\big)\,\mathrm{d}t+\sqrt{2p(\mathbf{x}^1_\xi(t))\mathbf{x}^2_\xi(t)(1-\mathbf{x}^2_\xi(t))}\,\mathrm{d}B^2_\xi(t),\end{aligned} \tag{4.1}$$

where $\alpha>0$ is a constant, p is a nonnegative function on $[0,1]$ satisfying $p(0)=0$ and $p(1)>0$, and $(B^i_\xi)^{i=1,2}_{\xi\in\mathbb{Z}^2}$ is a collection of independent Brownian motions. We call $\mathbf{x}$ a system of linearly interacting catalytic Wright-Fisher diffusions with catalyzation function p. It is expected that $\mathbf{x}$ *clusters*, i.e., $\mathbf{x}(t)$ converges in distribution as $t\to\infty$ to a limit $(\mathbf{x}_\xi(\infty))_{\xi\in\mathbb{Z}^2}$ such that $\mathbf{x}_\xi(\infty)=\mathbf{x}_0(\infty)$ for all $\xi\in\mathbb{Z}^2$ and $\mathbf{x}_0(\infty)$ takes values in the effective boundary associated with the diffusion matrix $w^{\alpha,p}$ (see (2.3)). Heuristic arguments, based on renormalization, yield a formula for the *clustering distribution* $\mathcal{L}(\mathbf{x}_0(\infty))$ in terms of the diffusion matrix w^* which is the unique solution of the asymptotic fixed point equation (2.16) (ii) in the renormalization class $\mathcal{W}^{0,1}_{\rm cat}$; see Conjecture A.3 in Appendix A.2 below.

The present paper is inspired by the work of Greven, Klenke and Wakolbinger [GKW01]. They study a model that is closely related to (4.1), but where $\mathbf{x}^1$ is replaced by a voter model. They show that their model clusters and determine its clustering distribution $\mathcal{L}(\mathbf{x}_0(\infty))$, which turns out to coincide with the mentioned prediction for (4.1) based on renormalization theory. In fact, they believe their results to hold for the model in (4.1) too, but they could not prove this due to certain technical difficulties that a $[0,1]$-valued catalyst would create, compared to the simpler $\{0,1\}$-valued voter model.

The work in [GKW01] not only provides the main motivation for the present paper, but also inspired some of our techniques for proving Theorem 1.4. This concerns in particular the proof of Proposition 3.1, which makes the connection between renormalization transformations and a branching process. We hope that conversely, our techniques may shed some light on the problems left open by [GKW01], in particular, the question whether their results stay true if the voter model catalyst is replaced by a Wright-Fisher catalyst. It seems plausible that their results may not hold for the model in (4.1) if the catalyzing function p grows too fast at 0. On the other hand, our proofs suggest that p with a finite slope at 0 should be OK. (In particular, while deriving formula (3.40), we use that p can be bounded from above by $r_+ h_{0,1}$ for some $r_+ > 0$, which requires that p has a finite slope at 0.)

Our results are also interesting in the wider program of studying renormalization classes in the sense of Definition 1.1. We conjecture that the class $\mathcal{W}^{0,1}_{\mathrm{cat}}$, unlike all renormalization classes studied previously, contains no fixed shapes (see the discussion following Lemma 2.8). In fact, we expect this to be the usual situation. In this sense, the renormalization classes studied so far were all of a special type.

4.2 Open problems

The general program of studying renormalization classes in the sense of Definition 1.1 contains a wealth of open problems. In our proofs, we make heavy use of the single-way nature of the catalyzation in (1.7), in particular, the fact that $\mathbf{y}^1$ is an autonomous process which allows one to condition on $\mathbf{y}^1$ and consider $\mathbf{y}^2$ as a process in a random environment created by $\mathbf{y}^1$. As soon as one leaves the single-way catalytic regime one runs into several difficulties, both technically (it is hard to prove that a given class of matrices is a renormalization class in the sense of Definition 1.1) and conceptually (it is not clear when solutions to the asymptotic fixed shape equation (2.16) (ii) are unique). Therefore, it seems at present hard to verify the complete picture for renormalization classes on the unit square that arises from the numerical simulations described in Section 2.2 and Figures 2 and 3, unless one or more essential new ideas are added.

In this context, the study of the nonlinear partial differential equation (2.18) and its fixed points seems to be a challenging problem. This may be a hard problem from an analytic point of view, since the equation is degenerate and not in divergence form. For the renormalization class $\mathcal{W}_{\mathrm{cat}}$, the quasilinear equation (2.18) reduces to the semilinear equation (3.15), which is analytically easier to treat and moreover has a probabilistic interpretation in terms of a superprocess. For a study of the semilinear equation (3.15) we refer to [FS03]. We do not know whether solutions to equation (2.18) can in general be represented in terms of a stochastic process of some sort.

Even for the renormalization class $\mathcal{W}_{\mathrm{cat}}$, several interesting problems are left open. One of the most urgent ones is to prove that the functions $p^*_{0,1,\gamma^*}$ are not constant in γ^*, and therefore, by Lemma 2.8 (c), $\mathcal{W}^{0,1}_{\mathrm{cat}}$ contains no fixed shapes. Moreover, we have not investigated the iterated renormalization transformations in the regime $\gamma^* = \infty$. Also, we believe that the convergence in (3.28) (ii) does not hold if the condition that p is Lipschitz is dropped, in particular, if p has an infinite slope at 0 or an infinite negative slope at 1. For $p \in \mathcal{H}_{0,0}$, it seems plausible that a properly rescaled version of the iterates $\mathcal{U}^{(n)}p$ converges to a universal limit, but we have not investigated this either. Finally, we have not investigated the convergence of the iterated kernels $K^{w,(n)}$ from (2.4) (in particular, we have not verified Conjecture A.2) for the renormalization class $\mathcal{W}_{\mathrm{cat}}$.

Our methods, combined with those in [BCGdH95], can probably be extended to study the action of iterated renormalization transformations on diffusion matrices of the following more general form (compared to (1.4)):

$$w(x) = \begin{pmatrix} g(x_1) & 0 \\ 0 & p(x_1)x_2(1-x_2) \end{pmatrix} \qquad (x =\in [0,1]^2), \tag{4.2}$$

where $g : [0,1] \to \mathbb{R}$ is Lipschitz, $g(0) = g(1) = 0$, $g > 0$ on $(0,1)$, and $p \in \mathcal{H}$ as before. This would, however, require a lot of extra technical work and probably not generate much new insight. The numerical simulations mentioned in Section 2.2 suggest that many diffusion matrices of an even more general form than (4.2) also converge under renormalization to the limit points w^* from Theorem 1.4, but we don't know how to prove this.

Part II

Outline of Part II In Section 5, we verify that $\mathcal{W}_{\rm cat}$ is a renormalization class, we prove Proposition 3.1, which connects the renormalization transformations F_c to the log-Laplace operators $\mathcal{U}_\gamma$, and we collect a number of technical properties of the operators $\mathcal{U}_\gamma$ that will be needed later on. In Section 6 we prove Theorem 3.2 about the convergence of the renormalization branching process to a time-homogeneous limit. In Section 7, we prove the statements from Section 3.5 about extinction versus unbounded growth of embedded particle systems, with the exception of Lemma 3.7, which is proved in Section 8. In Section 9, finally, we combine the results derived by that point to prove our main theorem.

5 The renormalization class $\mathcal{W}_{\rm cat}$

In this section we prove Theorem 1.4 (a) and Proposition 3.1, as well as Lemmas 2.1–2.8 from Section 2. The section is organized according to the techniques used. Section 5.1 collects some facts that hold for general renormalization classes on compact sets. In Section 5.2 we use the SDE (1.7) to couple catalytic Wright-Fisher diffusions. In Section 5.3 we apply the moment duality for the Wright-Fisher diffusion to the catalyst and to the reactant conditioned on the catalyst. In Section 5.4 we prove that monotone concave catalyzing functions form a preserved class under renormalization.

5.1 Renormalization classes on compact sets

In this section, we prove the lemmas stated in Section 2. Recall that $D \subset \mathbb{R}^d$ is open, bounded, and convex, and that $\mathcal{W}$ is a prerenormalization class on $\overline{D}$, equipped with the topology of uniform convergence.

Proof of Lemma 2.1 To see that $(x, c, w) \mapsto \nu_x^{c,w}$ is continuous, let (x_n, c_n, w_n) be a sequence converging in $\overline{D} \times (0, \infty) \times \mathcal{W}$ to a limit (x, c, w). By the compactness of $\overline{D}$, the sequence $(\nu_{x_n}^{c_n,w_n})_{n\geq 0}$ is tight, and each limit point ν^* satisfies

$$\langle \nu^*, A_x^{c,w} f\rangle = 0 \qquad (f \in \mathcal{C}^{(2)}(D)). \tag{5.1}$$

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

Proof of Lemma 2.2 Formula (2.1) (i) follows from the fact that rescaling the time in solutions $(\mathbf{y}_t)_{t\geq 0}$ to the martingale problem for $A_x^{c,w}$ by a factor λ has no influence on the invariant law. Formula (2.1) (ii) is a direct consequence of formula (2.1) (i). ∎

Proof of Lemma 2.3 This follows by inserting the functions $f(x) = x_i$ and $f(x) = x_i x_j$ into the equilibrium equation (5.1). ∎

Proof of Lemma 2.4 If $x \in \partial_w D$, then $\mathbf{y}_t := x$ $(t \geq 0)$ is a stationary solution to the martingale problem for $A_x^{c,w}$, and therefore $\nu_x^{c,w} = \delta_x$ and $F_c w(x) = w(x) = 0$. On the other hand, if $x \notin \partial_w D$, then $\mathbf{y}_t := x$ $(t \geq 0)$ is not a stationary solution to the martingale problem for $A^{c,w}_x$ and therefore $\int_{\overline{D}} \nu^{c,w}_x(\mathrm{d}y)|y-x|^2 > 0$. Let $\mathrm{tr}(w(y)) := \sum_i w_{ii}(y)$ denote the trace of $w(y)$. By (2.2) (ii), $\frac{1}{c}\mathrm{tr}(F_c w)(x) = \frac{1}{c}\int_{\overline{D}} \nu^{c,w}_x(\mathrm{d}y)\mathrm{tr}(w(y)) = \int_{\overline{D}} \nu^{c,w}_x(\mathrm{d}y)|y-x|^2 > 0$ and therefore $F_c w(x) \neq 0$. ∎

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

$$K^{w,(n)} = \nu^{c_{n-1},F^{(n-1)}w} \cdots \nu^{c_0,w} \qquad (n \geq 1), \tag{5.2}$$

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

$$(KL)_x(\mathrm{d}z) := \int_{\overline{D}} K_x(\mathrm{d}y) L_y(\mathrm{d}z). \tag{5.3}$$

Proof of Lemma 2.5 This is a direct consequence of Lemmas 2.1 and 2.3. In particular, the relations (2.6) follow by iterating the relations (2.2). ∎

Proof of Lemma 2.6 Recall that $\mathrm{tr}(w(y))$ denotes the trace of $w(y)$. Formulas (2.5) and (2.6) (ii) show that

$$\int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)\,|y-x|^2 = s_n \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)\,\mathrm{tr}(w(y)). \tag{5.4}$$

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

$$\lim_{n\to\infty} \sup_{x\in D} \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)\mathrm{tr}(w(y)) = 0. \tag{5.5}$$

Since w is symmetric and nonnegative definite, $\mathrm{tr}(w(y))$ is nonnegative, and zero if and only if $y \in \partial_w D$. If $f \in \mathcal{C}(\overline{D})$ satisfies $f = 0$ on $\partial_w D$, then, for every $\varepsilon > 0$, the sets $C_m := \{x \in \overline{D} : |f(x)| \geq \varepsilon + m\,\mathrm{tr}(w(x))\}$ are compact with $C_m \downarrow \emptyset$ as $m \uparrow \infty$, so there exists an m (depending on ε) such that $|f| < \varepsilon + m\,\mathrm{tr}(w)$. Therefore,

$$\begin{aligned}\limsup_{n\to\infty} \sup_{x\in\overline{D}} \Big| \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y) f(y) \Big| &\leq \limsup_{n\to\infty} \sup_{x\in\overline{D}} \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)|f(y)| \\ &\leq \varepsilon + m \limsup_{n\to\infty} \sup_{x\in\overline{D}} \int_{\overline{D}} K^{w,(n)}_x(\mathrm{d}y)\mathrm{tr}(w(y)) = \varepsilon.\end{aligned} \tag{5.6}$$

Since $\varepsilon > 0$ is arbitrary, (2.7) follows. ∎

Proof of Lemma 2.8 By (2.10), (2.12), and (2.13), $w^*_{\gamma^*} = \lim_{n\to\infty}(\overline{F}_{\gamma^*})^n w$ for each $w \in \mathcal{W}$. By Lemma 2.1 (b), $\overline{F}_{\gamma^*} : \mathcal{W} \to \mathcal{W}$ is continuous, so $w^*_{\gamma^*}$ is the unique fixed point of $\overline{F}_{\gamma^*}$. This proves part (a).

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

To complete the proof of part (b), note that if $w^* = w^*_{\gamma^*}$ does not depend on γ^*, then $w^* \in \mathcal{W}$ solves (2.16) (i) for all $0 < \gamma^* < \infty$, hence $F_c w^* = (1+\frac{1}{c})^{-1} w^*$ for all $c > 0$, and therefore, by scaling (Lemma 2.2), $F_c(\lambda w^*) = \lambda F_{c/\lambda}(w^*) = \lambda(1+\frac{\lambda}{c})^{-1} w^* = (\frac{1}{\lambda}+\frac{1}{c})^{-1} w^*$. ∎

5.2 Coupling of catalytic Wright-Fisher diffusions

In this section we verify condition (i) of Definition 1.1 for the class $\mathcal{W}_{\rm cat}$, and we prepare for the verification of conditions (ii)–(iv) in Section 5.3. In fact, we will show that the larger class $\overline{\mathcal{W}}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{C}_+[0,1]\}$ is also a renormalization class, and the equivalents of Theorem 1.4 (a) and Proposition 3.1 remain true for this larger class. (We do not know, however, if the convergence statements in Theorem 1.4 (b) also hold in this larger class; see the discussion in Section 4.2.)

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

We are therefore left with the task of verifying uniqueness of solutions to the martingale problem for $A^{c,w}_x$. By [EK86, Problem 4.19, Corollary 5.3.4, and Theorem 5.3.6], it suffices to show that solutions to (1.7) are pathwise unique.

Lemma 5.1 (Monotone coupling of Wright-Fisher diffusions) *Assume that $0 \leq x \leq \tilde x \leq 1$, $c \geq 0$ and that $(P_t)_{t\geq 0}$ is a progressively measurable, nonnegative process such that $\sup_{t\geq 0,\omega\in\Omega} P_t(\omega) < \infty$. Let $\mathbf{y}, \tilde{\mathbf{y}}$ be $[0,1]$-valued solutions to the SDE's*

$$\begin{aligned} \mathrm{d}\mathbf{y}_t &= c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2P_t\mathbf{y}_t(1-\mathbf{y}_t)}\mathrm{d}B_t,\\ \mathrm{d}\tilde{\mathbf{y}}_t &= c\,(\tilde x - \tilde{\mathbf{y}}_t)\mathrm{d}t + \sqrt{2P_t\tilde{\mathbf{y}}_t(1-\tilde{\mathbf{y}}_t)}\mathrm{d}B_t, \end{aligned} \tag{5.7}$$

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

$$\mathbf{y}_t \leq \tilde{\mathbf{y}}_t \quad \forall t \geq 0 \quad a.s. \tag{5.8}$$

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

$$0 \leq \rho_n(x) \leq \frac{1}{nx}1_{(0,1]}(x) \qquad\qquad (x \geq 0). \tag{5.9}$$

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

$$\phi_n(x) := \int_0^{x\vee 0}\mathrm{d}y\int_0^y \mathrm{d}z\,\rho_n(z). \tag{5.10}$$

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

$$\begin{aligned} E[\phi_n(\mathbf{y}_t - \tilde{\mathbf{y}}_t)] &= E[\phi_n(\mathbf{y}_0 - \tilde{\mathbf{y}}_0)] & \text{(i)}\\ &\quad + c\,(x-\tilde x)\int_0^t E[\phi_n'(\mathbf{y}_s - \tilde{\mathbf{y}}_s)]\mathrm{d}s - c\int_0^t E[(\mathbf{y}_s - \tilde{\mathbf{y}}_s)\phi_n'(\mathbf{y}_s - \tilde{\mathbf{y}}_s)]\mathrm{d}s & \text{(ii)}\\ &\quad + \int_0^t E\Big[P_s\Big(\sqrt{\mathbf{y}_s(1-\mathbf{y}_s)} - \sqrt{\tilde{\mathbf{y}}_s(1-\tilde{\mathbf{y}}_s)}\Big)^2\phi_n''(\mathbf{y}_s - \tilde{\mathbf{y}}_s)\Big]\mathrm{d}s. & \text{(iii)} \end{aligned} \tag{5.11}$$

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

$$|\sqrt{y(1-y)} - \sqrt{\tilde y(1-\tilde y)}| \leq |y - \tilde y|^{\frac{1}{2}} \qquad (y, \tilde y \in [0,1]), \tag{5.12}$$

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

$$E[0 \vee (\mathbf{y}_t - \tilde{\mathbf{y}}_t)] \leq E[0 \vee (\mathbf{y}_0 - \tilde{\mathbf{y}}_0)] = 0, \tag{5.13}$$

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

Corollary 5.2 (Pathwise uniqueness) *For all $c \geq 0$, $\alpha > 0$, $p \in \mathcal{C}_+[0,1]$ and $x \in [0,1]$, solutions to the SDE (1.7) are pathwise unique.*

Proof Let $(\mathbf{y}^1, \mathbf{y}^2)$ and $(\tilde{\mathbf{y}}^1, \tilde{\mathbf{y}}^2)$ be solutions to (1.7) relative to the same pair (B^1, B^2) of Brownian motions, with $(\mathbf{y}_0^1, \mathbf{y}_0^2) = (\tilde{\mathbf{y}}_0^1, \tilde{\mathbf{y}}_0^2)$. Applying Lemma 5.1, with inequality in both directions, we see that $\mathbf{y}^1 = \tilde{\mathbf{y}}^1$ a.s. Applying Lemma 5.1 two more times, this time using that $\mathbf{y}^1 = \tilde{\mathbf{y}}^1$ a.s., we see that also $\mathbf{y}^2 = \tilde{\mathbf{y}}^2$ a.s. ∎

Corollary 5.3 (Exponential coupling) *Assume that $x \in [0,1]$, $c \geq 0$, and $\alpha > 0$. Let $\mathbf{y}, \tilde{\mathbf{y}}$ be solutions to the SDE*

$$\mathrm{d}\mathbf{y}_t = c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2\alpha \mathbf{y}_t(1 - \mathbf{y}_t)}\mathrm{d}B_t, \tag{5.14}$$

relative to the same Brownian motion B. Then

$$E\big[|\tilde{\mathbf{y}}_t - \mathbf{y}_t|\big] = e^{-ct} E\big[|\tilde{\mathbf{y}}_0 - \mathbf{y}_0|\big]. \tag{5.15}$$

Proof If $\mathbf{y}_0 = y$ and $\tilde{\mathbf{y}}_0 = \tilde{y}$ are deterministic and $y \leq \tilde{y}$, then by Lemma 5.1 and a simple moment calculation

$$E\big[|\tilde{\mathbf{y}}_t - \mathbf{y}_t|\big] = E[\tilde{\mathbf{y}}_t - \mathbf{y}_t] = e^{-ct}|\tilde{y} - y|. \tag{5.16}$$

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

Corollary 5.4 (Ergodicity) *The Markov process defined by the SDE (3.6) has a unique invariant law Γ_x^γ and is ergodic, i.e, solutions to (3.6) started in an arbitrary initial law $\mathcal{L}(\mathbf{y}_0)$ satisfy $\mathcal{L}(\mathbf{y}_t) \underset{t\to\infty}{\Longrightarrow} \Gamma_x^\gamma$.*

Proof Since our process is a Feller diffusion on a compactum, the existence of an invariant law follows from a simple time averaging argument. Now start one solution $\tilde{\mathbf{y}}$ of (3.6) in this invariant law and let $\mathbf{y}$ be any other solution, relative to the same Brownian motion. Corollary 5.3 then gives ergodicity and, in particular, uniqueness of the invariant law. ∎

Remark 5.5 (Density of invariant law) It is well-known (see, for example [Ewe04, formula (5.70)]) that Γ_x^γ is a $\beta(\alpha_1, \alpha_2)$-distribution, where $\alpha_1 := x/\gamma$ and $\alpha_2 := (1-x)/\gamma$, i.e., $\Gamma_x^\gamma = \delta_x$ $(x \in \{0,1\})$ and

$$\Gamma_x^\gamma(\mathrm{d}y) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\, y^{\alpha_1 - 1}(1-y)^{\alpha_2 - 1}\mathrm{d}y \qquad (x \in (0,1)). \tag{5.17}$$

◇

We conclude this section with a lemma that prepares for the verification of condition (iv) in Definition 1.1 for the class $\mathcal{W}_{\rm cat}$.

Lemma 5.6 (Monotone coupling of stationary Wright-Fisher diffusions) *Assume that $c > 0$, $\alpha > 0$ and $0 \leq x \leq \tilde{x} \leq 1$. Then the pair of equations*

$$\begin{aligned} \mathrm{d}\mathbf{y}_t &= c\,(x - \mathbf{y}_t)\mathrm{d}t + \sqrt{2\alpha\mathbf{y}_t(1-\mathbf{y}_t)}\mathrm{d}B_t, \\ \mathrm{d}\tilde{\mathbf{y}}_t &= c\,(\tilde{x} - \tilde{\mathbf{y}}_t)\mathrm{d}t + \sqrt{2\alpha\tilde{\mathbf{y}}_t(1-\tilde{\mathbf{y}}_t)}\mathrm{d}B_t \end{aligned} \tag{5.18}$$

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

$$\mathbf{y}_t \leq \tilde{\mathbf{y}}_t \quad \forall t \in \mathbb{R} \quad a.s. \tag{5.19}$$

Proof Let $(\mathbf{y}_t, \tilde{\mathbf{y}}_t)_{t\geq 0}$ be a solution of (5.18) and let $(\mathbf{y}'_t, \tilde{\mathbf{y}}'_t)_{t\geq 0}$ be another one, relative to the same Brownian motion B. Then, by Lemma 5.3, $E[|\mathbf{y}_t - \mathbf{y}'_t|] \to 0$ and also $E[|\tilde{\mathbf{y}}_t - \tilde{\mathbf{y}}'_t|] \to 0$ as $t \to \infty$. Hence we may argue as in the proof of Corollary 5.4 that (5.18) has a unique invariant law and is ergodic. Now start a solution of (5.18) in an initial condition such that $\mathbf{y}_0 \leq \tilde{\mathbf{y}}_0$. By ergodicity, the law of this solution converges as $t \to \infty$ to the invariant law of (5.18) and using Lemma 5.1 we see that this invariant law is concentrated on $\{(y, \tilde{y}) \in [0,1]^2 : y \leq \tilde{y}\}$. Now consider, on the whole real time axis, the stationary solution to (5.18) with this invariant law. Applying Lemma 5.1 once more, we see that (5.19) holds. ∎

5.3 Duality for catalytic Wright-Fisher diffusions

In this section we prove Theorem 1.4 (a) and Proposition 3.1. Moreover, we will show that their statements remain true if the renormalization class $\mathcal{W}_{\rm cat}$ is replaced by the larger class $\overline{\mathcal{W}}_{\rm cat} := \{w^{\alpha,p} : \alpha > 0,\ p \in \mathcal{C}_+[0,1]\}$. We begin by recalling the usual moment duality for Wright-Fisher diffusions.

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

$$\mathrm{d}\mathbf{y}(t) = \tfrac{1}{\gamma}\,(x - \mathbf{y}(t))\mathrm{d}t + \sqrt{2\mathbf{y}(t)(1-\mathbf{y}(t))}\mathrm{d}B(t), \tag{5.20}$$

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

$$\begin{aligned} (\phi_t, \psi_t) &\to (\phi_t - 1, \psi_t) && \text{with rate } \phi_t(\phi_t - 1) \\ (\phi_t, \psi_t) &\to (\phi_t - 1, \psi_t + 1) && \text{with rate } \tfrac{1}{\gamma}\phi_t. \end{aligned} \tag{5.21}$$

Then one has the following *duality relation* (see for example Lemma 2.3 in [Shi80] or Proposition 1.5 in [GKW01])

$$E^y\big[\mathbf{y}_t^n x^m\big] = E^{(n,m)}\big[y^{\phi_t} x^{\psi_t}\big] \qquad (y \in [0,1],\ (n,m) \in \mathbb{N}^2), \tag{5.22}$$

where $0^0 := 1$. The duality in (5.22) has the following heuristic explanation. Consider a population containing a fixed, large number of organisms, that come in two genetic types, say I and II. Each pair of organisms in the population is *resampled* with rate 2. This means that one organism of the pair (chosen at random) dies, while the other organism produces one child of its own genetic type. Moreover, each organism is replaced with rate $\frac{1}{\gamma}$ by an organism chosen from an infinite reservoir where the frequency of type I has the fixed value x. In the limit that the number of organisms in the population is large, the relative frequency $\mathbf{y}_t$ of type

I organisms follows the SDE (5.20). Now $E[\mathbf{y}_t^n]$ is the probability that n organisms sampled from the population at time t are all of type I. In order to find this probability, we follow the ancestors of these organisms back in time. Viewed backwards in time, these ancestors live for a while in the population, until, with rate $\frac{1}{\gamma}$, they jump to the infinite reservoir. Moreover, due to resampling, each pair of ancestors coalesces with rate 2 to one common ancestor. Denoting the number of ancestors that lived at time $t-s$ in the population and in the reservoir by ϕ_s and ψ_s, respectively, we see that the probability that all ancestors are of type I is $E^y[\mathbf{y}_t^n] = E^{(n,0)}[y^{\phi_t} x^{\psi_t}]$. This gives a heuristic explanation of (5.22).

Since eventually all ancestors of the process (ϕ, ψ) end up in the reservoir, we have $(\phi_t, \psi_t) \to (0, \psi_\infty)$ as $t \to \infty$ a.s. for some $\mathbb{N}$-valued random variable ψ_∞. Taking the limit $t \to \infty$ in (5.22), we see that the moments of the invariant law Γ_x^γ from Corollary 5.4 are given by:

$$\int \Gamma_x^\gamma(\mathrm{d}y) y^n = E^{(n,0)}[x^{\psi_\infty}] \qquad (n \geq 0). \tag{5.23}$$

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

$$\int \Gamma_x^\gamma(\mathrm{d}y) y^n = \prod_{k=0}^{n-1} \frac{x + k\gamma}{1 + k\gamma} \qquad (n \geq 1). \tag{5.24}$$

In particular, it follows that

$$\int \Gamma_x^\gamma(\mathrm{d}y) y(1-y) = \frac{1}{1+\gamma} x(1-x). \tag{5.25}$$

This is the important *fixed shape property* of the Wright-Fisher diffusion (see formula (2.17)).

We now consider catalytic Wright-Fisher diffusions $(\mathbf{y}^1, \mathbf{y}^2)$ as in (1.7) with $p \in \mathcal{C}_+[0,1]$ and apply duality to the catalyst $\mathbf{y}^2$ conditioned on the reactant $\mathbf{y}^1$. Let $(\mathbf{y}_t^1, \mathbf{y}_t^2)_{t \in \mathbb{R}}$ be a stationary solution to the SDE (1.7) with $c = 1/\gamma$. Let $(\tilde\phi, \tilde\psi)$ be a $\mathbb{N}^2$-valued process, defined on the same probability space as $(\mathbf{y}^1, \mathbf{y}^2)$, such that conditioned on the past path $(\mathbf{y}_{-t}^1)_{t \leq 0}$, the process $(\tilde\phi, \tilde\psi)$ is a (time-inhomogeneous) Markov process that jumps as:

$$\begin{array}{lll} (\tilde\phi_t, \tilde\psi_t) \to (\tilde\phi_t - 1, \tilde\psi_t) & \quad \text{with rate} & p(\mathbf{y}_{-t}^1)\tilde\phi_t(\tilde\phi_t - 1), \\ (\tilde\phi_t, \tilde\psi_t) \to (\tilde\phi_t - 1, \tilde\psi_t + 1) & \quad \text{with rate} & \frac{1}{\gamma}\tilde\phi_t. \end{array} \tag{5.26}$$

Then, in analogy with (5.22),

$$E[(\mathbf{y}_0^2)^n x_2^m | (\mathbf{y}_{-t}^1)_{t \leq 0}] = E^{(n,m)}[(\mathbf{y}_{-t}^2)^{\tilde\phi_t} x_2^{\tilde\psi_t} | (\mathbf{y}_{-t}^1)_{t \leq 0}] \qquad ((n,m) \in \mathbb{N}^2,\ t \geq 0). \tag{5.27}$$

We may interpret (5.26) by saying that pairs of ancestors in a finite population coalesce with time-dependent rate $2p(\mathbf{y}_{-t}^1)$ and ancestors jump to an infinite reservoir with constant rate $\frac{1}{\gamma}$. Again, eventualy all ancestors end up in the reservoir, and therefore $(\tilde\phi_t, \tilde\psi_t) \to (0, \tilde\psi_\infty)$ as $t \to \infty$ a.s. for some $\mathbb{N}$-valued random variable $\tilde\psi_\infty$. Taking the limit $t \to \infty$ in (5.27) we find that

$$E[(\mathbf{y}_0^2)^n x_2^m | (\mathbf{y}_{-t}^1)_{t \leq 0}] = E^{(n,m)}[x_2^{\tilde\psi_\infty} | (\mathbf{y}_{-t}^1)_{t \leq 0}] \qquad ((n,m) \in \mathbb{N}^2,\ t \geq 0). \tag{5.28}$$

Lemma 5.7 (Uniqueness of invariant law) *For each $c > 0$, $w \in \overline{\mathcal{W}}_{\mathrm{cat}}$, and $x \in [0,1]^2$, there exists a unique invariant law $\nu_x^{c,w}$ for the martingale problem for $A_x^{c,w}$.*

Proof Our process being a Feller diffusion on a compactum, the existence of an invariant law follows from time averaging. We need to show uniqueness. If $(\mathbf{y}^1,\mathbf{y}^2)=\mathbf{y}^1_t,\mathbf{y}^2_t)_{t\in\mathbb{R}}$ is a stationary solution, then $\mathbf{y}^1$ is an autonomous process, and $\mathcal{L}(\mathbf{y}^1_0)=\Gamma^{1/c}_x$, the unique invariant law from Corollary 5.4. Therefore, $\mathcal{L}((\mathbf{y}^1_t)_{t\in\mathbb{R}})$ is determined uniquely by the requirement that $(\mathbf{y}^1,\mathbf{y}^2)$ be stationary. By (5.28), the conditional distribution of $\mathbf{y}^2_0$ given $(\mathbf{y}^1_t)_{t\leq 0}$ is determined uniquely, and therefore the joint distribution of $\mathbf{y}^2_0$ and $(\mathbf{y}^1_t)_{t\leq 0}$ is determined uniquely. In particular, $\mathcal{L}(\mathbf{y}^1_0,\mathbf{y}^2_0)=\nu^{c,w}_x$ is determined uniquely. ∎

Remark 5.8 (Reversibility) It seems that the invariant law $\nu^{c,w}_x$ from Lemma 5.7 is reversible. In many cases (densities of) reversible invariant measures can be obtained in closed form by solving the equations of detailed balance. This is the case, for example, for the one-dimensional Wright-Fisher diffusion. We have not attempted this for the catalytic Wright-Fisher diffusion. ◇

The next proposition implies Proposition 3.1 and prepares for the proof of Theorem 1.4 (a).

Proposition 5.9 (Extended renormalization class) *The set $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class on $[0,1]^2$, and*

$$\overline{F}_\gamma w^{1,p}=w^{1,\mathcal{U}_\gamma p}\qquad (p\in\mathcal{C}_+[0,1],\ \gamma>0). \tag{5.29}$$

Proof To see that $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class we need to check conditions (i)–(iv) from Definition 1.1. By Lemma 5.2, the martingale problem for $A^{c,w}_x$ is well-posed for all $c\geq 0$, $w\in\mathcal{W}_{\rm cat}$ and $x\in[0,1]^2$. By Lemma 5.7, the corresponding Feller process on $[0,1]^2$ has a unique invariant law $\nu^{c,w}_x$. This shows that conditions (i) and (ii) from Definition 1.1 are satisfied. Note that by the compactness of $[0,1]^2$, any continuous function on $[0,1]^2$ is bounded, so condition (iii) is automatically satisfied. Hence $\mathcal{W}$ is a prerenormalization class. As a consequence, for any $p\in\mathcal{C}_+[0,1]$, $\overline{F}_\gamma w^{1,p}$ is well-defined by (1.2) and (2.9). We will now first prove (5.29) and then show that $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class.

Fix $\gamma>0$, $p\in\mathcal{C}_+[0,1]$, and $x\in[0,1]^2$. Let $(\mathbf{y}^1_t,\mathbf{y}^2_t)_{t\in\mathbb{R}}$ be a stationary solution to the SDE (1.7) with $\alpha=1$ and $c=1/\gamma$. Then

$$\overline{F}_\gamma w^{1,p}_{ij}(x)=(1+\gamma)E[w^{1,p}_{ij}(\mathbf{y}^1_0,\mathbf{y}^2_0)]\qquad (i,j=1,2). \tag{5.30}$$

Since $w^{1,p}_{ij}=0$ if $i\neq j$, it is clear that $\overline{F}_\gamma w^{1,p}_{ij}(x)=0$ if $i\neq j$. Since $\mathcal{L}(\mathbf{y}^1_0)=\Gamma^\gamma_x$ it follows from (5.25) that $\overline{F}_\gamma w^{1,p}_{11}(x)=x_1(1-x_1)$. We are left with the task of showing that

$$\overline{F}_\gamma w^{1,p}_{22}(x)=\mathcal{U}_\gamma p(x_1)x_2(1-x_2). \tag{5.31}$$

Here, by (2.2) (ii),

$$\begin{aligned}\overline{F}_\gamma w^{1,p}_{22}(x)&=(1+\gamma)E[p(\mathbf{y}^1_0)\mathbf{y}^2_0(1-\mathbf{y}^2_0)]\\ &=(\tfrac{1}{\gamma}+1)E[(\mathbf{y}^2_0-x_2)^2].\end{aligned} \tag{5.32}$$

By (5.28), using the fact that $E[\mathbf{y}^2_0]=x_2$ (which follows from (5.27) or more elementary from (2.6) (i)), we find that

$$E[(\mathbf{y}^2_0-x_2)^2]=E[(\mathbf{y}^2_0)^2]-(x_2)^2=E^{(2,0)}[x_2^{\tilde\psi_\infty}]-(x_2)^2=P^{(2,0)}[\tilde\psi_\infty=1]x_2(1-x_2)\qquad (t\geq 0). \tag{5.33}$$

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

$$P^{(2,0)}[\tilde{\psi}_\infty = 2] = E\big[e^{-\int_0^{\frac{1}{2}\tau_\gamma} 2p(y^1_{-t})\mathrm{d}t}\big], \tag{5.34}$$

where τ_γ is an exponentially distributed random variable with mean γ. Combining this with (5.32) and (5.33) we find that

$$\begin{aligned}\overline{F}_\gamma w^{1,p}_{22}(x) &= (\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} p(y^1_{-t/2})\mathrm{d}t}\big]x_2(1-x_2)\\ &= q_\gamma E\big[1 - e^{-\langle \mathcal{Z}^\gamma_x, p\rangle}\big]x_2(1-x_2)\\ &= \mathcal{U}_\gamma p(x_1)x_2(1-x_2),\end{aligned} \tag{5.35}$$

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

We still have to show that $\overline{\mathcal{W}}_{\rm cat}$ satisfies condition (iv) from Definition 1.1. For any $\alpha > 0$ and $p \in \mathcal{C}_+[0,1]$, by scaling (Lemma 2.2) and (5.29),

$$F_c w^{\alpha, p} = \alpha F_{\frac{c}{\alpha}} w^{1,\frac{p}{\alpha}} = \alpha(1+\frac{\alpha}{c})^{-1}\overline{F}_{\frac{c}{\alpha}} w^{1,\frac{p}{\alpha}} = w^{(\frac{1}{\alpha}+\frac{1}{c})^{-1}, (\frac{1}{\alpha}+\frac{1}{c})^{-1}\mathcal{U}_{\frac{c}{\alpha}}(\frac{p}{\alpha})}. \tag{5.36}$$

By Lemma 2.1, this diffusion matrix is continuous, which implies that $\mathcal{U}_{\frac{c}{\alpha}}(\frac{p}{\alpha})$ is continuous. ∎

Our proof of Propostion 5.9 has a corollary.

Corollary 5.10 (Continuity in parameters) *The map* $(x,\gamma) \mapsto \mathcal{Q}_\gamma(x,\cdot)$ *from* $[0,1] \times (0,\infty)$ *to* $\mathcal{M}_1(\mathcal{M}[0,1])$ *and the map* $(x,\gamma,p) \mapsto \mathcal{U}_\gamma p(x)$ *from* $[0,1]\times(0,\infty)\times\mathcal{C}_+[0,1]$ *to* $\mathbb{R}$ *are continuous.*

Proof By Lemma 2.1, the diffusion matrix in (5.36) is continuous in x, γ, and p, which implies the continuity of $\mathcal{U}_\gamma p(x)$. It follows that the map $(x,\gamma) \mapsto \int \mathcal{Q}_\gamma(x,\mathrm{d}\chi)e^{-\langle \chi, f\rangle}$ is continuous for all $f \in \mathcal{C}_+[0,1]$, so by [Kal76, Theorem 4.2], $(x,\gamma) \mapsto \mathcal{Q}_\gamma(x,\cdot)$ is continuous. ∎

Proof of Theorem 1.4 (a) We need to show that $\mathcal{W}_{\rm cat}$ is a renormalization class and that F_c maps the subclasses $\mathcal{W}^{l,r}_{\rm cat}$ into themselves. It has already been explained in Section 2 that the latter fact is a consequence of Lemma 2.4. Since in Proposition 5.9 it has been shown that $\overline{\mathcal{W}}_{\rm cat}$ is a renormalization class, we are left with the task to show that F_c maps $\mathcal{W}_{\rm cat}$ into itself. By (5.29) and scaling, it suffices to show that $\mathcal{U}_\gamma$ maps $\mathcal{H}$ into itself.

Fix $0 \leq x \leq \tilde{x} \leq 1$. By Lemma 5.6, we can couple the processes $\mathbf{y}^\gamma_x$ and $\mathbf{y}^\gamma_{\tilde{x}}$ from (3.6) such that

$$\mathbf{y}^\gamma_x(t) \leq \mathbf{y}^\gamma_{\tilde{x}}(t) \quad \forall t \leq 0 \quad \text{a.s.} \tag{5.37}$$

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

$$\begin{aligned}&\big|\mathcal{U}_\gamma p(\tilde{x}) - \mathcal{U}_\gamma p(x)\big|\\ &= \Big|(\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} p(\mathbf{y}^\gamma_{\tilde{x}}(-t/2))\mathrm{d}t}\big] - (\tfrac{1}{\gamma}+1)E\big[1 - e^{-\int_0^{\tau_\gamma} p(\mathbf{y}^\gamma_x(-t/2))\mathrm{d}t}\big]\Big|\\ &\leq (\tfrac{1}{\gamma}+1)E\Big[\int_0^{\tau_\gamma} \big|p(\mathbf{y}^\gamma_{\tilde{x}}(-t/2)) - p(\mathbf{y}^\gamma_x(-t/2))\big|\mathrm{d}t\Big]\\ &\leq (\tfrac{1}{\gamma}+1)LE\Big[\int_0^{\tau_\gamma} \big|\mathbf{y}^\gamma_{\tilde{x}}(-t/2) - \mathbf{y}^\gamma_x(-t/2)\big|\mathrm{d}t\Big]\\ &= (\tfrac{1}{\gamma}+1)L\gamma(\tilde{x}-x) = L(1+\gamma)|\tilde{x}-x|,\end{aligned} \tag{5.38}$$

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

5.4 Monotone and concave catalyzing functions

In this section we prove that the log-Laplace operators $\mathcal{U}_\gamma$ from (3.9) map monotone functions into monotone functions, and monotone concave functions into monotone concave functions. We do not know if in general $\mathcal{U}_\gamma$ maps concave functions into concave functions.

Proposition 5.11 (Preservation of monotonicity and concavity) *Let $\gamma > 0$. Then:*

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

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

Proof Our proof of Proposition 5.11 is in part based on ideas from [BCGdH97, Appendix A]. The proof is quite long and will depend on several lemmas. We remark that part (a) can be proved in a more elementary way using Lemma 5.6.

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

$$\begin{array}{rl}\text{(i)} & \mathcal{D}(A) \text{ is dense,}\\ \text{(ii)} & A \text{ is dissipative,}\\ \text{(iii)} & \mathcal{R}(1-\alpha A) \text{ is dense for some, and hence for all } \alpha > 0.\end{array} \tag{5.39}$$

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

$$(1-\alpha\overline{A})^{-1}u = \int_0^\infty S_t u\, \alpha^{-1} e^{-t/\alpha} \mathrm{d}t \qquad (u \in V,\ \alpha > 0). \tag{5.40}$$

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

$$\begin{array}{rl}\text{(i)} & 1 \in \mathcal{D}(\overline{A}) \text{ and } \overline{A}1 = 0,\\ \text{(ii)} & \mathcal{D}(A) \text{ is dense,}\\ \text{(iii)} & A \text{ satisfies the positive maximum principle,}\\ \text{(iv)} & \mathcal{R}(1-\alpha A) \text{ is dense for some, and hence for all } \alpha > 0.\end{array} \tag{5.41}$$

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

$$S^g_t u(x) = E^x\big[u(\xi(t))\, e^{\int_0^t g(\xi(s))\mathrm{d}s}\big] \qquad (t \geq 0,\ x \in E,\ g, u \in \mathcal{C}(E)). \tag{5.42}$$

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

$$Bu(x,y) := y(1-y)\tfrac{\partial^2}{\partial y^2}u(x,y) + \tfrac{1}{\gamma}(x-y)\tfrac{\partial}{\partial y}u(x,y). \tag{5.43}$$

Below, we will prove:

Lemma 5.12 (Feller semigroup) *The closure in $\mathcal{C}([0,1]^2)$ of the operator B generates a Feller semigroup on $\mathcal{C}([0,1]^2)$.*

Write

$$\begin{aligned}
\mathcal{C}_+ &:= \big\{u \in \mathcal{C}([0,1]^2) : u \geq 0\big\},\\
\mathcal{C}_{1+} &:= \big\{u \in \mathcal{C}^{(1)}([0,1]^2) : \tfrac{\partial}{\partial y}u,\, \tfrac{\partial}{\partial x}u \geq 0\big\},\\
\mathcal{C}_{2+} &:= \big\{u \in \mathcal{C}^{(2)}([0,1]^2) : \tfrac{\partial^2}{\partial y^2}u,\, \tfrac{\partial^2}{\partial x \partial y}u,\, \tfrac{\partial^2}{\partial x^2}u \geq 0\big\}.
\end{aligned} \tag{5.44}$$

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

Lemma 5.13 (Preserved classes) *Let $g \in \mathcal{C}([0,1]^2)$ and let $(S^g_t)_{t\geq 0}$ be the strongly continuous semigroup with generator $\overline{B} + g$. Then, for each $t \geq 0$:*

(a) *If $g \in \overline{\mathcal{C}_{1+}}$, then S^g_t maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+}}$ into itself.*

(b) *If $g \in \overline{\mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$, then S^g_t maps $\overline{\mathcal{C}_+ \cap \mathcal{C}_{1+} \cap \mathcal{C}_{2+}}$ into itself.*

To see why Lemma 5.13 implies Proposition 5.11, let $(\mathbf{x}(t), \mathbf{y}(t))_{t\geq 0}$ denote the Feller process in $[0,1]^2$ generated by $\overline{B}$. It is easy to see that $\mathbf{x}(t) = \mathbf{x}(0)$ a.s. for all $t \geq 0$. For fixed $\mathbf{x}(0) = x$, the process $(\mathbf{y}(t))_{t\geq 0}$ is the diffusion given by the SDE (5.20). Therefore, by Feynman-Kac, for each $g \in \mathcal{C}([0,1]^2)$,

$$E^y\big[e^{\int_0^t g(x,\mathbf{y}(s))\mathrm{d}s}\big] = S^g_t 1(x,y), \tag{5.45}$$

where 1 denotes the constant function $1 \in \mathcal{C}([0,1]^2)$. By (3.9),

$$\mathcal{U}_\gamma f(x) = (\tfrac{1}{\gamma}+1)\Big(1 - \int \Gamma^\gamma_x(\mathrm{d}y) E^y\big[e^{-\int_0^{\tau_\gamma} f(\mathbf{y}_x(s))\mathrm{d}s}\big]\Big) \qquad (f \in \mathcal{C}_+[0,1]), \tag{5.46}$$

where Γ^γ_x is the invariant law of $(\mathbf{y}(t))_{t\geq 0}$ from Corollary 5.4 and τ_γ is an exponential time with mean γ, independent of $(\mathbf{y}(t))_{t\geq 0}$. Setting $g(x,y) := -f(y)$ in (5.45), using the ergodicity of $(\mathbf{y}(t))_{t\geq 0}$ (see Corollary 5.4), we find that for each $z \in [0,1]$ and $t \geq 0$,

$$\begin{aligned}
\int \Gamma^\gamma_x(\mathrm{d}y) E^y\big[e^{-\int_0^t f(\mathbf{y}(s))\mathrm{d}s}\big] &= \lim_{r\to\infty} \int P^z[\mathbf{y}(r) \in \mathrm{d}y]\, E^y\big[e^{-\int_0^t g(x,\mathbf{y}(s))\mathrm{d}s}\big]\\
&= \lim_{r\to\infty} S^0_r S^g_t 1(x,z).
\end{aligned} \tag{5.47}$$

It follows from Lemma 5.13 that for each fixed r, t, and z, the function $x \mapsto S^0_r S^g_t 1(x,z)$ is nondecreasing if f is nonincreasing, and nondecreasing and convex if f is nonincreasing and concave. Therefore, taking the expectation over the randomness of τ_γ, the claims follow from (5.46) and (5.47). ∎

We still need to prove Lemmas 5.12 and 5.13.

Proof of Lemma 5.12 It is easy to see that the operator B from (5.43) is densely defined, satisfies the positive maximum principle, and maps the constant function 1 into 0. Therefore, by Hille-Yosida (5.41), we must show that the range $\mathcal{R}(1-\alpha B)$ is dense in $\mathcal{C}([0,1]^2)$ for some, and hence for all $\alpha > 0$. Let $\mathcal{P}_n$ denote the space of polynomials on $[0,1]^2$ of n-th and lower order, i.e., the space of functions $f : [0,1]^2 \to \mathbb{R}$ of the form

$$f(x,y) = \sum_{k,l\geq 0} a_{kl}\, x^k y^l \quad \text{with } a_{k,l} = 0 \text{ for } k+l > n. \tag{5.48}$$

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

As a first step towards proving Lemma 5.13, we prove:

Lemma 5.14 (Smooth solutions to Laplace equation) *Let $\alpha > 0$, $g \in \mathcal{C}^{(2)}([0,1])$, $g \leq 0$, $v \in \mathcal{C}([0,1]^2)$, and assume that $u \in \mathcal{C}^{(\infty)}([0,1]^2)$ solves the Laplace equation*

$$(1-\alpha(B+g))u = v. \tag{5.49}$$

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

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

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

$$\begin{aligned} B' &:= y(1-y)\frac{\partial^2}{\partial y^2} + \big(c(x-y) + 2(\tfrac{1}{2}-y)\big)\frac{\partial}{\partial y}, \\ B'' &:= y(1-y)\frac{\partial^2}{\partial y^2} + \big(c(x-y) + 4(\tfrac{1}{2}-y)\big)\frac{\partial}{\partial y}. \end{aligned} \tag{5.50}$$

Then

$$\begin{aligned} \tfrac{\partial}{\partial y}Bu &= (B'-c)u^y, \quad & \tfrac{\partial}{\partial y}B'u &= (B''-c-2)u^y, \\ \tfrac{\partial}{\partial x}Bu &= Bu^x + cu^y, \quad & \tfrac{\partial}{\partial x}B'u &= B'u^x + cu^y. \end{aligned} \tag{5.51}$$

Therefore, it is easy to see that

$$\begin{array}{ll} \text{(i)} & (1-\alpha(B'-c+g))u^y = v^y + \alpha g^y u, \\ \text{(ii)} & (1-\alpha(B+g))u^x = v^x + \alpha(cu^y + g^x u), \\ \text{(iii)} & (1-\alpha(B''-2c-2+g))u^{yy} = v^{yy} + \alpha(2g^y u^y + g^{yy}u), \\ \text{(iv)} & (1-\alpha(B'-c+g))u^{xy} = v^{xy} + \alpha(cu^{yy} + g^y u^x + g^{xy}u + g^x u^y), \\ \text{(v)} & (1-\alpha(B+g))u^{xx} = v^{xx} + \alpha(2cu^{xy} + 2g^x u^x + g^{xx}u), \end{array} \tag{5.52}$$

where in (i) and (ii) we assume that $v \in \mathcal{C}^{(1)}([0,1]^2)$ and in (iii)–(v) we assume that $v \in \mathcal{C}^{(2)}([0,1]^2)$. By Lemma 5.12, the closure of the operator B generates a Feller processes in $[0,1]^2$. Exactly the same proof shows that B' and B'' also generate Feller processes on $[0,1]^2$. Therefore, by Feynman-Kac, u is nonnegative if v is nonnegative and $u^y, \ldots, u^{xx}$ are nonnegative if the right-hand sides of the equations (i)–(v) are well-defined and nonnegative. (Instead of using Feynman-Kac, this follows more elementarily from the fact that B, B', and B'' satisfy the positive maximum principle.) In particular, if $g^y, g^x \geq 0$ and $v \in \mathcal{C}^{(1)}([0,1]^2)$, $v, v^y, v^x \geq 0$, then it follows that $u, u^y, u^x \geq 0$. If moreover $g^{yy}, g^{xy}, g^{xx} \geq 0$ and $v \in \mathcal{C}^{(2)}([0,1]^2)$, $v^{yy}, v^{xy}, v^{yy} \geq 0$, then also $u^{yy}, u^{xy}, u^{yy} \geq 0$. ∎

In order to prove Lemma 5.13, based on Lemma 5.14, we will show that the Laplace equation (5.49) has smooth solutions u for sufficiently many functions v. Here 'suffiently many' will mean dense in the topology of uniform convergence of functions and their derivatives up to second order. To this aim, we make $\mathcal{C}^{(2)}([0,1]^2)$ into a Banach space by equipping it with the norm

$$\|u\|_{(2)} := \|u\| + \|u^y\| + \|u^x\| + \|u^{yy}\| + 2\|u^{xy}\| + \|u^{xx}\|. \tag{5.53}$$

Here, to reduce notation, we denote the supremumnorm by $\|f\| := \|f\|_\infty$. Note the factor 2 in the second term from the right in (5.53), which is crucial for the next key lemma.

Lemma 5.15 (Semigroup on twice diffferentiable functions) *The closure in $\mathcal{C}^{(2)}([0,1]^2)$ of the operator B generates a strongly continuous contraction semigroup on $\mathcal{C}^{(2)}([0,1]^2)$.*

Proof We must check the conditions (i)–(iii) from (5.39). It is well-known (see for example [EK86, Proposition 7.1 from the appendix]) that the space $\mathcal{P}_\infty$ of polynomials is dense in $\mathcal{C}^{(2)}([0,1]^2)$. Therefore $\mathcal{D}(B) = \mathcal{C}^{(\infty)}([0,1]^2)$ is dense, and copying the proof of Lemma 5.12 we see that $\mathcal{R}(1-\alpha B)$ is dense for all but countably many α. To complete the proof, we must show that B is dissipative, i.e., that

$$\|(1-\varepsilon B)u\|_{(2)} \geq \|u\|_{(2)} \qquad (\varepsilon > 0,\ u \in \mathcal{C}^{(\infty)}([0,1]^2)). \tag{5.54}$$

Using (5.51), we calculate

$$\begin{aligned}
\tfrac{\partial}{\partial y}(1-\varepsilon B)u &= (1-\varepsilon(B'-c))u^y,\\
\tfrac{\partial}{\partial x}(1-\varepsilon B)u &= (1-\varepsilon B)u^x - \varepsilon c u^y,\\
\tfrac{\partial^2}{\partial y^2}(1-\varepsilon B)u &= (1-\varepsilon(B''-2c-2))u^{yy},\\
\tfrac{\partial^2}{\partial x \partial y}(1-\varepsilon B)u &= (1-\varepsilon(B'-c))u^{xy} - \varepsilon c u^{yy},\\
\tfrac{\partial^2}{\partial x^2}(1-\varepsilon B)u &= (1-\varepsilon B)u^{xx} - 2\varepsilon c u^{xy}.
\end{aligned} \tag{5.55}$$

Using the disipativity of B, B', and B'' with respect to the supremumnorm (which follows from the positive maximum principle) we see that $\|(1-\varepsilon(B'-c))u^y\| = (1+\varepsilon c)\|(1-\frac{\varepsilon}{1+\varepsilon c}B)u^y\| \geq (1+\varepsilon c)\|u^y\|$ etc. We conclude therefore from (5.55) that

$$\begin{aligned}
\|(1-\varepsilon B)u\|_{(2)} \geq\ & \|(1-\varepsilon B)u\| + \|(1-\varepsilon(B'-c))u^y\| + \|(1-\varepsilon B)u^x\| - \varepsilon c\|u^y\|\\
&+\|(1-\varepsilon(B''-2c-2))u^{yy}\| + 2\|(1-\varepsilon(B'-c))u^{xy}\| - 2\varepsilon c\|u^{yy}\|\\
&+\|(1-\varepsilon B)u^{xx}\| - 2\varepsilon c\|u^{xy}\|\\
\geq\ & \|u\| + (1+\varepsilon c)\|u^y\| + \|u^x\| - \varepsilon c\|u^y\|\\
&+(1+\varepsilon(2c+2))\|u^{yy}\| + 2(1+\varepsilon c)\|u^{xy}\| - 2\varepsilon c\|u^{yy}\|\\
&+\|u^{xx}\| - 2\varepsilon c\|u^{xy}\| \geq \|u\|_{(2)}
\end{aligned} \tag{5.56}$$

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

Proof of Lemma 5.13 Let $g \in \mathcal{C}^{(2)}([0,1]^2)$. Then $u \mapsto gu$ is a bounded operator on both $\mathcal{C}([0,1]^2)$ and $\mathcal{C}^{(2)}([0,1]^2)$, so we can choose a $\lambda > 0$ such that

$$\|gu\| \leq \lambda\|u\| \quad \text{and} \quad \|gu\|_{(2)} \leq \lambda\|u\|_{(2)} \tag{5.57}$$

for all u in $\mathcal{C}([0,1]^2)$ and $\mathcal{C}^{(2)}([0,1]^2)$, respectively. Put $\tilde{g} := g - \lambda$. By Lemma 5.12, $\overline{B} + \tilde{g}$ generates a strongly continuous contraction semigroup $(S^{\tilde{g}}_t)_{t\geq 0} = (e^{-\lambda t}S^g_t)_{t\geq 0}$ on $\mathcal{C}([0,1]^2)$. Note that $\mathcal{R}(1-\alpha(B+\tilde{g}))$ is the space of all $v \in \mathcal{C}([0,1]^2)$ for which the Laplace equation $(1-\alpha(B+\tilde{g}))u = v$ has a solution $u \in \mathcal{C}^{(\infty)}([0,1]^2)$. Therefore, by Lemma 5.14, for each $\alpha > 0$:

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

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

(5.58)

By Lemma 5.15, the restriction of the semigroup $(S_t^{\tilde{g}})_{t\geq 0}$ to $\mathcal{C}^{(2)}([0,1]^2)$ is strongly continuous and contractive in the norm $\|\cdot\|_{(2)}$. Therefore, by Hille-Yosida (5.39), $\mathcal{R}(1-\alpha(B+\tilde{g}))$ is dense in $\mathcal{C}^{(2)}([0,1]^2)$ for each $\alpha>0$. It follows that $\mathcal{R}(1-\alpha(B+\tilde{g}))\cap\mathcal{C}_+\cap\mathcal{C}_{1+}$ is dense in $\mathcal{C}_+\cap\mathcal{C}_{1+}$ and likewise $\mathcal{R}(1-\alpha(B+\tilde{g}))\cap\mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+}$ is dense in $\mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+}$, both in the norm $\|\cdot\|_{(2)}$. Note that we need density in the norm $\|\cdot\|_{(2)}$ here: if we would only know that $\mathcal{R}(1-\alpha(B+\tilde{g}))$ is a dense subset of $\mathcal{C}([0,1]^2)$ in the norm $\|\cdot\|$, then $\mathcal{R}(1-\alpha(B+\tilde{g}))\cap\mathcal{C}_+\cap\mathcal{C}_{1+}$ might be empty. By approximation in the norm $\|\cdot\|_{(2)}$ it follows from (5.58) that:

$$\begin{array}{rl} \text{(i)} & \text{If } g\in\mathcal{C}_{1+}, \text{ then } (1-\alpha(\overline{B}+\tilde{g}))^{-1} \text{ maps } \mathcal{C}_+\cap\mathcal{C}_{1+} \text{ into itself.} \\ \text{(ii)} & \text{If } g\in\mathcal{C}_{1+}\cap\mathcal{C}_{2+}, \text{ then } (1-\alpha(\overline{B}+\tilde{g}))^{-1} \text{ maps } \mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+} \text{ into itself.} \end{array} \tag{5.59}$$

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

$$\begin{array}{rl} \text{(i)} & \text{If } g\in\mathcal{C}_{1+}, \text{ then } (1-\alpha(\overline{B}+\tilde{g}))^{-1} \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}} \text{ into itself.} \\ \text{(ii)} & \text{If } g\in\mathcal{C}_{1+}\cap\mathcal{C}_{2+}, \text{ then } (1-\alpha(\overline{B}+\tilde{g}))^{-1} \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+}} \text{ into itself.} \end{array} \tag{5.60}$$

For $\varepsilon>0$ let

$$G_\varepsilon := \varepsilon^{-1}\big((1-\varepsilon(\overline{B}+\tilde{g}))^{-1}-1\big) \tag{5.61}$$

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

$$e^{G_\varepsilon t} = e^{-\varepsilon^{-1}t}\sum_{n=0}^\infty \frac{t^n}{n!}(1-\varepsilon(\overline{B}+\tilde{g}))^{-n} \qquad (t\geq 0), \tag{5.62}$$

and therefore, by (5.60), for each $t\geq 0$:

$$\begin{array}{rl} \text{(i)} & \text{If } g\in\mathcal{C}_{1+}, \text{ then } e^{G_\varepsilon t} \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}} \text{ into itself.} \\ \text{(ii)} & \text{If } g\in\mathcal{C}_{1+}\cap\mathcal{C}_{2+}, \text{ then } e^{G_\varepsilon t} \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+}} \text{ into itself.} \end{array} \tag{5.63}$$

Finally

$$e^{-\lambda t}S_t^g u = S_t^{\tilde{g}} u = \lim_{\varepsilon\to 0} e^{G_\varepsilon t} u \qquad (t\geq 0,\ u\in\mathcal{C}([0,1]^2)), \tag{5.64}$$

so (5.63) implies that for each $t\geq 0$:

$$\begin{array}{rl} \text{(i)} & \text{If } g\in\mathcal{C}_{1+}, \text{ then } S_t^g \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}} \text{ into itself.} \\ \text{(ii)} & \text{If } g\in\mathcal{C}_{1+}\cap\mathcal{C}_{2+}, \text{ then } S_t^g \text{ maps } \overline{\mathcal{C}_+\cap\mathcal{C}_{1+}\cap\mathcal{C}_{2+}} \text{ into itself.} \end{array} \tag{5.65}$$

Using the continuity of S_t^g in g (which follows from Feynman-Kac (5.42)) we arrive at the statements in Lemma 5.13. ∎

6 Convergence to a time-homogeneous process

6.1 Convergence of certain Markov chains

Section 6 is devoted to the proof of Theorem 3.2. In the present subsection, we start by formulating a theorem about the convergence of certain Markov chains to continuous-time processes. In Section 6.2 we specialize to Poisson-cluster branching processes and superprocesses. In Section 6.3, finally, we carry out the necessary calculations for the specific processes from Theorem 3.2.

Let E be a *compact* metrizable space. Let $A : \mathcal{D}(A) \to \mathcal{C}(E)$ be an operator defined on a domain $\mathcal{D}(A) \subset \mathcal{C}(E)$. We say that a process $\mathbf{y} = (\mathbf{y}_t)_{t\geq 0}$ solves the martingale problem for A if $\mathbf{y}$ has sample paths in $\mathcal{D}_E[0,\infty)$ and for each $f \in \mathcal{D}(A)$, the process $(M^f_t)_{t\geq 0}$ given by

$$M^f_t := f(\mathbf{y}_t) - \int_0^t Af(\mathbf{y}_s)\mathrm{d}s \qquad (t \geq 0) \tag{6.1}$$

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

$$P_k(x, \mathrm{d}y) = P\big[X^{(n)}_k \in \mathrm{d}y \big| X^{(n)}_{k-1} = x\big] \qquad (1 \leq k \leq m(n)). \tag{6.2}$$

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

$$A^{(n)}_k f(x) := (\varepsilon^{(n)}_k)^{-1}\Big(\int_E P_k(x,\mathrm{d}y)f(y) - f(x)\Big) \qquad (1 \leq k \leq m(n),\ f \in \mathcal{C}(E)). \tag{6.3}$$

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

$$t^{(n)}_k := \sum_{l=1}^k \varepsilon^{(n)}_l \qquad (1 \leq k \leq m(n)), \tag{6.4}$$

and put

$$k^{(n)}(t) := \max\big\{k \ :\ 0 \leq k \leq m(n),\ t^{(n)}_k \leq t\big\} \qquad (t \geq 0). \tag{6.5}$$

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

$$\mathbf{y}^{(n)}_t := X^{(n)}_{k^{(n)}(t)} \qquad (t \geq 0). \tag{6.6}$$

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

Theorem 6.1 (Convergence of Markov chains) *Assume that $\mathcal{L}(X^{(n)}_0) \Rightarrow \mu$ as $n \to \infty$ for some probability law μ on E. Suppose that there exists at most one (in law) solution to the martingale problem for A with initial law μ. Assume that the linear span of $\mathcal{D}(A)$ contains an algebra that separates points. Assume that*

$$\text{(i)}\ \lim_{n\to\infty} \sum_{k=1}^{m(n)} \varepsilon^{(n)}_k = \infty, \qquad \text{(ii)}\ \lim_{n\to\infty} \sup_{k:\ t^{(n)}_k \leq T} \varepsilon^{(n)}_k = 0, \tag{6.7}$$

and

$$\lim_{n\to\infty} \sup_{k:\ t^{(n)}_k\leq T} \big\|A^{(n)}_k f - Af\|_\infty = 0 \qquad (f \in \mathcal{D}(A)) \tag{6.8}$$

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

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

$$f(X_k^{(n)}) - \sum_{i=1}^{k} \varepsilon_i^{(n)} A_i^{(n)} f(X_{i-1}^{(n)}) \qquad (0 \leq k \leq m(n)) \tag{6.9}$$

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

$$f(\mathbf{y}_t^{(n)}) \ - \sum_{i=1}^{k^{(n)}(t)} \varepsilon_i^{(n)} A_i^{(n)} f(\mathbf{y}^{(n)}_{t^{(n)}_{i-1}}) \qquad (t \geq 0) \tag{6.10}$$

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

$$\lfloor t \rfloor^{(n)} := t^{(n)}_{k^{(n)}(t)} \qquad (t \geq 0) \tag{6.11}$$

and set

$$\phi_t^{(n)} := A^{(n)}_{k^{(n)}(t)+1} f(\mathbf{y}^{(n)}_{\lfloor t \rfloor^{(n)}}) 1_{\{t < t^{(n)}_{m(n)}\}} \qquad (t \geq 0) \tag{6.12}$$

and

$$\xi_t^{(n)} := f(\mathbf{y}_t^{(n)}) + \int_{\lfloor t \rfloor^{(n)}}^{t} \phi_s^{(n)} \mathrm{d}s \qquad (t \geq 0). \tag{6.13}$$

Then we can rewrite the martingale in (6.10) as

$$\xi_t^{(n)} - \int_0^t \phi_s^{(n)} \mathrm{d}s. \tag{6.14}$$

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

$$\begin{array}{rl}
\text{(i)} & \displaystyle\sup_{n \geq N} \sup_{t \leq T} E\big[|\xi_t^{(n)}|\big] < \infty, \\
\text{(ii)} & \displaystyle\sup_{n \geq N} \sup_{t \leq T} E\big[|\phi_t^{(n)}|\big] < \infty, \\
\text{(iii)} & \displaystyle\lim_{n \to \infty} E\Big[\big(\xi_T^{(n)} - f(\mathbf{y}_T^{(n)})\big) \prod_{i=1}^{r} h_i(\mathbf{y}_{s_i}^{(n)})\Big] = 0, \\
\text{(iv)} & \displaystyle\lim_{n \to \infty} E\Big[\big(\phi_T^{(n)} - Af(\mathbf{y}_T^{(n)})\big) \prod_{i=1}^{r} h_i(\mathbf{y}_{s_i}^{(n)})\Big] = 0, \\
\text{(v)} & \displaystyle\lim_{n \to \infty} E\Big[\sup_{t \in \mathbb{Q} \cap [0,T]} \big|\xi_t^{(n)} - f(\mathbf{y}_t^{(n)})\big|\Big] = 0, \\
\text{(vi)} & \displaystyle\sup_{n \geq N} E\big[\|\phi^{(n)}\|_{p,T}\big] < \infty \qquad \text{for some } p \in (1, \infty],
\end{array} \tag{6.15}$$

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

$$\|g\|_{p,T} := \Big(\int_0^T |g(t)|^p \mathrm{d}t\Big)^{1/p} \qquad (1 \leq p < \infty) \tag{6.16}$$

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

The conditions (6.15) (i)–(vi) are implied by the stronger conditions

$$\begin{aligned} &\text{(i)} \quad \lim_{n\to\infty} \sup_{0\leq t\leq T} \big\|\xi^{(n)}_t - f(\mathbf{y}^{(n)}_t)\big\|_\infty = 0,\\ &\text{(ii)} \quad \lim_{n\to\infty} \sup_{0\leq t\leq T} \big\|\phi^{(n)}_t - Af(\mathbf{y}^{(n)}_t)\big\|_\infty = 0, \end{aligned} \tag{6.17}$$

where we denote the essential supremumnorm of a real-valued random variable X by $\|X\|_\infty := \inf\{K \geq 0 : |X| \leq K \text{ a.s.}\}$. Condition (6.17) (ii) is implied by (6.7) (i) and (6.8). To see that also (6.17) (i) holds, set

$$M_n := \sup_{0\leq t\leq T} \big\|\phi^{(n)}_t\big\|_\infty, \tag{6.18}$$

and estimate

$$\sup_{0\leq t\leq T} \big\|\xi^{(n)}_t - f(\mathbf{y}^{(n)}_t)\big\|_\infty \leq M_n \sup\{\varepsilon^{(n)}_k : 1 \leq k \leq m(n),\ t^{(n)}_k \leq T\}. \tag{6.19}$$

Condition (6.17) (ii) implies that $\limsup_n M_n < \infty$ and therefore the right-hand side of (6.19) tends to zero by assumption (6.7) (ii). ∎

6.2 Convergence of certain branching processes

In this section we apply Theorem 6.1 to certain branching processes and superprocesses.

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

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

$$\begin{cases} \frac{\partial}{\partial t} u_t = \overline{A} u_t + \beta u_t - \alpha u_t^2 & (t \geq 0),\\ u_0 = f \end{cases} \tag{6.20}$$

if $t \mapsto u_t$ is continuously differentiable (in $\mathcal{C}(E)$), $u_t \in \mathcal{D}(\overline{A})$ for all $t \geq 0$, and (6.20) holds. We say that u is a *mild* solution to (6.20) if $t \mapsto u_t$ is continuous and

$$u_t = P_t f + \int_0^t P_{t-s}(\beta u_s - \alpha u_s^2)\mathrm{d}s \qquad (t \geq 0). \tag{6.21}$$

Lemma 6.2 (Mild and classical solutions) *Equation (6.20) has a unique $\mathcal{C}_+(E)$-valued mild solution u for each $f \in \mathcal{C}_+(E)$, and $f > 0$ implies that $u_t > 0$ for all $t \geq 0$. If moreover $f \in \mathcal{D}(\overline{A})$ then u is a classical solution. For each $t \geq 0$, u_t depends continuously on $f \in \mathcal{C}_+(E)$.*

Proof It follows from [Paz83, Theorems 6.1.2, 6.1.4, and 6.1.5] that for each $f \in \mathcal{C}(E)$, (6.20) has a unique solution $(u_t)_{0\leq t<T}$ up to an explosion time T, and that this is a classical solution if $f \in \mathcal{D}(\overline{A})$. Moreover, u_t depends continuously on f. Using comparison arguments based on the fact that $\overline{A}$ satisfies the positive maximum principle (which follows from Hille-Yosida (5.41)) one easily proves the other statements; compare [FS04, Lemmas 23 and 24]. ∎

We denote the (mild or classical) solution of (6.20) by $\mathcal{U}_t f := u_t$; then $\mathcal{U}_t : \mathcal{C}_+(E) \to \mathcal{C}_+(E)$ are continuous operators and $\mathcal{U} = (\mathcal{U}_t)_{t\geq 0}$ is a (nonlinear) semigroup on $\mathcal{C}_+(E)$.

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

$$F_f(\mu) := e^{-\langle \mu, f \rangle} \qquad (f \in \mathcal{C}_+(E),\ f > 0,\ \mu \in \mathcal{M}(E)). \tag{6.22}$$

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

$$\mathcal{D}(\mathcal{G}) := \{F_f : f \in \mathcal{D}(A),\ f > 0\}, \tag{6.23}$$

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

$$\mathcal{G}F_f(\mu) := -\langle \mu, Af + \beta f - \alpha f^2 \rangle\, e^{-\langle \mu, f \rangle} \qquad (\mu \in \mathcal{M}(E)). \tag{6.24}$$

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

Proposition 6.3 ($(\overline{A}, \alpha, \beta)$-superprocesses) *The martingale problem for the operator $\mathcal{G}$ is well-posed. The solutions to this martingale problem define a Feller process $\mathcal{Y} = (\mathcal{Y}_t)_{t \geq 0}$ in $\mathcal{M}(E)_\infty$ with continuous sample paths, called the $(\overline{A}, \alpha, \beta)$-superprocess. If $\mathcal{Y}_0 = \infty$ then $\mathcal{Y}_t = \infty$ for all $t \geq 0$. If $\mathcal{Y}_0 = \mu \in \mathcal{M}(E)$ then*

$$E^\mu\big[e^{-\langle \mathcal{Y}_t, f \rangle}\big] = e^{-\langle \mu, \mathcal{U}_t f \rangle} \qquad (f \in \mathcal{C}_+(E)). \tag{6.25}$$

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

We are going to extend $\mathcal{G}$ to an operator $\hat{\mathcal{G}}$ that is linear and satisfies the conditions of the Hille-Yosida Theorem (5.41). For any $\gamma \in \mathcal{C}_+(E)$ and $\mu \in \mathcal{M}(E)$, let $\mathrm{Clust}_\gamma(\mu)$ denote a random measure such that on $\{\gamma = 0\}$, $\mathrm{Clust}_\gamma(\mu)$ is equal to μ, and on $\{\gamma > 0\}$, $\mathrm{Clust}_\gamma(\mu)$ is a Poisson cluster measure with intensity $\frac{1}{\gamma}\mu$ and cluster mechanism $\mathcal{Q}(x, \cdot) = \mathcal{L}(\tau_{\gamma(x)}\delta_x)$, where $\tau_{\gamma(x)}$ is exponentially distributed with mean $\gamma(x)$. It is not hard to see that

$$E\big[e^{-\langle \mathrm{Clust}_\gamma(\mu), f \rangle}\big] = e^{-\langle \mu, \mathcal{V}_\gamma f \rangle} \qquad (f \in \mathcal{C}(E),\ f > 0), \tag{6.26}$$

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

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

$$\mathcal{G}_\alpha F(\mu) := \lim_{\varepsilon \to 0} \varepsilon^{-1}\big(E[F(\mathrm{Clust}_{\varepsilon\alpha}(\mu))] - F(\mu)\big) \tag{6.27}$$

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

$$\mathcal{G}_\beta F(\mu) := \lim_{\varepsilon \to 0} \varepsilon^{-1}\big(F((1 + \varepsilon\beta)\mu) - F(\mu)\big) \tag{6.28}$$

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

$$\mathcal{G}_{\overline{A}} F(\mu) := \lim_{\varepsilon \to 0} \varepsilon^{-1}\big(F(P_\varepsilon^*\mu) - F(\mu)\big), \tag{6.29}$$

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

$$\hat{\mathcal{G}} := \mathcal{G}_\alpha + \mathcal{G}_\beta + \mathcal{G}_{\overline{A}}, \tag{6.30}$$

with domain $\mathcal{D}(\hat{\mathcal{G}}) := \mathcal{D}(\mathcal{G}_\alpha) \cap \mathcal{D}(\mathcal{G}_{\overline{A}})$. If $f \in \mathcal{D}(\overline{A})$, $f > 0$, and F_f is as in (6.22), then it is not hard to see that $\hat{\mathcal{G}} F_f(\infty) = 0$ and

$$\hat{\mathcal{G}} F_f(\mu) := -\langle \mu, \overline{A} f + \beta f - \alpha f^2\rangle\, e^{-\langle \mu, f\rangle} \qquad (\mu \in \mathcal{M}(E)). \tag{6.31}$$

In particular, $\hat{\mathcal{G}}$ extends the operator $\mathcal{G}$ from (6.24). Since $\mathcal{D}(\overline{A})$ is dense in $\mathcal{C}(E)$, it is easy to see that $\{F_f : f \in \mathcal{D}(\overline{A}),\ f > 0\}$ is dense in $\mathcal{C}(\mathcal{M}(E)_\infty)$. Hence $\mathcal{D}(\hat{\mathcal{G}})$ is dense. Using (6.27)–(6.29) it is not hard to show that $\hat{\mathcal{G}}$ satisfies the positive maximum principle. Moreover, by Lemma 6.2, for $f \in \mathcal{D}(\overline{A})$ with $f > 0$, the function $t \mapsto F_{\mathcal{U}_t f}$ from $[0, \infty)$ into $\mathcal{C}(\mathcal{M}(E)_\infty)$ is continuously differentiable, satisfies $F_{\mathcal{U}_t f} \in \mathcal{D}(\hat{\mathcal{G}})$ for all $t \geq 0$, and

$$\tfrac{\partial}{\partial t} F_{\mathcal{U}_t f} = \hat{\mathcal{G}} F_{\mathcal{U}_t f} \qquad (t \geq 0). \tag{6.32}$$

From this it is not hard to see that $\hat{\mathcal{G}}$ also satisfies condition (5.41) (ii), so the closure of $\hat{\mathcal{G}}$ generates a Feller semigroup $(S_t)_{t \geq 0}$ on $\mathcal{C}(\mathcal{M}(E)_\infty)$. It is easy to see that $S_t F_f = F_{\mathcal{U}_t f}$ $(t \geq 0)$. By [EK86, Theorem 4.2.7], this semigroup corresponds to a Feller process $\mathcal{Y}$ with cadlag sample paths in $\mathcal{M}(E)_\infty$. This means that $E^\mu[F_f(\mathcal{Y}_t)] = F_{\mathcal{U}_t f}(\mu)$ for all $f \in \mathcal{D}(\overline{A})$ with $f > 0$. If $\mu = \infty$ this shows that $\mathcal{Y}_t = \infty$ for all $t \geq 0$. If $\mu \in \mathcal{M}(E)$ we obtain (6.25) for $f \in \mathcal{D}(\overline{A})$, $f > 0$; the general case follows by approximation. ∎

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

$$Z_\varepsilon(x) := \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\langle \chi, 1\rangle < \infty \qquad (x \in E) \tag{6.33}$$

and define probability kernels K_ε on E by

$$\int K_\varepsilon(x, \mathrm{d}y) f(y) := \frac{1}{Z_\varepsilon(x)} \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\langle \chi, f\rangle \qquad (f \in B(E)). \tag{6.34}$$

For each $n \geq 0$, let $(\varepsilon_k^{(n)})_{1 \leq k \leq m(n)}$ (with $1 \leq m(n) < \infty$) be positive constants. Let $\mathcal{X}^{(n)} = (\mathcal{X}_0^{(n)}, \ldots, \mathcal{X}_{m(n)}^{(n)})$ be a Poisson-cluster branching process with weight functions $q_{\varepsilon_1^{(n)}}, \ldots, q_{\varepsilon_{m(n)}^{(n)}}$ and cluster mechanisms $\mathcal{Q}_{\varepsilon_1^{(n)}}, \ldots, \mathcal{Q}_{\varepsilon_{m(n)}^{(n)}}$. Define $t_k^{(n)}$ and $k^{(n)}(t)$ as in (6.4)–(6.5). Define processes $\mathcal{Y}^{(n)}$ by

$$\mathcal{Y}_t^{(n)} := \mathcal{X}_{k^{(n)}(t)}^{(n)} \qquad (t \geq 0). \tag{6.35}$$

Theorem 6.4 (Convergence of Poisson-cluster branching processes) *Assume that $\mathcal{L}(\mathcal{X}_0^{(n)}) \Rightarrow \rho$ as $n \to \infty$ for some probability law ρ on $\mathcal{M}(E)$. Suppose that the constants $\varepsilon_k^{(n)}$ fulfill (6.7). Assume that*

$$\begin{array}{ll} \text{(i)} & q_\varepsilon(x) \displaystyle\int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\langle \chi, 1\rangle = 1 + \varepsilon\beta(x) + o(\varepsilon), \\ \text{(ii)} & q_\varepsilon(x) \displaystyle\int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\langle \chi, 1\rangle^2 = \varepsilon\, 2\alpha(x) + o(\varepsilon), \\ \text{(iii)} & q_\varepsilon(x) \displaystyle\int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\langle \chi, 1\rangle^2 1_{\{\langle \chi, 1\rangle > \delta\}} = o(\varepsilon) \end{array} \tag{6.36}$$

for each $\delta > 0$, and

$$\int K_\varepsilon(x, \mathrm{d}y) f(y) = f(x) + \varepsilon A f(x) + o(\varepsilon) \tag{6.37}$$

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

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

Proof We apply Theorem 6.1 to the operator $\mathcal{G}$, where we use the fact that if we view $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)}[0, \infty))$ as a subspace of $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)_\infty}[0, \infty))$ (note the compactification), equipped with the topology of weak convergence, then the induced topology on $\mathcal{M}_1(\mathcal{D}_{\mathcal{M}(E)}[0, \infty))$ is again the topology of weak convergence.

By Proposition 6.3, solutions to the martingale problem for $\mathcal{G}$ are unique. Since $F_f F_g = F_{f+g}$ and $\mathcal{D}(A)$ is a linear space, the linear span of the domain of $\mathcal{G}$ is an algebra. Using the fact that $\mathcal{D}(A)$ is dense in $\mathcal{C}(E)$ we see that this algebra separates points. Therefore, we are left with the task to check (6.8).

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

$$\mathcal{U}_\varepsilon f(x) := q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\big(1 - e^{-\langle \chi, f\rangle}\big) \qquad (x \in E,\ f \in \mathcal{C}_+[0,1],\ f > 0,\ \varepsilon > 0), \tag{6.38}$$

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

$$\int P_\varepsilon(\mu, \mathrm{d}\nu) e^{-\langle \nu, f\rangle} = e^{-\langle \mu, \mathcal{U}_\varepsilon f\rangle}. \tag{6.39}$$

We will show that

$$\lim_{\varepsilon \to 0} \big\| \varepsilon^{-1}(\mathcal{U}_\varepsilon f - f) - (Af + \beta f - \alpha f^2)\big\|_\infty = 0 \qquad (f \in \mathcal{D}(A),\ f > 0). \tag{6.40}$$

Together with (6.39) this implies that

$$\int P_\varepsilon(\mu, \mathrm{d}\nu) F_f(\nu) = F_f(\mu) + \varepsilon \mathcal{G} F_f(\mu) + o(\varepsilon) \qquad (f \in \mathcal{D}(A),\ f > 0), \tag{6.41}$$

uniformly in $\mu \in \mathcal{M}(E)_\infty$ as $\varepsilon \to 0$. Therefore, the result follows from Theorem 6.1.

It remains to prove (6.40). Set $g(z) := 1 - z + \frac{1}{2}z^2 - e^{-z}$ $(z \geq 0)$ and write

$$\mathcal{U}_\varepsilon f(x) = q_\varepsilon(x) \int \mathcal{Q}_\varepsilon(x, \mathrm{d}\chi)\big(\langle \chi, f\rangle - \tfrac{1}{2}\langle \chi, f\rangle^2 + g(\langle \chi, f\rangle)\big). \tag{6.42}$$

Since

$$g(z)=\int_0^z \mathrm{d}y\int_0^y \mathrm{d}x\int_0^x \mathrm{d}t\,e^{-t}\qquad (z\geq 0),\tag{6.43}$$

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

$$0\leq g(z)\leq \tfrac{1}{2}z^2\wedge\tfrac{1}{6}z^3\qquad (z\geq 0).\tag{6.44}$$

Using these facts and (6.36) (ii) and (iii), we find that

$$\begin{aligned}
&q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)g(\langle\chi,f\rangle)\\
&\quad\leq \|f\|_\infty q_\varepsilon(x)\Big\{\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)g(\langle\chi,1\rangle)1_{\{\langle\chi,1\rangle\leq\delta\}}+\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)g(\langle\chi,1\rangle)1_{\{\langle\chi,1\rangle>\delta\}}\Big\}\\
&\quad\leq \|f\|_\infty q_\varepsilon(x)\Big\{\tfrac{1}{6}\delta\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 1_{\{\langle\chi,1\rangle\leq\delta\}}+\tfrac{1}{2}\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}}\Big\}\\
&\quad=\tfrac{1}{6}\delta\|f\|_\infty\big(\varepsilon\,2\alpha(x)+o(\varepsilon)\big)+o(\varepsilon).
\end{aligned}\tag{6.45}$$

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

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)g(\langle\chi,f\rangle)=o(\varepsilon)\tag{6.46}$$

uniformly in x as $\varepsilon\to 0$. By (6.36) (i) and (6.37),

$$\begin{aligned}
q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f\rangle&=\Big(q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle\Big)\Big(\int K_\varepsilon(x,\mathrm{d}y)f(y)\Big)\\
&=\big(1+\varepsilon\beta(x)+o(\varepsilon)\big)\big(f(x)+\varepsilon Af(x)+o(\varepsilon)\big)\\
&=f(x)+\varepsilon\beta(x)f(x)+\varepsilon Af(x)+o(\varepsilon).
\end{aligned}\tag{6.47}$$

Finally, write

$$\begin{aligned}
&q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f\rangle^2\\
&\quad=q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\big(\langle\chi,f(x)\rangle^2+2\langle\chi,f(x)\rangle\langle\chi,f-f(x)\rangle+\langle\chi,f-f(x)\rangle^2\big).
\end{aligned}\tag{6.48}$$

Then, by (6.36) (ii),

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f(x)\rangle^2=f(x)^2\big(\varepsilon\,2\alpha(x)+o(\varepsilon)\big).\tag{6.49}$$

We will prove that

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle^2=o(\varepsilon).\tag{6.50}$$

Then, by Hölder's inequality, (6.36) (ii), and (6.50),

$$\begin{aligned}
&\Big|q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle\langle\chi,f(x)\rangle\Big|\\
&\quad\leq\Big(q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle^2\Big)^{1/2}\Big(q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f(x)\rangle^2\Big)^{1/2}\\
&\quad\leq\big(o(\varepsilon)(2\alpha(x)\varepsilon+o(\varepsilon))\big)^{1/2}=o(\varepsilon).
\end{aligned}\tag{6.51}$$

Inserting (6.49), (6.50) and (6.51) into (6.48) we find that

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f\rangle^2 = \varepsilon\, 2\alpha(x)f(x)^2 + o(\varepsilon). \tag{6.52}$$

Inserting (6.46), (6.47) and (6.52) into (6.42), we arrive at (6.40). We still need to prove (6.50). To this aim, we estimate, using (6.47),

$$\begin{aligned} &q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle^2 1_{\{\langle\chi,1\rangle\leq\delta\}}\\ &\quad\leq \delta\|f-f(x)\|_\infty q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle\\ &\quad= \delta\|f-f(x)\|_\infty\big(\varepsilon Af(x)+o(\varepsilon)\big) \end{aligned} \tag{6.53}$$

and, using (6.36) (iii),

$$\begin{aligned} &q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}}\\ &\quad\leq \|f-f(x)\|_\infty q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,1\rangle^2 1_{\{\langle\chi,1\rangle>\delta\}} = o(\varepsilon). \end{aligned} \tag{6.54}$$

It follows that

$$q_\varepsilon(x)\int \mathcal{Q}_\varepsilon(x,\mathrm{d}\chi)\langle\chi,f-f(x)\rangle^2 \leq \delta\varepsilon\|f-f(x)\|_\infty Af(x)+o(\varepsilon) \tag{6.55}$$

for any $\delta>0$. This implies (6.50) and completes the proof of (6.40). ∎

6.3 Application to the renormalization branching process

Proof of Theorem 3.2 (a) For any $f_0,\ldots,f_k\in\mathcal{C}_+[0,1]$ one has

$$\begin{aligned} &E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k},f_k\rangle}\big]\\ &\quad= E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k-1},f_{k-1}+\mathcal{U}_{\gamma_{n-k}}f_k\rangle}\big]\\ &\quad=\cdots=\quad E\big[e^{-\langle\mathcal{X}_{-n},g_k\rangle}\big], \end{aligned} \tag{6.56}$$

where we define inductively

$$g_0:=f_k\quad\text{and}\quad g_{m+1}:=f_{k-m-1}+\mathcal{U}_{\gamma_{n-k+m}}g_m. \tag{6.57}$$

By the compactness of $[0,1]$ and Corollary 5.10, the map $(\gamma,f)\mapsto\mathcal{U}_\gamma f$ from $(0,\infty)\times\mathcal{C}_+[0,1]$ to $\mathcal{C}_+[0,1]$ (equipped with the supremumnorm) is continuous. Using this fact and (6.56) we find that

$$E\big[e^{-\langle\mathcal{X}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{X}_{-n+k},f_k\rangle}\big]\underset{n\to\infty}{\longrightarrow}E\big[e^{-\langle\mathcal{Y}^{\gamma^*}_{-n},f_0\rangle}\cdots e^{-\langle\mathcal{Y}^{\gamma^*}_{-n+k},f_k\rangle}\big]. \tag{6.58}$$

Since $f_1,\ldots,f_k$ are arbitrary, (3.12) follows. ∎

Proof of Theorem 3.2 (b) We apply Theorem 6.4 to the weight functions q_γ and cluster mechanisms $\mathcal{Q}_\gamma$ from (3.8) and to $A_{\mathrm{WF}}=x(1-x)\frac{\partial^2}{\partial x^2}$ with domain $\mathcal{D}(A_{\mathrm{WF}})=\mathcal{C}^{(2)}[0,1]$, and

$\alpha = \beta = 1$. It is well-known that $\overline{A}_{\mathrm{WF}}$ generates a Feller semigroup [EK86, Theorem 8.2.8]. We observe that

$$\int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, f\rangle = E\big[2\int_0^{\tau_\gamma} f(\mathbf{y}^\gamma_x(-t))\big] = 2E[\tau_\gamma]E\big[f(\mathbf{y}^\gamma_x(0))\big] = \gamma \int \Gamma^\gamma_x(\mathrm{d}y)f(y), \qquad (6.59)$$

where Γ^γ_x is the equilibrium law of the process $\mathbf{y}^\gamma_x$ from Corollary 5.4. It follows from (5.24) that

$$\begin{aligned} &\text{(i)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x) = 0, \\ &\text{(ii)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x)^2 = \frac{\gamma x(1-x)}{1+\gamma}, \\ &\text{(iii)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x)^4 = O(\gamma^2), \end{aligned} \qquad (6.60)$$

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

$$\begin{aligned} &\text{(i)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x) = 0, \\ &\text{(ii)} && \int \Gamma^\gamma_x(\mathrm{d}y)(y-x)^2 = \gamma x(1-x) + o(\gamma), \\ &\text{(iii)} && \int \Gamma^\gamma_x(\mathrm{d}y)1_{\{|y-x|>\delta\}} = o(\gamma), \end{aligned} \qquad (6.61)$$

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

$$\int \Gamma^\gamma_x(\mathrm{d}y)f(x) = f(x) + \gamma\tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2}f(x) + o(\gamma) \qquad (f \in \mathcal{C}^{(2)}[0,1]), \qquad (6.62)$$

uniformly in x as $\gamma \to 0$. (For details, in particular the uniformity in x, see for example Proposition [Swa99, B.1.1].) This shows that condition (6.37) is satisfied. Moreover,

$$\begin{aligned} \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle &= E[2\tau_\gamma] = \gamma, \\ \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^2 &= E[(2\tau_\gamma)^2] = \int_0^\infty z^2\tfrac{1}{\gamma}e^{-z/\gamma}\mathrm{d}z = 2\gamma^2, \\ \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^3 &= E[(2\tau_\gamma)^3] = \int_0^\infty z^3\tfrac{1}{\gamma}e^{-z/\gamma}\mathrm{d}z = 6\gamma^3, \end{aligned} \qquad (6.63)$$

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

$$\begin{aligned} q_\gamma \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle &= 1 + \gamma, \\ q_\gamma \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^2 &= 2\gamma + o(\gamma), \\ q_\gamma \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^3 &= o(\gamma). \end{aligned} \qquad (6.64)$$

This shows that (6.36) is fulfilled. In particular,

$$q_\gamma \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^2 1_{\{\langle \chi, 1\rangle > \delta\}} \leq \delta^{-1} q_\gamma \int \mathcal{Q}_\gamma(x, \mathrm{d}\chi)\langle \chi, 1\rangle^3 = o(\gamma) \qquad (6.65)$$

for all $\delta > 0$. ∎

7 Embedded particle systems

In this section we use embedded particle systems to prove Proposition 3.5. An essential ingredient in the proofs is Proposition 7.15 (a), which will be proved in the Section 8.

7.1 Weighting and Poissonization

Proof of Proposition 3.3 Obviously $q^h_k \in \mathcal{C}_+(E^h)$ for each $k = 1, \ldots, n$. Since $h \in \mathcal{C}_+(E)$ and h is bounded, it is easy to see that the map $\mu \mapsto h\mu$ from $\mathcal{M}(E)$ into $\mathcal{M}(E^h)$ is continuous, and therefore the cluster mechanisms defined in (3.21) are continuous. Since

$$\mathcal{U}^h_k f(x) = \frac{q_k(x)}{h(x)} E\big[1 - e^{-\langle h\mathcal{Z}_x, f\rangle}\big] = \frac{\mathcal{U}_k(hf)(x)}{h(x)} \qquad (x \in E^h,\ f \in B_+(E^h)), \tag{7.1}$$

formula (3.22) holds on E^h. To see that (3.22) holds on $E\backslash E^h$, note that by assumption $\mathcal{U}_k h \leq Kh$ for some $K < \infty$, so if $x \in E\backslash E^h$, then $\mathcal{U}_k h(x) = 0$. By monotonicity also $\mathcal{U}_k(hf)(x) = 0$, while $h\mathcal{U}^h_k f(x) = 0$ by definition. Since $\sup_{x\in E^h} \mathcal{U}^h_k 1(x) = \sup_{x\in E^h} \frac{\mathcal{U}_k h(x)}{h(x)} \leq K < \infty$, the log-Laplace operators $\mathcal{U}^h_k$ satisfy (3.3). If $\mathcal{X}$ is started in an initial state $\mathcal{X}_0$, then the Poisson-cluster branching process $\mathcal{X}^h$ with log-Laplace operators $\mathcal{U}^h_1, \ldots, \mathcal{U}^h_n$ started in $\mathcal{X}^h_0 = h\mathcal{X}_0$ satisfies

$$\begin{aligned} E\big[e^{-\langle h\mathcal{X}_k, f\rangle}\big] &= E\big[e^{-\langle \mathcal{X}_0, \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_k(hf)\rangle}\big] \\ &= E\big[e^{-\langle \mathcal{X}_0, h\mathcal{U}^h_1 \circ \cdots \circ \mathcal{U}^h_k(f)\rangle}\big] = E\big[e^{-\langle \mathcal{X}^h_k, f\rangle}\big] \qquad (f \in B_+(E^h)), \end{aligned} \tag{7.2}$$

which proves (3.23). ∎

Proof of Proposition 3.4 We start by noting that by (3.2),

$$\mathcal{U}_k f(x) = q(x)E\big[1 - e^{-\langle \mathcal{Z}^k_x, f\rangle}\big] = q_k(x)P[\mathrm{Pois}(f\mathcal{Z}^k_x) \neq 0] \qquad (x \in E,\ f \in B_+(E)). \tag{7.3}$$

Into (3.24), we insert

$$\begin{aligned} &P\big[\mathrm{Pois}(h\mathcal{Z}^k_x) \in \cdot\,\big] \\ &\quad = P\big[\mathrm{Pois}(h\mathcal{Z}^k_x) \in \cdot \,\big|\, \mathrm{Pois}(h\mathcal{Z}^k_x) \neq 0\big] P[\mathrm{Pois}(h\mathcal{Z}^k_x) \neq 0] + \delta_0 P[\mathrm{Pois}(h\mathcal{Z}^k_x) = 0]. \end{aligned} \tag{7.4}$$

Here and in similar formulas below, if in a conditional probability the symbol Pois(·) occurs twice with the same argument, then it always refers to the same random variable (and not to independent Poisson point measures with the same intensity, for example). Using moreover (7.3) we can rewrite (3.24) as

$$Q^h_k(x, \cdot\,) = \frac{\mathcal{U}_k h(x)}{h(x)} P\big[\mathrm{Pois}(h\mathcal{Z}^k_x) \in \cdot \,\big|\, \mathrm{Pois}(h\mathcal{Z}^k_x) \neq 0\big] + \frac{h(x) - \mathcal{U}_k h(x)}{h(x)} \delta_0(\,\cdot\,). \tag{7.5}$$

In particular, since we are assuming that h is $\mathcal{U}_k$-subharmonic, this shows that $Q^h_k(x, \cdot\,)$ is a probability measure. Let X^h be the branching particle system with offspring mechanisms $Q^h_1, \ldots, Q^h_k$. Let $Z^{h,k}_x$ be random variables such that $\mathcal{L}(Z^{h,k}_x) = Q^h_k(x, \cdot\,)$. Then, by (3.18), (3.24), (3.20), and (7.3),

$$\begin{aligned} U^h_k f(x) = P[\mathrm{Thin}_f(Z^{h,k}_x) \neq 0] &= \frac{q_k(x)}{h(x)} P[\mathrm{Thin}_f(\mathrm{Pois}(h\mathcal{Z}^k_x)) \neq 0] \\ = \frac{q_k(x)}{h(x)} P[\mathrm{Pois}(hf\mathcal{Z}^k_x) \neq 0] &= \frac{1}{h(x)} \mathcal{U}_k(hf)(x) \qquad (x \in E^h). \end{aligned} \tag{7.6}$$

If $x \in E\backslash E^h$, then $\mathcal{U}_k(hf)(x) \leq \mathcal{U}_k(h)(x) \leq h(x) = 0 =: h\mathcal{U}^h(f)(x)$. This proves (3.25). To see that Q_k^h is a *continuous* offspring mechanism, by [Kal76, Theorem 4.2] it suffices to show that $x \mapsto \int Q_k^h(x, \mathrm{d}\nu)e^{-\langle \nu, g\rangle}$ is continuous for all bounded $g \in \mathcal{C}_+(E^h)$. Indeed, setting $f := 1-e^{-g}$, one has $\int Q_k^h(x, \mathrm{d}\nu)e^{-\langle \nu, g\rangle} = \int Q_k^h(x, \mathrm{d}\nu)(1-f)^\nu = 1-\mathcal{U}_k^h f(x) = 1-\mathcal{U}_k(hf)(x)/h(x)$ which is continuous on E^h by the continuity of q_k and $\mathcal{Q}_k$.

To see that also (3.26) holds, just note that by (3.19), (3.25), and (3.5),

$$\begin{aligned}
&P^{\mathcal{L}(\mathrm{Pois}(h\mu))}[\mathrm{Thin}_f(X_n^h) = 0] = P[\mathrm{Thin}_{U_1^h \circ \cdots \circ U_n^h f}(\mathrm{Pois}(h\mu)) = 0] \\
&\quad = P[\mathrm{Pois}((hU_1^h \circ \cdots \circ U_n^h f)\mu) = 0] = P[\mathrm{Pois}((\mathcal{U}_1 \circ \cdots \circ \mathcal{U}_n(hf))\mu) = 0] \\
&\quad = P^\mu[\mathrm{Pois}(hf\mathcal{X}_n) = 0] = P^\mu[\mathrm{Thin}_f(\mathrm{Pois}(h\mathcal{X}_n)) = 0].
\end{aligned} \tag{7.7}$$

Since this formula holds for all $f \in B_{[0,1]}(E^h)$, formula (3.26) follows. ∎

Remark 7.1 (Boundedness of h) Propositions 3.3 and 3.4 generalize to the case that h is unbounded, except that in this case the cluster mechanism in (3.21) and the offspring mechanism in (3.24) need in general not be continuous. Here, in order for (3.22) and (3.25) to be well-defined, one needs to extend the definition of $\mathcal{U}_k f$ to unbounded functions f, but this can always be done unambiguously [FS03, Lemma 9]. ◇

7.2 Sub- and superharmonic functions

This section contains a number of pivotal calculations involving the log-Laplace operators $\mathcal{U}_\gamma$ from (3.9). In particular, we will prove that the functions $h_{1,1}$, $h_{0,0}$, and $h_{0,1}$ from Lemmas 3.6, 3.7, and 3.8, respectively, are $\mathcal{U}_\gamma$-superharmonic.

We start with an observation that holds for general log-Laplace operators.

Lemma 7.2 (Constant multiples) *Let $\mathcal{U}$ be a log-Laplace operator of the form (3.2) satisfying (3.3) and let $f \in B_+(E)$. Then $\mathcal{U}(rf) \leq r\mathcal{U}f$ for all $r \geq 1$, and $\mathcal{U}(rf) \geq r\mathcal{U}f$ for all $0 \leq r \leq 1$. In particular, if f is $\mathcal{U}$-superharmonic then rf is $\mathcal{U}$-superharmonic for each $r \geq 1$, and if f is $\mathcal{U}$-subharmonic then rf is $\mathcal{U}$-superharmonic for each $0 \leq r \leq 1$.*

Proof If $\mathcal{X}$ is a branching process and $\mathcal{U}$ is the log-Laplace operator of the transition law from $\mathcal{X}_0$ to $\mathcal{X}_1$ then, using Jensen's inequality, for all $r \geq 1$,

$$e^{-\langle \mu, \mathcal{U}(rf)\rangle} = E^\mu\big[e^{-\langle \mathcal{X}_1, rf\rangle}\big] = E^\mu\big[\big(e^{-\langle \mathcal{X}_1, f\rangle}\big)^r\big] \geq \big(E^\mu\big[e^{-\langle \mathcal{X}_1, f\rangle}\big]\big)^r = e^{-\langle \mu, r\mathcal{U}f\rangle}. \tag{7.8}$$

Since this holds for all $\mu \in \mathcal{M}(E)$, it follows that $\mathcal{U}(rf) \leq r\mathcal{U}f$. The proof of the statements for $0 \leq r \leq 1$ is the same but with the inequality signs reversed. ∎

We next turn our attention to the functions $h_{1,1}$ and $h_{0,0}$.

Lemma 7.3 (The catalyzing function $h_{1,1}$) *One has*

$$\mathcal{U}_\gamma(rh_{1,1})(x) = \frac{1+\gamma}{\frac{1}{r}+\gamma} \qquad (\gamma, r > 0,\ x \in [0,1]). \tag{7.9}$$

In particular, $h_{1,1}$ is $\mathcal{U}_\gamma$-harmonic for each $\gamma > 0$.

Proof Recall (3.7)–(3.9). Let $\sigma_{1/r}$ be an exponentially distributed random variable with mean $1/r$, independent of τ_γ. Then

$$\mathcal{U}_\gamma(rh_{1,1})(x) = (\tfrac{1}{\gamma}+1)E\big[1-e^{-\int_0^{\tau_\gamma} r\mathrm{d}t}\big] = (\tfrac{1}{\gamma}+1)P[\sigma_{1/r}<\tau_\gamma] = (\tfrac{1}{\gamma}+1)\frac{\gamma}{\frac{1}{r}+\gamma}, \tag{7.10}$$

which yields (7.9). ∎

Lemma 7.4 (The catalyzing function $h_{0,0}$) *One has $\mathcal{U}_\gamma(rh_{0,0})\leq rh_{0,0}$ for each $\gamma,r>0$.*

Proof Let Γ^γ_x be the invariant law from Corollary 5.4. Then, for any $\gamma>0$ and $f\in B_+[0,1]$,

$$\begin{aligned}\mathcal{U}_\gamma f(x) &= (\tfrac{1}{\gamma}+1)E\big[1-e^{-\langle\mathcal{Z}^\gamma_x,\,f\rangle}\big]\leq(\tfrac{1}{\gamma}+1)E[\langle\mathcal{Z}^\gamma_x,f\rangle]\\ &= (\tfrac{1}{\gamma}+1)E\Big[\int_0^{\tau_\gamma} f(\mathbf{y}^\gamma_x(-t/2))\,\mathrm{d}t\Big] = (1+\gamma)\langle\Gamma^\gamma_x,f\rangle \qquad (x\in[0,1]),\end{aligned} \tag{7.11}$$

where we have used that τ_γ is independent of $\mathbf{y}^\gamma_x$ and has mean γ. In particular, setting $f=rh_{0,0}$ and using (5.25) we find that $\mathcal{U}_\gamma(rh_{0,0})\leq rh_{0,0}$. ∎

The aim of the remainder of this section is to derive various bounds on $\mathcal{U}_\gamma f$ for $f\in\mathcal{H}_{0,1}$. We start with a formula for $\mathcal{U}_\gamma f$ that holds for general $[0,1]$-valued functions f.

Lemma 7.5 (Action of $\mathcal{U}_\gamma$ on $[0,1]$-valued functions) *Let $\mathbf{y}^\gamma_x$ be the stationary solution to (3.6) and let $\tau_{\gamma/2}$ be an independent exponentially distributed random variable with mean $\gamma/2$. Let $(\beta_i)_{i\geq1}$ be independent exponentially distributed random variables with mean $\frac{1}{2}$, independent of $\mathbf{y}^\gamma_x$ and $\tau_{\gamma/2}$, and let $\sigma_k:=\sum_{i=1}^k\beta_i$ $(k\geq0)$. Then*

$$1-\mathcal{U}_\gamma f(x) = E\Big[\prod_{k\geq0:\ \sigma_k<\tau_\gamma}\big(1-f(\mathbf{y}^\gamma_x(-\sigma_k))\big)\Big] \qquad (\gamma>0,\ f\in B_{[0,1]}[0,1],\ x\in[0,1]). \tag{7.12}$$

Proof By Lemma 7.3, the constant function $h_{1,1}(x):=1$ satisfies $\mathcal{U}_\gamma h_{1,1}=h_{1,1}$ for all $\gamma>0$. Therefore, by Proposition 3.4, Poissonizing the Poisson-cluster branching process $\mathcal{X}$ with the density $h_{1,1}$ yields a branching particle system $X^{h_{1,1}}=(X^{h_{1,1}}_{-n},\ldots,X^{h_{1,1}}_0)$ with generating operators $U^{h_{1,1}}_{\gamma_{n-1}},\ldots,U^{h_{1,1}}_{\gamma_0}$, where

$$U^{h_{1,1}}_\gamma f = \mathcal{U}_\gamma f \qquad (f\in B_{[0,1]}[0,1],\ \gamma>0). \tag{7.13}$$

By (3.18) and (7.5),

$$U^{h_{1,1}}_\gamma f(x) = 1-E\big[(1-f)^{\mathrm{Pois}(\mathcal{Z}^\gamma_x)}\,\big|\,\mathrm{Pois}(\mathcal{Z}^\gamma_x)\neq0\big] \quad (f\in B_{[0,1]}[0,1],\ x\in[0,1],\ \gamma>0). \tag{7.14}$$

Therefore, (7.12) will follow provided that

$$P\big[\mathrm{Pois}(\mathcal{Z}^\gamma_x)\in\cdot\,\big|\,\mathrm{Pois}(\mathcal{Z}^\gamma_x)\neq0\big] = \mathcal{L}\Big(\sum_{k\geq0:\ \sigma_k<\tau_{\gamma/2}}\delta_{\mathbf{y}^\gamma_x(-\sigma_k)}\Big). \tag{7.15}$$

Indeed, it is not hard to see that

$$\mathrm{Pois}(\mathcal{Z}^\gamma_x) \stackrel{\mathcal{D}}{=} \sum_{k>0:\ \sigma_k<\tau_{\gamma/2}}\delta_{\mathbf{y}^\gamma_x(-\sigma_k)}. \tag{7.16}$$

This follows from the facts that $\mathcal{Z}^\gamma_x = 2\int_0^{\tau_{\gamma/2}} \delta_{\mathbf{y}^\gamma_x(-s)}\mathrm{d}s$ and

$$\sum_{k>0:\,\sigma_k<\tau_{\gamma/2}} \delta_{-\sigma_k} \stackrel{\mathcal{D}}{=} \mathrm{Pois}(2\,1_{(-\tau_{\gamma/2},0]}). \tag{7.17}$$

Conditioning $\mathrm{Pois}(2\,1_{(-\tau_{\gamma/2},0]})$ on being nonzero means conditioning on $\tau_{\gamma/2} > \sigma_1$. Since $\tau_{\gamma/2} - \sigma_1$, conditioned on being nonnegative, is exponentially distributed with mean $\gamma/2$, using the stationarity of $\mathbf{y}^\gamma_x$, we arrive at (7.15). ∎

The next lemma generalizes the duality (5.22) to mixed moments of the Wright-Fisher diffusion $\mathbf{y}$ at multiple times. We can interpret the left-hand side of (7.18) as the probability that $m_1, \ldots, m_n$ organisms sampled from the population at times $t_1, \ldots, t_n$ are all of the genetic type I.

Lemma 7.6 (Sampling at multiple times) *Fix $0 \le t_1 < \cdots < t_n = t$ and nonnegative integers $m_1, \ldots, m_n$. Let $\mathbf{y}$ be the diffusion in (5.20). Then*

$$E^y\Big[\prod_{k=1}^n \mathbf{y}_{t_k}^{m_k}\Big] = E\big[y^{\phi_t} x^{\psi_t}\big], \tag{7.18}$$

where $(\phi_s, \psi_s)_{s\in[0,t]}$ is a Markov process in $\mathbb{N}^2$ started in $(\phi_0, \psi_0) = (m_n, 0)$, that jumps deterministically as

$$(\phi_s, \psi_s) \to (\phi_s + m_k, \psi_s) \quad \text{at time} \quad t - t_k \quad (k < n), \tag{7.19}$$

and between these deterministic times jumps with rates as in (5.21).

Proof Induction, with repeated application of (5.22). ∎

For any $m \ge 1$, we put

$$h_m(x) := 1 - (1-x)^m \qquad (x \in [0,1]). \tag{7.20}$$

The next lemma shows that we have particular good control on the action of $\mathcal{U}_\gamma$ on the functions h_m.

Lemma 7.7 (Action of $\mathcal{U}_\gamma$ on the functions h_m) *Let $m \ge 1$ and let τ_γ be an exponentially distributed random variable with mean γ. Conditional on τ_γ, let $(\phi'_t, \psi'_t)_{t\ge0}$ be a Markov process in $\mathbb{N}^2$, started in $(\phi'_0, \psi'_0) = (m, 0)$ that jumps at time t as:*

$$\begin{array}{ll} (\phi'_t, \psi'_t) \to (\phi'_t - 1, \psi'_t) & \textit{with rate } \phi'_t(\phi'_t - 1), \\ (\phi'_t, \psi'_t) \to (\phi'_t - 1, \psi'_t + 1) & \textit{with rate } \frac{1}{\gamma}\phi'_t, \\ (\phi'_t, \psi'_t) \to (\phi'_t + m, \psi'_t) & \textit{with rate } 1_{\{\tau_{\gamma/2} < t\}}. \end{array} \tag{7.21}$$

Then the limit $\lim_{t\to\infty} \psi'_t =: \psi'_\infty$ exists a.s., and

$$\mathcal{U}_\gamma h_m(x) = E^{(m,0)}\big[1 - (1-x)^{\psi'_\infty}\big] \qquad (m \ge 1,\ x \in [0,1]). \tag{7.22}$$

Proof Let $\mathbf{y}^\gamma_x$, $\tau_{\gamma/2}$, and $(\sigma_k)_{k\ge0}$ be as in Lemma 7.5. Then, by (7.12),

$$\mathcal{U}_\gamma h_m(x) = 1 - E\Big[\prod_{k\ge0:\,\sigma_k<\tau_{\gamma/2}} \big(1 - \mathbf{y}^\gamma_x(-\sigma_k)\big)^m\Big]. \tag{7.23}$$

Let $(\phi',\psi')=(\phi'_t,\psi'_t)_{t\geq 0}$ be a $\mathbb{N}^2$-valued process started in $(\phi'_0,\psi'_0)=(m,0)$ such that conditioned on τ_γ and $(\sigma_k)_{k\geq 0}$, (ϕ',ψ') is a Markov process that jumps deterministically as

$$(\phi'_t,\psi'_t)\to(\phi'_t+m,\psi'_s)\quad\text{at time}\quad\sigma_k\quad(k\geq 1:\ \sigma_k<\tau_{\gamma/2}) \tag{7.24}$$

and between these times jumps with rates as in (5.21). Then $(\phi'_t,\psi'_t)\to(0,\psi'_\infty)$ as $t\to\infty$ a.s. for some $\mathbb{N}$-valued random variable ψ'_∞, and (7.22) follows from Lemma 7.6, using the symmetry $y\leftrightarrow 1-y$. Since $\sigma_{k+1}-\sigma_k$ are independent exponentially distributed random variables with mean one, (ϕ',ψ') is the Markov process with jump rates as in (7.21). ∎

The next result is a simple application of Lemma 7.7.

Lemma 7.8 (The catalyzing function h_1) *The function $h_1(x):=x$ $(x\in[0,1])$ is $\mathcal{U}_\gamma$-subharmonic for each $\gamma>0$.*

Proof Since $\psi'_\infty\geq 1$ a.s., one has $1-(1-x)^{\psi'_\infty}\geq x$ a.s. $(x\in[0,1])$ in (7.22). In particular, setting $m=1$ yields $\mathcal{U}_\gamma h_1\geq h_1$. ∎

We now set out to prove that h_7, which is the function $h_{0,1}$ from Lemma 3.8, is $\mathcal{U}_\gamma$-superharmonic. In order to do so, we will derive upper bounds on the expectation of ψ'_∞. We derive two estimates: one that is good for small γ and one that is good for large γ.

In order to avoid tedious formal arguments, it will be convenient to recall the interpretation of the process (ϕ',ψ') and Lemma 7.6. Recall from the discussion following (5.22) that $(\mathbf{y}^\gamma_x(t))_{t\in\mathbb{R}}$ describes the equilibrium frequency of genetic type I as a function of time in a population that is in genetic exchange with an infinite reservoir. From this population we sample at times $-\sigma_k$ $(k\geq 0,\ \sigma_k<\tau_{\gamma/2})$ each time m individuals, and ask for the probability that they are not all of the genetic type II. In order to find this probability, we follow the ancestors of the sampled individuals back in time. Then ϕ'_t and ψ'_t are the number of ancestors that lived at time $-t$ in the population and the reservoir, respectively, and $E[1-(1-x)^{\psi'_\infty}]$ is the probability that at least one ancestor is of type I.

Lemma 7.9 (Bound for small γ) *For each $\gamma\in(0,\infty)$ and $m\geq 1$,*

$$\frac{1}{m}E^{(m,0)}[\psi'_\infty]\leq\frac{1}{m}\sum_{i=0}^{m-1}\frac{1+\gamma}{1+i\gamma}=:\chi_m(\gamma). \tag{7.25}$$

The function χ_m is concave and satisfies $\chi_m(0)=1$ for each $m\geq 1$.

Proof Note that

$$E\big[\big|\{k\geq 0:\ \sigma_k<\tau_{\gamma/2}\}\big|\big]=1+\gamma. \tag{7.26}$$

We can estimate (ϕ',ψ') from above by a process where ancestors from individuals sampled at different times cannot coalesce. Therefore,

$$E^{(m,0)}[\psi'_\infty]\leq(1+\gamma)E^{(m,0)}[\psi_\infty], \tag{7.27}$$

where (ϕ,ψ) is the Markov process in (5.21). Note that if (ϕ,ψ) is in the state $(m+1,0)$, then the next jump is to $(m,1)$ with probability

$$\frac{\frac{1}{\gamma}(m+1)}{\frac{1}{\gamma}(m+1)+m(m+1)}=\frac{1}{1+m\gamma} \tag{7.28}$$

and to $(m,0)$ with one minus this probability. Therefore,

$$\begin{aligned}E^{(m+1,0)}[\psi_\infty] &= \frac{1}{1+m\gamma}E^{(m,1)}[\psi_\infty] + \Big(1-\frac{1}{1+m\gamma}\Big)E^{(m,0)}[\psi_\infty] \\ &= \frac{1}{1+m\gamma}\Big(E^{(m,0)}[\psi_\infty]+1\Big) + \Big(1-\frac{1}{1+m\gamma}\Big)E^{(m,0)}[\psi_\infty] \\ &= E^{(m,0)}[\psi_\infty] + \frac{1}{1+m\gamma}.\end{aligned} \tag{7.29}$$

By induction, it follows that

$$E^{(m,0)}[\psi_\infty] = \sum_{i=0}^{m-1}\frac{1}{1+i\gamma}. \tag{7.30}$$

Inserting this into (7.27) we arrive at (7.25). Finally, since

$$\frac{\partial^2}{\partial\gamma^2}\frac{1+\gamma}{1+i\gamma} = \frac{2i(i-1)}{(1+i\gamma)^3} \geq 0 \qquad (i\geq 0,\ \gamma\geq 0), \tag{7.31}$$

the function χ_m is convex. ∎

Lemma 7.10 (Bound for large γ) *For each $\gamma\in(0,\infty)$ and $m\geq 1$,*

$$E^{(m,0)}[\psi'_\infty] \leq (\tfrac{1}{\gamma}+1)\sum_{k=1}^{m}\frac{1}{k}+\frac{3}{2}. \tag{7.32}$$

Proof We start by observing that $\frac{\partial}{\partial t}E[\psi'_t] = \frac{1}{\gamma}E[\phi'_t]$, and therefore

$$E[\psi'_\infty] = \tfrac{1}{\gamma}\int_0^\infty E[\phi'_t]\mathrm{d}t. \tag{7.33}$$

Unlike in the proof of the last lemma, this time we cannot fully ignore the coalescence of ancestors sampled at different times. In order to deal with this we use a trick: at time zero we introduce an extra ancestor that can only jump to the reservoir when $t\geq\tau_\gamma$ and there are no other ancestors left in the population. We further assume that all other ancestors do not jump to the reservoir on their own. Let ξ_t be one as long as this extra ancestor is in the population and zero otherwise, and let ϕ''_t be the number of other ancestors in the population according to these new rules. Then we have at a Markov process (ξ,ϕ'') started in $(\xi_0,\phi''_0)=(1,m)$ that jumps as:

$$\begin{array}{ll}(\xi_t,\phi''_t)\to(\xi_t,\phi''_t-1) & \text{with rate } (\phi''_t+1)\phi''_t, \\ (\xi_t,\phi''_t)\to(\xi_t,\phi''_t+m) & \text{with rate } 1_{\{\tau_{\gamma/2}<t\}}, \\ (\xi_t,\phi''_t)\to(\xi_t-1,\phi''_t) & \text{with rate } \frac{1}{\gamma}1_{\{\tau_{\gamma/2}\geq t\}}1_{\{\phi''_t=0\}}.\end{array} \tag{7.34}$$

It is not hard to show that (ξ,ϕ'') and ϕ' can be coupled such that $\xi_t+\phi''_t\geq\phi'_t$ for all $t\geq 0$. We now simplify even further and ignore all coalescence between ancestors belonging to the process ϕ'' that are introduced at different times. Let $\phi^{(k)}_t$ be the number of ancestors in the population that were introduced at the time σ_k $(k\geq 0)$. Thus, for $t<\sigma_k$ one has $\phi^{(k)}_t=0$, for $t=\sigma_k$ one has $\phi^{(k)}_t=m$, while for $t>\sigma_k$, the process $\phi^{(k)}_t$ jumps from n to $n-1$ with rate

$(n+1)n$. Then it is not hard to see that, for an appropriate coupling, $\phi''_t \le \sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} \phi^{(k)}_t$ for all $t \ge 0$. We let ξ' be a process such that $\xi'_0 = 1$ and ξ'_t jumps to zero with rate

$$\frac{1}{\gamma} 1_{\{\tau_{\gamma/2}\ge t\}} \prod_{k\ge 0:\sigma_k<\tau_{\gamma/2}} 1_{\{\phi^{(k)}_t=0\}}. \tag{7.35}$$

Then for an appropriate coupling $\xi'_t \ge \xi_t$ $(t \ge 0)$. Thus, we can estimate

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \le \int_0^\infty E[\xi'_t]\mathrm{d}t + \int_0^\infty E\Big[\sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} \phi^{(k)}_t\Big]\mathrm{d}t. \tag{7.36}$$

Set $\rho := \inf\{t \ge \tau_{\gamma/2} : \phi^{(k)}_t = 0\ \forall k \ge 0 \text{ with } \sigma_k < \tau_{\gamma/2}\}$ and $\pi := \inf\{t \ge 0 : \xi'_t = 0\}$. Then

$$\int_0^\infty E[\xi'_t]\mathrm{d}t = E[\tau_{\gamma/2}] + E[\rho - \tau_{\gamma/2}] + E[\pi - \rho] = \frac{3}{2}\gamma + E[\rho - \tau_{\gamma/2}]. \tag{7.37}$$

Since

$$\begin{aligned} E[\rho - \tau_{\gamma/2}] &\le \int_0^\infty E\Big[1_{\{\sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} \phi^{(k)}_t \ne 0\}}\Big]\mathrm{d}t \\ &\le \int_0^\infty E\Big[\sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} 1_{\{\phi^{(k)}_t\ne 0\}}\Big]\mathrm{d}t, \end{aligned} \tag{7.38}$$

using moreover (7.36) and (7.37), we can estimate

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \le \frac{3}{2}\gamma + \int_0^\infty E\Big[\sum_{k\ge 0:\sigma_k<\tau_{\gamma/2}} (\phi^{(k)}_t + 1_{\{\phi^{(k)}_t\ne 0\}})\Big]\mathrm{d}t. \tag{7.39}$$

Since $E\big[\big|\{k \ge 0 :\ \sigma_k < \tau_{\gamma/2}\}\big|\big] = 1 + \gamma$, we obtain

$$\int_0^\infty E[\phi'_t]\mathrm{d}t \le \frac{3}{2}\gamma + (1+\gamma)\int_0^\infty E[\phi^{(0)}_t + 1_{\{\phi^{(0)}_t\ne 0\}}]\mathrm{d}t. \tag{7.40}$$

Since $\phi^{(0)}_t$ jumps from n to $n-1$ with rate $(n+1)n$, the expected total time that $\phi^{(0)}_t = n$ equals $1/((n+1)n)$, and therefore

$$\int_0^\infty E[\phi^{(0)}_t + 1_{\{\phi^{(0)}_t\ne 0\}}]\mathrm{d}t = \sum_{n=1}^m \frac{1}{(n+1)n}(n + 1_{\{n\ne 0\}}) = \sum_{n=1}^m \frac{1}{n}. \tag{7.41}$$

Inserting this into (7.40), using (7.33), we arrive at (7.32). ∎

Lemma 7.11 (The catalyzing function $h_{0,1}$) *One has $\mathcal{U}_\gamma(h_{0,1}) \le h_{0,1}$ for each $\gamma > 0$. Moreover, for each $r > 1$ and $\gamma > 0$,*

$$\sup_{x\in(0,1]} \frac{\mathcal{U}_\gamma(rh_{0,1})(x)}{rh_{0,1}(x)} < 1. \tag{7.42}$$

Proof Recall that $h_{0,1}(x) = h_7(x) = 1 - (1-x)^7$ $(x \in [0,1])$. We will show that

$$E^{(7,0)}[\psi'_\infty] < 7 \tag{7.43}$$

for each $\gamma \in (0,\infty)$. The function $\chi_m(\gamma)$ from Lemma 7.9 satisfies

$$\chi_m(1) = \frac{1}{m}\sum_{n=1}^{m}\frac{2}{n} < 1 \qquad (m \geq 5). \tag{7.44}$$

Since $\chi_m(\gamma)$ is concave in γ and satisfies $\chi_m(0) = 1$, it follows that $\chi_m(\gamma) < 1$ for all $0 < \gamma \leq 1$ and $m \geq 5$. By Lemma 7.10, for all $\gamma \geq 1$,

$$E^{(m,0)}[\psi'_\infty] \leq 2\sum_{k=1}^{m}\frac{1}{k} + \frac{3}{2} < m \qquad (m \geq 7). \tag{7.45}$$

Therefore, if $m \geq 7$, then $m' := E^{(m,0)}[\psi'_\infty] < m$. It follows by (7.22) and Jensen's inequality applied to the concave function $z \mapsto 1-(1-x)^z$ that

$$\mathcal{U}_\gamma h_m(x) \leq 1-(1-x)^{E^{(m,0)}[\psi'_\infty]} = 1-(1-x)^{m'} \leq h_m(x) \qquad (x \in [0,1],\ \gamma > 0). \tag{7.46}$$

This shows that h_m is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$. By Lemma 7.2, for each $r > 1$,

$$\frac{\mathcal{U}_\gamma(rh_m)(x)}{rh_m(x)} \leq \frac{r\mathcal{U}_\gamma(h_m)(x)}{rh_m(x)} \leq \frac{1-(1-x)^{m'}}{1-(1-x)^m} \qquad (x \in (0,1]). \tag{7.47}$$

By Lemma 7.3 and the monotonicity of $\mathcal{U}_\gamma$,

$$\frac{\mathcal{U}_\gamma(rh_m)(x)}{rh_m(x)} \leq \frac{\mathcal{U}_\gamma(r)(x)}{rh_m(x)} \leq \frac{1+\gamma}{1+r\gamma}\frac{1}{(1-(1-x)^m)} \qquad (x \in (0,1]). \tag{7.48}$$

Since the right-hand side of (7.47) is smaller than 1 for $x \in (0,1)$ and tends to $m'/m < 1$ as $x \to 0$, since the right-hand side of (7.48) is smaller than 1 for x in an open neighborhood of 1, and since both bounds are continuous, (7.42) follows. ∎

7.3 Extinction versus unbounded growth

In this section we show that Lemmas 3.6–3.8 are equivalent to Proposition 3.9. (This follows from the equivalence of conditions (i) and (ii) in Lemma 7.12 below.) We moreover prove Lemmas 3.6 and 3.8 and prepare for the proof of Lemma 3.7. We start with some general facts about log-Laplace operators and branching processes.

For the next lemma, let E be a separable, locally compact, metrizable space. For $n \geq 0$, let $q_n \in \mathcal{C}_+(E)$ be continuous weight functions, let $\mathcal{Q}_n$ be continuous cluster mechanisms on E, and assume that the associated log-Laplace operators $\mathcal{U}_n$ defined in (3.2) satisfy (3.3). Assume that $0 \neq h \in \mathcal{C}_+(E)$ is bounded and $\mathcal{U}_n$-superharmonic for all n, let $E^h := \{x \in E : h(x) > 0\}$, and define generating operators $U^h_n : B_{[0,1]}(E^h) \to B_{[0,1]}(E)$ as in (3.25). For each $n \geq 0$, let $(\mathcal{X}^{(n)}_0, \mathcal{X}^{(n)}_1)$ be a one-step Poisson cluster branching process with log-Laplace operator $\mathcal{U}_n$, and let $(X^{(n),h}_0, X^{(n),h}_1)$ be the one-step branching particle system with generating operator U^h_n. (In a typical application of this lemma, the operators $\mathcal{U}_n$ will be iterates of other log-Laplace operators, and $\mathcal{X}^{(n)}_0, \mathcal{X}^{(n)}_1$ will be the initial and final state, respectively, of a Poisson cluster branching process with many time steps.)

Lemma 7.12 (Extinction versus unbounded growth) *Assume that* $\rho \in \mathcal{C}_{[0,1]}(E^h)$ *and put*

$$p(x) := \begin{cases} h(x)\rho(x) & \text{if } x \in E^h, \\ 0 & \text{if } x \in E \backslash E^h. \end{cases} \tag{7.49}$$

Then the following statements are equivalent:

$$\begin{aligned} &\text{(i)} \quad P^{\delta_x}\big[|X_1^{(n),h}| \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \rho(x)\delta_\infty + (1-\rho(x))\delta_0 \\ &\qquad\qquad \textit{locally uniformly for } x \in E^h, \\ &\text{(ii)} \quad P^{\delta_x}\big[\langle \mathcal{X}_1^{(n)}, h\rangle \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} e^{-p(x)}\delta_0 + \big(1-e^{-p(x)}\big)\delta_\infty \\ &\qquad\qquad \textit{locally uniformly for } x \in E, \\ &\text{(iii)} \quad \mathcal{U}_n(\lambda h)(x) \underset{n\to\infty}{\longrightarrow} p(x) \\ &\qquad\qquad \textit{locally uniformly for } x \in E \quad \forall \lambda > 0, \\ &\text{(iv)} \quad \exists 0 < \lambda_1 < \lambda_2 < \infty : \quad \mathcal{U}_n(\lambda_i h)(x) \underset{n\to\infty}{\longrightarrow} p(x) \\ &\qquad\qquad \textit{locally uniformly for } x \in E \qquad (i = 1, 2). \end{aligned}$$

Proof of Lemma 7.12 It is not hard to see that (i) is equivalent to

$$P^{\delta_x}[\mathrm{Thin}_\lambda(X_1^{(n),h}) \neq 0] \underset{n\to\infty}{\longrightarrow} \rho(x) \tag{7.50}$$

locally uniformly for $x \in E^h$, for all $0 < \lambda \leq 1$. It follows from (3.19) and (3.25) that $h(x)P^{\delta_x}[\mathrm{Thin}_\lambda(X_1^{(n),h}) \neq 0] = hU^h(\lambda)(x) = \mathcal{U}(\lambda h)(x)$ $(x \in E)$, so (i) is equivalent to

$$\begin{aligned} &\text{(i)}' \quad \mathcal{U}_n(\lambda h)(x) \underset{n\to\infty}{\longrightarrow} p(x) \\ &\qquad\qquad \text{locally uniformly for } x \in E \quad \forall 0 < \lambda \leq 1. \end{aligned}$$

By (3.4), condition (ii) implies that

$$e^{-\mathcal{U}_n(\lambda h)(x)} = E^{\delta_x}\big[e^{-\lambda\langle \mathcal{X}_1, h\rangle}\big] \underset{n\to\infty}{\longrightarrow} e^{-p(x)} \tag{7.51}$$

locally uniformly for $x \in E$ for all $\lambda > 0$, and therefore (ii) implies (iii). Obviously (iii)$\Rightarrow$ (i)$'$ $\Rightarrow$(iv) so we are done if we show that (iv)$\Rightarrow$(ii). Indeed, (iv) implies that

$$E^{\delta_x}\big[e^{-\lambda_1\langle \mathcal{X}_1^{(n)}, h\rangle} - e^{-\lambda_2\langle \mathcal{X}_1^{(n)}, h\rangle}\big] \underset{n\to\infty}{\longrightarrow} 0 \tag{7.52}$$

locally uniformly for $x \in E$, which shows that

$$P^{\delta_x}\big[c < \langle \mathcal{X}_1^{(n)}, h\rangle < C\big] \underset{n\to\infty}{\longrightarrow} 0 \tag{7.53}$$

for all $0 < c < C < \infty$. Using (iv) once more we arive at (ii). ∎

Our next lemma gives sufficient conditions for the n-th iterates of a single log-Laplace operator $\mathcal{U}$ to satisfy the equivalent conditions of Lemma 7.12. Let E (again) be separable, locally compact, and metrizable. Let $q \in \mathcal{C}_+(E)$ be a weight function, $\mathcal{Q}$ a continuous cluster mechanism on E, and assume that the associated log-Laplace operator $\mathcal{U}$ defined in (3.2) satisfies (3.3). Let $\mathcal{X} = (\mathcal{X}_0, \mathcal{X}_1, \ldots)$ be the Poisson-cluster branching process with log-Laplace operator $\mathcal{U}$ in each step, let $0 \neq h \in \mathcal{C}_+(E)$ be bounded and $\mathcal{U}$-superharmonic, and let $X^h = (X_0^h, X_1^h, \ldots)$ denote the branching particle system on E^h obtained from $\mathcal{X}$ by Poissonization with a $\mathcal{U}$-superharmonic function h, in the sense of Proposition 3.4.

Lemma 7.13 (Sufficient condition for extinction versus unbounded growth) *Assume that*

$$\sup_{x\in E^h} \frac{\mathcal{U}h(x)}{h(x)} < 1. \tag{7.54}$$

Then the process X^h started in any initial law $\mathcal{L}(X^h_0) \in \mathcal{M}_1(E^h)$ satisfies

$$\lim_{k\to\infty} |X^h_k| = \infty \quad \text{or} \quad \exists k \geq 0 \text{ s.t. } X^h_k = 0 \qquad \text{a.s.} \tag{7.55}$$

Moreover, if the function $\rho : E^h \to [0,1]$ defined by

$$\rho(x) := P^{\delta_x}[X^h_n \neq 0 \quad \forall n \geq 0] \qquad (x \in E^h) \tag{7.56}$$

satisfies $\inf_{x\in E^h} \rho(x) > 0$, then ρ is continuous.

Proof of Lemma 7.13 Let $\mathcal{A}$ denote the tail event $\mathcal{A} = \{X^h_n \neq 0 \ \forall n \geq 0\}$ and let $(\mathcal{F}_k)_{k\geq 0}$ be the filtration generated by X^h. Then, by the Markov property and continuity of the conditional expectation with respect to increasing limits of σ-fields (see Complement 10(b) from [Loe63, Section 29] or [Loe78, Section 32])

$$P[X^h_n \neq 0 \ \forall n \geq 0 | X_k] = P(\mathcal{A}|\mathcal{F}_k) \underset{k\to\infty}{\longrightarrow} 1_{\mathcal{A}} \qquad \text{a.s.} \tag{7.57}$$

In particular, this implies that a.s. on the event $\mathcal{A}$ one must have $P[X^h_{k+1} = 0 | X^h_k] \to 0$ a.s. By (3.19) and (3.25), $P^{\delta_x}[X^h_1 \neq 0] = U^h 1(x) = (\mathcal{U}h(x))/h(x)$, which is uniformly bounded away from one by (7.54). Therefore, $P[X^h_{k+1} = 0 | X^h_k] \to 0$ a.s. on $\mathcal{A}$ is only possible if the number of particles tends to infinity.

The continuity of ρ can be proved by a straightforward adaptation of the proof of [FS04, Proposition 5 (d)] to the present setting with discrete time and noncompact space E. An essential ingredient in the proof, apart from (7.54), is the fact that the map $\nu \mapsto P^\nu[X^h_n \in \cdot]$ from $\mathcal{N}(E)$ to $\mathcal{M}_1(\mathcal{N}(E))$ is continuous, which follows from the continuity of Q^h. ∎

We now turn our attention more specifically to the renormalization branching process $\mathcal{X}$. In the remainder of this section, $(\gamma_k)_{k\geq 0}$ is a sequence of positive constants such that $\sum_n \gamma_n = \infty$ and $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$, and $\mathcal{X} = (\mathcal{X}_{-n}, \ldots, \mathcal{X}_0)$ is the Poisson cluster branching process on $[0,1]$ defined in Section 3.2. We put $\mathcal{U}^{(n)} := \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}$. If $0 \neq h \in \mathcal{C}[0,1]$ is $\mathcal{U}_{\gamma_k}$-superharmonic for all $k \geq 0$, then $\mathcal{X}^h$ and X^h denote the branching process and the branching particle system on $\{x \in [0,1] : h(x) > 0\}$ obtained from $\mathcal{X}$ by weighting and Poissonizing with h in the sense of Propositions 3.3 and 3.4, respectively.

Proof of Lemma 3.6 By induction, it follows from Lemma 7.3 that

$$\mathcal{U}^{(n)}(\lambda h_{1,1}) = \frac{\prod_{k=0}^{n-1}(1+\gamma_k)}{\prod_{k=0}^{n-1}(1+\gamma_k) - 1 + \frac{1}{\lambda}} \qquad (\lambda > 0). \tag{7.58}$$

It is not hard to see (compare the footnote at (2.12)) that

$$\prod_{k=0}^{\infty}(1+\gamma_k) = \infty \quad \text{if and only if} \quad \sum_{k=0}^{\infty} \gamma_k = \infty. \tag{7.59}$$

Therefore, since we are assuming that $\sum_n \gamma_n = \infty$,

$$\mathcal{U}^{(n)}(\lambda h_{1,1}) \underset{n\to\infty}{\longrightarrow} h_{1,1}, \tag{7.60}$$

uniformly on $[0,1]$ for all $\lambda > 0$. The result now follows from Lemma 7.12 (with $h = h_{1,1}$ and $\rho(x) = 1$ $(x \in [0,1])$). ■

Remark 7.14 (Conditions on $(\gamma_n)_{n\geq 0}$) Our proof of Lemma 3.6 does not use that $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$. On the other hand, the proof shows that $\sum_n \gamma_n = \infty$ is a necessary condition for (3.29). ◇

We do not know if the assumption that $\gamma_n \to \gamma^*$ for some $\gamma^* \in [0,\infty)$ is needed in Lemma 3.7. We guess that it can be dropped, but it will greatly simplify proofs to have it around.

We will show that in order to prove Lemmas 3.7 and 3.8, it suffices to prove their analogues for embedded particle systems in the time-homogeneous processes $\mathcal{Y}^{\gamma^*}$ $(\gamma^* \in [0,\infty))$. More precisely, we will derive Lemmas 3.7 and 3.8 from the following two results. Below, $(\mathcal{U}^0_t)_{t\geq 0}$ is the log-Laplace semigroup of the super-Wright-Fisher diffusion $\mathcal{Y}^0$, defined in (3.15). The functions $p^*_{0,1,\gamma^*}$ $(\gamma^* \in [0,\infty))$ are defined in (3.34).

Proposition 7.15 (Time-homogeneous embedded particle system with $h_{0,0}$)
(a) *For any $\gamma^* > 0$, one has $(\mathcal{U}_{\gamma^*})^n h_{0,0} \underset{n\to\infty}{\longrightarrow} 0$ uniformly on $[0,1]$.*
(b) *One has $\mathcal{U}^0_t h_{0,0} \underset{t\to\infty}{\longrightarrow} 0$ uniformly on $[0,1]$.*

Proposition 7.16 (Time-homogeneous embedded particle system with $h_{0,1}$)
(a) *For any $\gamma^* > 0$, one has $(\mathcal{U}_{\gamma^*})^n(\lambda h_{0,1}) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}$ uniformly on $[0,1]$, for all $\lambda > 0$.*
(b) *One has $\mathcal{U}^0_t(\lambda h_{0,1}) \underset{t\to\infty}{\longrightarrow} p^*_{0,1,0}$ uniformly on $[0,1]$, for all $\lambda > 0$.*

Proposition 7.15 (a) will be proved in Section 8.2.

Proof of Proposition 7.16 (a) By formula (7.42) from Lemma 7.11, for each $r > 1$ the function $rh_{0,1}$ satisfies condition (7.54) from Lemma 7.13. Set $\rho(x) := P^{\delta_x}[Y^{\gamma^*,rh_{0,1}}_n \neq 0\ \forall n]$. Then, by (3.19) and (3.25),

$$\begin{aligned}\rho(x) &= \lim_{n\to\infty} P^{\delta_x}[Y^{\gamma^*,rh_{0,1}}_n \neq 0] = \lim_{n\to\infty}(U^{rh_{0,1}}_{\gamma^*})^n 1(x)\\ &= \lim_{n\to\infty}\frac{(\mathcal{U}_{\gamma^*})^n(rh_{0,1})(x)}{rh_{0,1}(x)} \geq \frac{h_1(x)}{rh_{0,1}(x)} \qquad (x \in (0,1]),\end{aligned} \tag{7.61}$$

where $h_1(x) = x$ $(x \in [0,1])$ is the $\mathcal{U}_{\gamma^*}$-subharmonic function from Lemma 7.8. It follows that $\inf_{x\in(0,1]}\rho(x) > 0$ and therefore, by Lemma 7.13, ρ is continuous in x.

By Lemma 7.13, we see that the Poissonized particle system $X^{rh_{0,1}}$ exhibits extinction versus unbounded growth in the sense of Lemma 7.12, which implies the statement in Proposition 7.16 (a). ■

Proof of Propositions 7.15 (b) and 7.16 (b) These statements follow from results in [FS03]. Indeed, [FS03, Proposition 2] implies that for any $f \in B_+[0,1]$ and $x \in [0,1]$,

$$\begin{aligned}&\mathcal{U}^0_t f(x) \underset{t\to\infty}{\longrightarrow} 0 && \text{if } f(0) = f(1) = 0,\\ &\mathcal{U}^0_t f(x) \underset{t\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}(x) && \text{if } f(0) = 0,\ f(1) > 0.\end{aligned} \tag{7.62}$$

To see that the convergence in (7.62) is in fact uniform in $x \in [0,1]$ we use the fact that each function $f \in B_+[0,1]$ with $f(0) = f(1) = 0$ can be bounded as $f \leq r1_{(0,1)}$ for some

$r \geq 1$, and that each function $f \in B_+[0,1]$ with $f(0) = 0$ and $f(1) > 0$ can be bounded as $\varepsilon 1_{\{1\}} \leq f \leq r1_{(0,1]}$ for some $0 < \varepsilon \leq 1$ and $r \geq 1$. Therefore, by the monotonity of $\mathcal{U}^0_t$, it suffices to show that $\mathcal{U}^0_t(r1_{(0,1)})$, $\mathcal{U}^0_t(r1_{(0,1]})$, and $\mathcal{U}^0_t(\varepsilon 1_{\{1\}})$ converge uniformly on $[0,1]$. By [FS03, Lemma 15], these functions are continuous for each $t > 0$, and since moreover the limit functions are continuous, it suffices to show that the convergence is monotone. Thus, we claim that

$$\begin{aligned} \mathcal{U}^0_t(r1_{(0,1)}) &\downarrow 0 && (r \geq 1), \\ \mathcal{U}^0_t(r1_{(0,1]}) &\downarrow p^*_{0,1,\gamma^*} && (r \geq 1), \\ \mathcal{U}^0_t(\varepsilon 1_{\{1\}}) &\uparrow p^*_{0,1,\gamma^*} && (0 < \varepsilon \leq 1). \end{aligned} \tag{7.63}$$

By (an obvious analogue of) Lemma 7.2, it suffices to show that $1_{(0,1)}$ and $1_{(0,1]}$ are $\mathcal{U}^0_t$-superharmonic, while $1_{\{1\}}$ is $\mathcal{U}^0_t$-subharmonic for each $t \geq 0$. Let $(\mathcal{Y}^{0,h_{1,1}}_t)_{t\geq 0}$ be the branching particle system obtained from $(\mathcal{Y}^0_t)_{t\geq 0}$ by Poissonization with the constant function $h_{1,1} := 1$. Then $\mathcal{Y}^{0,h_{1,1}}$ is a system of binary splitting Wright-Fisher diffusions, which was also studied in [FS03]. One has (compare (3.19))

$$\mathcal{U}^0_t 1_{(0,1)}(x) = P[\mathrm{Thin}_{\mathcal{U}^0_t 1_{(0,1)}}(\delta_x) \neq 0] = P^{\delta_x}[\mathrm{Thin}_{1_{(0,1)}}(Y^{0,h_{1,1}}_t) \neq 0] = P^{\delta_x}[Y^{0,h_{1,1}}_t((0,1)) > 0]. \tag{7.64}$$

Likewise,

$$\mathcal{U}^0_t 1_{(0,1]}(x) = P^{\delta_x}[Y^{0,h_{1,1}}_t((0,1]) > 0] \quad \text{and} \quad \mathcal{U}^0_t 1_{\{1\}}(x) = P^{\delta_x}[Y^{0,h_{1,1}}_t(\{1\}) > 0]. \tag{7.65}$$

Using the fact that the points $0,1$ are traps for the Wright-Fisher diffusion and that in a binary splitting Wright-Fisher diffusion, particles never die, it is easy to see that $P^{\delta_x}[Y^{0,h_{1,1}}_t((0,1)) > 0]$ and $P^{\delta_x}[Y^{0,h_{1,1}}_t((0,1]) > 0]$ are nonincreasing in t, while $P^{\delta_x}[Y^{0,h_{1,1}}_t(\{1\}) > 0]$ is nondecreasing in t. ∎

We now show that Propositions 7.15 and 7.16 imply Lemmas 3.7 and 3.8, respectively.

Proof of Lemma 3.7 We start with the proof that the embedded particle system $X^{h_{0,0}}$ is critical. For any $f \in B_+[0,1]$ and $k \geq 1$, we have, by Poissonization (Proposition 3.4) and the definition of $\mathcal{X}$,

$$\begin{aligned} h_{0,0}(x) E^{-k,\delta_x}[\langle X^{h_{0,0}}_{-k+1}, f\rangle] &= E^{-k,\mathcal{L}(\mathrm{Pois}(h_{0,0}\delta_x))}[\langle X^{h_{0,0}}_{-k+1}, f\rangle] = E^{-k,\delta_x}[\langle \mathrm{Pois}(h_{0,0}\mathcal{X}_{-k+1}), f\rangle] \\ &= E^{-k,\delta_x}[\langle \mathcal{X}_{-k+1}, h_{0,0} f\rangle] = (\tfrac{1}{\gamma} + 1) E[\langle \mathcal{Z}^\gamma_x, h_{0,0} f\rangle] = (\tfrac{1}{\gamma} + 1)\langle \Gamma^{\gamma_{k-1}}_x, h_{0,0} f\rangle, \end{aligned} \tag{7.66}$$

where Γ^γ_x is the invariant law of $\mathbf{y}^\gamma_x$ from Corollary 5.4. In particular, setting $f = 1$ gives $h_{0,0}(x) E^{-k,\delta_x}[|X^{h_{0,0}}_{-k+1}|] = h_{0,0}(x)$ by (5.25).

To prove (3.30), by Lemma 7.12 it suffices to show that

$$\mathcal{U}^{(n)}(\lambda h_{0,0}) \underset{n\to\infty}{\longrightarrow} 0 \tag{7.67}$$

uniformly on $[0,1]$ for all $0 < \lambda \leq 1$. We first treat the case $\gamma^* > 0$. Then, by Theorem 3.2 (a), for each fixed $l \geq 1$ and $f \in \mathcal{C}_+[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l}} f \underset{n\to\infty}{\longrightarrow} (\mathcal{U}_{\gamma^*})^l f \tag{7.68}$$

uniformly on $[0,1]$. Therefore, by a diagonal argument, we can find $l(n) \to \infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)} h_{0,0} - \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}} h_{0,0}\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{7.69}$$

Using the fact that the function $h_{0,0}$ is $\mathcal{U}_\gamma$-superharmonic for each $\gamma > 0$ and the monotonicity of the operators $\mathcal{U}_\gamma$, we derive from Proposition 7.15 (a) that

$$\mathcal{U}^{(n)}(\lambda h_{0,0}) \leq \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}} h_{0,0} \underset{n\to\infty}{\longrightarrow} 0 \tag{7.70}$$

uniformly on $[0,1]$ for all $0 < \lambda \leq 1$. This proves (7.67) in the case $\gamma^* > 0$.

The proof in the case $\gamma^* = 0$ is similar. In this case, by Theorem 3.2 (b), for each fixed $t > 0$ and $f \in \mathcal{C}_+[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{k_n(t)}} f(x_n) \underset{n\to\infty}{\longrightarrow} \mathcal{U}^0_t f(x) \qquad \forall x_n \to x \in [0,1], \tag{7.71}$$

which shows that $\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{k_n(t)}} f$ converges to $\mathcal{U}^0_t f$ uniformly on $[0,1]$. By a diagonal argument, we can find $t(n) \to \infty$ such that

$$\big\|\mathcal{U}^0_t(h_{0,0}) - \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{k_n(t(n))}}(h_{0,0})\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0, \tag{7.72}$$

and the proof proceeds in the same way as before. ∎

Proof of Lemma 3.8 By Lemma 7.12 and the monotonicity of the operators $\mathcal{U}_\gamma$ it suffices to show that

$$\begin{aligned} &\text{(i)} \quad \limsup_{n\to\infty} \mathcal{U}^{(n)}(h_{0,1}) \leq p^*_{0,1,\gamma^*}, \\ &\text{(ii)} \quad \liminf_{n\to\infty} \mathcal{U}^{(n)}(\tfrac{1}{2}h_{0,1}) \geq p^*_{0,1,\gamma^*}, \end{aligned} \tag{7.73}$$

uniformly on $[0,1]$. We first consider the case $\gamma^* > 0$. By (7.68) and a diagonal argument, we can find $l(n) \to \infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)} h_{0,1} - \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}} h_{0,1}\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{7.74}$$

Therefore, by Proposition 7.16 (a), the fact that $h_{0,1}$ is $\mathcal{U}_{\gamma_k}$-superharmonic for each $k \geq 0$, and the monotonicity of the operators $\mathcal{U}_\gamma$, we find that

$$\mathcal{U}^{(n)} h_{0,1} \leq \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}} h_{0,1} \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}, \tag{7.75}$$

uniformly on $[0,1]$. This proves (7.73) (i). To prove also (7.73) (ii) we use the $\mathcal{U}_\gamma$-*sub*harmonic (for each $\gamma > 0$) function h_1 from Lemma 7.8. By Lemma 7.2 also $\frac{1}{2}h_1$ is $\mathcal{U}_\gamma$-subharmonic. By bounding $\frac{1}{2}h_1$ from above and below with multiples of $h_{0,1}$ it is easy to derive from Proposition 7.16 (a) that

$$(\mathcal{U}_{\gamma^*})^n(\tfrac{1}{2}h_1) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*} \tag{7.76}$$

uniformly on $[0,1]$. Arguing as before, we can find $l(n) \to \infty$ such that

$$\big\|(\mathcal{U}_{\gamma^*})^{l(n)}(\tfrac{1}{2}h_1) - \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}}(\tfrac{1}{2}h_1)\big\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \tag{7.77}$$

Therefore, by (7.76) and the facts that $\frac{1}{2}h_1$ is $\mathcal{U}_{\gamma_k}$-subharmonic for each $k \geq 0$ and $\frac{1}{2}h_1 \leq \frac{1}{2}h_{0,1}$,

$$\mathcal{U}^{(n)}(\tfrac{1}{2}h_{0,1}) \geq \mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_{n-l(n)}}(\tfrac{1}{2}h_1) \underset{n\to\infty}{\longrightarrow} p^*_{0,1,\gamma^*}, \tag{7.78}$$

uniformly on $[0,1]$, which proves (7.73) (ii). The proof of (7.73) in case $\gamma^* = 0$ is completely analogous. ∎

8 Extinction on the interior

8.1 Basic facts

In this section we prove Proposition 7.15 (a). To simplify notation, throughout this section h denotes the function $h_{0,0}$. We fix $0 < \gamma^* < \infty$, we let $Y^h := Y^{\gamma^*,h}$ denote the branching particle system on $(0,1)$ obtained from $\mathcal{Y}^{\gamma^*} = (\mathcal{Y}_0^{\gamma^*}, \mathcal{Y}_1^{\gamma^*}, \ldots)$ by Poissonization with h in the sense of Proposition 3.4, and we denote its log-Laplace operator by $U^h_{\gamma^*}$. We will prove that

$$\rho(x) := P^{\delta_x}\big[Y^h_n \neq 0 \ \forall n \geq 0\big] = 0 \qquad (x \in (0,1)). \tag{8.1}$$

Since for each n fixed, $x \mapsto \rho_n(x) := P^{\delta_x}[Y^h_n \neq 0]$ is a continuous function that decreases to $\rho(x)$, (8.1) implies that $\rho_n(x) \to 0$ locally uniformly on $(0,1)$, which, by an obvious analogon of Lemma 7.12, yields Proposition 7.15 (a).

As a first step, we prove:

Lemma 8.1 (Continuous survival probability) *One has either $\rho(x) = 0$ for all $x \in (0,1)$ or there exists a continuous function $\tilde\rho : (0,1) \to [0,1]$ such that $\rho(x) \geq \tilde\rho(x) > 0$ for all $x \in (0,1)$.*

Proof Put $p(x) := h(x)\rho(x)$. We will show that either $p = 0$ on $(0,1)$ or there exists a continuous function $\tilde p : (0,1) \to (0,1]$ such that $p \geq \tilde p$ on $(0,1)$. Indeed,

$$\begin{aligned} p(x) &= h(x)P^{\delta_x}\big[Y^h_n \neq 0 \ \forall n \geq 0\big] = \lim_{n\to\infty} h(x)P^{\delta_x}\big[Y^h_n \neq 0\big] \\ &= h(x)\lim_{n\to\infty}(U^h_{\gamma^*})^n 1(x) = \lim_{n\to\infty}(\mathcal{U}_{\gamma^*})^n h(x) \qquad (x \in (0,1)), \end{aligned} \tag{8.2}$$

where we have used (3.19) and (3.25) in the last two steps. Using the continuity of $\mathcal{U}_{\gamma^*}$ with respect to decreasing sequences, it follows that

$$\mathcal{U}_{\gamma^*}p = p. \tag{8.3}$$

We claim that for any $f \in B_{[0,1]}[0,1]$, one has the bounds

$$\langle \Gamma^\gamma_x, f\rangle \leq \mathcal{U}_\gamma f(x) \leq (1+\gamma)\langle \Gamma^\gamma_x, f\rangle \qquad (\gamma > 0,\ x \in [0,1]). \tag{8.4}$$

Indeed, by Lemma 7.5, $\mathcal{U}_\gamma f(x) \geq 1 - E[(1 - f(\mathbf{y}^\gamma_x(0)))] = \langle \Gamma^\gamma_x, f\rangle$, while the upper bound in (8.4) follows from (7.11).

By Remark 5.5, $(0,1) \ni x \mapsto \langle \Gamma^\gamma_x, f\rangle$ is continuous for all $f \in B_{[0,1]}[0,1]$. Moreover, $\langle \Gamma^\gamma_x, f\rangle = 0$ for some $x \in (0,1)$ if and only if $f = 0$ almost everywhere with respect to Lebesgue measure.

Applying these facts to $f = p$ and $\gamma = \gamma^*$, using (8.3), we see that there are two possibilities. Either $p = 0$ a.s. with respect to Lebesgue measure, and in this case $p = 0$ by the upper bound in (8.4), or p is not almost everywhere zero with respect to Lebesgue measure, and in this case the function $x \mapsto \tilde p(x) := \langle \Gamma^\gamma_x, f\rangle$ is continuous, positive on $(0,1)$, and estimates p from below by the lower bound in (8.4). ∎

8.2 A representation for the Campbell law

(Local) extinction properties of critical branching processes are usually studied using Palm laws. Our proof of formula (8.1) is no exception, except that we will use the closely related Campbell laws. Loosely speaking, Palm laws describe a population that is size-biased at a given position, plus 'typical' particle sampled from that position, while Campbell laws describe a population that is size-biased as a whole, plus a 'typical' particle sampled from a random position.

Let $\mathcal{P}$ be a probability law on $\mathcal{N}(0,1)$ with $\int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)|\nu| = 1$. Then the *size-biased law* $\mathcal{P}_{\rm size}$ associated with $\mathcal{P}$ is the probability law on $\mathcal{N}(0,1)$ defined by

$$\mathcal{P}_{\rm size}(\,\cdot\,) := \int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)\,|\nu| 1_{\{\nu\,\in\,\cdot\,\}}. \tag{8.5}$$

The *Campbell law* associated with $\mathcal{P}$ is the probability law on $(0,1)\times\mathcal{N}(0,1)$ defined by

$$\mathcal{P}_{\rm Camp}(A\times B) := \int_{\mathcal{N}(0,1)} \mathcal{P}(\mathrm{d}\nu)\,\nu(A) 1_{\{\nu\,\in\,B\}} \tag{8.6}$$

for all Borel-measurable $A\subset(0,1)$ and $B\subset\mathcal{N}(0,1)$. If (v,V) is a $(0,1)\times\mathcal{N}(0,1)$-valued random variable with law $\mathcal{P}_{\rm Camp}$, then $\mathcal{L}(V)=\mathcal{P}_{\rm size}$, and v is the position of a 'typical' particle chosen from V.

Let

$$\mathcal{P}^{x,n}(\,\cdot\,) := P^{\delta_x}\big[Y^h_n \in \cdot\,\big] \tag{8.7}$$

denote the law of Y^h at time n, started at time 0 with one particle at position $x\in(0,1)$. Note that by criticality, $\int_{\mathcal{N}(0,1)} \mathcal{P}^{x,n}(\mathrm{d}\nu)|\nu| = 1$. Using again criticality, it is easy to see that in order to prove the extinction formula (8.1), it suffices to show that

$$\lim_{n\to\infty} \mathcal{P}^{x,n}_{\rm size}\big(\{1,\ldots,N\}\big) = 0 \qquad (x\in(0,1),\ N\geq 1). \tag{8.8}$$

In order to prove (8.8), we will write down an expression for $\mathcal{P}^{x,n}_{\rm Camp}$. Let Q^h denote the offspring mechanism of Y^h, and, for fixed $x\in(0,1)$, let $Q^h_{\rm Camp}(x,\,\cdot\,)$ denote the Campbell law associated with $Q^h(x,\,\cdot\,)$. The next proposition is a time-inhomogeneous version of Kallenberg's famous backward tree technique; see [Lie81, Satz 8.2].

Proposition 8.2 (Representation of Campbell law) *Let $(\mathbf{v}_k,V_k)_{k\geq 0}$ be the Markov process in $(0,1)\times\mathcal{N}(0,1)$ with transition laws*

$$P\big[(\mathbf{v}_{k+1},V_{k+1})\in\cdot\,\big|\,(\mathbf{v}_k,V_k)=(x,\nu)\big] = Q^h_{\rm Camp}(x,\,\cdot\,) \qquad ((x,\nu)\in(0,1)\times\mathcal{N}(0,1)), \tag{8.9}$$

started in $(\mathbf{v}_0,V_0)=(\delta_x,0)$. Let $(Y^{h,(k)})^{k\geq 1}$ be branching particle systems with offspring mechanism Q^h, conditionally independent given $(\mathbf{v}_k,V_k)_{k\geq 0}$, started in $Y^{h,(k)}_0 = V_k-\delta_{\mathbf{v}_k}$. Then

$$\mathcal{P}^{x,n}_{\rm Camp} = \mathcal{L}\Big(\mathbf{v}_n, \delta_{\mathbf{v}_n} + \sum_{k=1}^n Y^{h,(k)}_{n-k}\Big). \tag{8.10}$$

Formula (8.10) says that the Campbell law at time n arises in such a way, that an 'immortal' particle at positions $\mathbf{v}_0,\ldots,\mathbf{v}_n$ sheds off offspring $V_1-\delta_{\mathbf{v}_1},\ldots,V_n-\delta_{\mathbf{v}_n}$, distributed according to the size-biased law with one 'typical' particle taken out, and this offspring then evolve under the usual forward dynamics till time n. Note that the position of the immortal particle $(\mathbf{v}_k)_{k\geq 0}$ is an autonomous Markov chain.

We need a bit of explicit control on $Q^h_{\rm Camp}$.

Lemma 8.3 (Campbell law) *One has*

$$Q^h_{\rm Camp}(x,A\times B)=\frac{\frac{1}{\gamma^*}+1}{h(x)}\int P[{\rm Pois}(h\mathcal{Z}^{\gamma^*}_x)\in{\rm d}\chi]\chi(A)1_{\{\chi\in A\}},\tag{8.11}$$

where the random measures $\mathcal{Z}^{\gamma^*}_x$ are defined in (3.7).

Proof By the definition of the Campbell law (8.6), and (3.24),

$$\begin{aligned}Q^h_{\rm Camp}(x,A\times B)&=\int Q^h(x,{\rm d}\chi)\chi(A)1_{\{\chi\in B\}}\\&=\frac{\frac{1}{\gamma^*}+1}{h(x)}\int P[{\rm Pois}(h\mathcal{Z}^{\gamma^*}_x)\in{\rm d}\chi]\chi(A)1_{\{\chi\in B\}}+\Big(1-\frac{\frac{1}{\gamma^*}+1}{h(x)}\Big)\cdot 0.\end{aligned}\tag{8.12}$$

∎

Recall that by (3.7),

$$\mathcal{Z}^{\gamma^*}_x:=\int_0^{\tau_{\gamma_*}}\delta_{\mathbf{y}^{\gamma^*}_x(-t/2)}{\rm d}t,\tag{8.13}$$

where $(\mathbf{y}^{\gamma^*}_x(t))_{t\in\mathbb{R}}$ is a stationary solution to the SDE (3.6) with $\gamma=\gamma^*$. By Lemma 8.3, the transition law of the Markov chain $(\mathbf{v}_k)_{k\geq 0}$ from Proposition 8.2 is given by

$$P[\mathbf{v}_{k+1}\in{\rm d}y|\mathbf{v}_k=x]=\frac{\frac{1}{\gamma^*}+1}{h(x)}E[{\rm Pois}(h\mathcal{Z}^{\gamma^*}_x)({\rm d}y)]=\frac{1+\gamma^*}{h(x)}h(y)\Gamma^{\gamma^*}_x({\rm d}y),\tag{8.14}$$

where $\Gamma^{\gamma^*}_x$ is the invariant law of $\mathbf{y}^{\gamma^*}_x$ from Corollary 5.4. In the next section we will prove the following lemma.

Lemma 8.4 (Immortal particle stays in interior) *The Markov chain $(\mathbf{v}_k)_{k\geq 0}$ started in any $\mathbf{v}_0=x\in(0,1)$ satisfies*

$$(\mathbf{v}_k)_{k\geq 0}\ \textit{has a cluster point in}\ (0,1)\quad\textit{a.s.}\tag{8.15}$$

We now show that Lemma 8.4, together with our previous results, implies Proposition 7.15 (a).

Proof of Proposition 7.15 (a) We need to prove (8.1). By our previous analysis, it suffices to prove (8.8) under the assumption that $\rho\neq 0$. By Proposition 8.2,

$$\mathcal{P}^{x,n}_{\rm size}=\mathcal{L}\Big(\delta_{\mathbf{v}_n}+\sum_{k=1}^n Y^{h,(k)}_{n-k}\Big).\tag{8.16}$$

Conditioned on $(\mathbf{v}_k,V_k)_{k\geq 0}$, the $(Y^{h,(k)}_{n-k})_{k=1,\ldots,n}$ are independent random variables with

$$P\big[Y^{h,(k)}_{n-k}\neq 0\big]\geq P\big[Y^{h,(k)}_m\neq 0\ \forall m\geq 0\big]=P[{\rm Thin}_\rho(V_k-\delta_{\mathbf{v}_k})\neq 0].\tag{8.17}$$

Therefore, (8.8) will follow by Borel-Cantelli provided that we can show that

$$\sum_{k=1}^\infty P[{\rm Thin}_\rho(V_k-\delta_{\mathbf{v}_k})\neq 0|\mathbf{v}_{k-1}]=\infty\quad\text{a.s.}\tag{8.18}$$

Define $f(x) := P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1} = x]$ $(x \in (0,1))$. We need to show that $\sum_{k=1}^\infty f(x) = \infty$ a.s. Using Lemma 8.1 and Lemma 8.3 we can estimate

$$f(x) \geq P[\mathrm{Thin}_{\tilde\rho}(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1} = x] = \int_{\mathcal{N}(0,1)} Q^h_{\mathrm{Camp}}(x, \mathrm{d}y, \mathrm{d}\nu)\{1 - (1-\tilde\rho)^{\nu - \delta_y}\} > 0 \tag{8.19}$$

for all $x \in (0,1)$. Since $\mathcal{Q}_{\gamma^*}$, defined in (3.8), is a continuous cluster mechanism, also $Q^h_{\mathrm{Camp}}(x, \cdot)$ is continuous as a function of x, hence the bound in (8.19) is locally uniform on $(0,1)$, hence Lemma 8.4 implies that there is an $\varepsilon > 0$ such that

$$P[\mathrm{Thin}_\rho(V_k - \delta_{\mathbf{v}_k}) \neq 0 | \mathbf{v}_{k-1}] \geq \varepsilon \tag{8.20}$$

at infinitely many times $k-1$, which in turn implies (8.18). ∎

8.3 The immortal particle

Proof of Lemma 8.4 Let $K(x, \mathrm{d}y)$ denote the transition kernel (on $(0,1)$) of the Markov chain $(\mathbf{v}_k)_{k\geq 0}$, i.e., by (8.14),

$$K(x, \mathrm{d}y) = (1+\gamma^*)\frac{y(1-y)}{x(1-x)}\Gamma^{\gamma^*}_x(\mathrm{d}y). \tag{8.21}$$

It follows from (5.24) that

$$\int K(x, \mathrm{d}y) y(1-y) = \frac{x(1-x) + \gamma^*(1+\gamma^*)}{(1+2\gamma^*)(1+3\gamma^*)}. \tag{8.22}$$

Set

$$g(x) := \int K(x, \mathrm{d}y) y(1-y) - x(1-x) \qquad (x \in (0,1)). \tag{8.23}$$

Then

$$M_n := \mathbf{v}_n(1-\mathbf{v}_n) - \sum_{k=0}^{n-1} g(\mathbf{v}_k) \qquad (n \geq 0) \tag{8.24}$$

defines a martingale $(M_n)_{n\geq 0}$. Since $g > 0$ in an open neighborhood of $\{0,1\}$,

$$P[(\mathbf{v}_k)_{k\geq 0} \text{ has no cluster point in } (0,1)] \leq P[\lim_{n\to\infty} M_n = -\infty] = 0, \tag{8.25}$$

where in the last equality we have used that $(M_n)_{n\geq 0}$ is a martingale. ∎

9 Proof of the main result

Proof of Theorem 1.4 Part (a) has been proved in Section 5.3. It follows from (2.12), (2.13), (3.10), and (3.11) that part (b) is equivalent to the following statement. Assuming that

$$\text{(i)} \quad \sum_{n=1}^\infty \gamma_n = \infty \qquad \text{and} \qquad \text{(ii)} \quad \gamma_n \underset{n\to\infty}{\longrightarrow} \gamma^* \tag{9.1}$$

for some $\gamma^* \in [0,\infty)$, one has, uniformly on $[0,1]$,

$$\mathcal{U}_{\gamma_{n-1}} \circ \cdots \circ \mathcal{U}_{\gamma_0}(p) \underset{n\to\infty}{\longrightarrow} p^*_{l,r,\gamma^*}, \tag{9.2}$$

where p^*_{l,r,γ^*} is the unique solution in $\mathcal{H}_{l,r}$ of

$$\begin{array}{llll} \text{(i)} & \mathcal{U}_{\gamma^*} p^* = p^* & & \text{if } 0 < \gamma^* < \infty, \\ \text{(ii)} & \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}p^*(x) - p^*(x)(1-p^*(x)) = 0 & (x \in [0,1]) & \text{if } \gamma^* = 0. \end{array} \tag{9.3}$$

It follows from Proposition 3.5 that the left-hand side of (9.2) converges uniformly to a limit p^*_{l,r,γ^*} which is given by (3.34). We must show 1° that $p^*_{l,r,\gamma^*} \in \mathcal{H}_{l,r}$ and 2° that p^*_{l,r,γ^*} is the unique solution in this class to (9.3). We first treat the case $\gamma^* > 0$.

1° Since $p^*_{0,0,\gamma^*} \equiv 0$ and $p^*_{1,1,\gamma^*} \equiv 1$, it is obvious that $p^*_{0,0,\gamma^*} \in \mathcal{H}_{0,0}$ and $p^*_{1,1,\gamma^*} \in \mathcal{H}_{1,1}$. Therefore, by symmetry, it suffices to show that $p^*_{0,1,\gamma^*} \in \mathcal{H}_{0,1}$. By Lemmas 7.8 and 7.11, $x \leq p \leq 1-(1-x)^7$ implies $x \leq \mathcal{U}_{\gamma_k} p \leq 1-(1-x)^7$ for each k. Iterating this relation, using (9.2), we find that

$$x \leq p^*_{0,1,\gamma^*}(x) \leq 1-(1-x)^7. \tag{9.4}$$

By Proposition 5.11, the left-hand side of (9.2) is nondecreasing and concave in x if p is, so taking the limit we find that $p^*_{0,1,\gamma^*}$ is nondecreasing and concave. Combining this with (9.4) we conclude that $p^*_{0,1,\gamma^*}$ is Lipschitz continuous. Moreover $p^*_{0,1,\gamma^*}(0) = 0$ and $p^*_{0,1,\gamma^*}(1) = 1$ so $p^*_{0,1,\gamma^*} \in \mathcal{H}_{0,1}$.

2° Taking the limit $n \to \infty$ in $(\mathcal{U}_{\gamma^*})^n p = \mathcal{U}_{\gamma^*}(\mathcal{U}_{\gamma^*})^{n-1} p$, using the continuity of $\mathcal{U}_{\gamma^*}$ (Corollary 5.10) and (9.2), we find that $\mathcal{U}_{\gamma^*} p^*_{l,r,\gamma^*} = p^*_{l,r,\gamma^*}$. It follows from (9.2) that p^*_{l,r,γ^*} is the only solution in $\mathcal{H}_{l,r}$ to this equation.

For $\gamma^* = 0$, it has been shown in [FS03, Proposition 3] that $p^*_{l,r,0}$ is the unique solution in $\mathcal{H}_{l,r}$ to (9.3) (ii). In particular, it has been shown there that $p^*_{0,1,0}$ is twice continuously differentiable on $[0,1]$ (including the boundary). This proves parts (b) and (c) of the theorem. ∎

A Appendix: Infinite systems of linearly interacting diffusions

A.1 Hierarchically interacting diffusions

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

$$\|\xi\| := \min\{n \geq 0 : \xi_k = 0 \ \forall k > n\} \qquad (\xi \in \Omega_N), \tag{A.1}$$

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

Let $D \subset \mathbb{R}^d$ be open and convex, and let $\mathcal{W}$ be a renormalization class on $\overline{D}$. Let σ be a continuous root of a diffusion matrix $w \in \mathcal{W}$ as in Remark 1.2. Consider a collection $\mathbf{x} = (\mathbf{x}_\xi)_{\xi\in\Omega_N}$ of $\overline{D}$-valued processes, solving a system of SDE's of the form

$$\mathrm{d}\mathbf{x}_\xi(t) = \sum_{k=0}^{\infty} \frac{c_k}{N^k}\Big(\mathbf{x}^{k+1}_\xi(t) - \mathbf{x}_\xi(t)\Big)\mathrm{d}t + \sqrt{2}\sigma(\mathbf{x}_\xi(t))\mathrm{d}B_\xi(t) \qquad (t \geq 0,\ \xi \in \Omega_N), \tag{A.2}$$

where $(B_\xi)_{\xi\in\Omega_N}$ is a collection of independent standard Brownian motions, with initial condition

$$\mathbf{x}_\xi(0) = \theta \in D \qquad (\xi \in \Omega_N). \tag{A.3}$$

Here the $(c_k)_{k\geq 0}$ are positive constants satisfying $\sum_k c_k/N^k < \infty$, and $\mathbf{x}_\xi^k(t)$ denotes the *k-block average* around ξ:

$$\mathbf{x}_\xi^k(t) := \frac{1}{N^k} \sum_{\eta:\|\xi-\eta\|\leq k} \mathbf{x}_\eta(t) \qquad (k \geq 0). \tag{A.4}$$

(Note that $|\{\eta : \|\xi - \eta\| \leq k\}| = N^k$.) Under suitable additional assumptions on σ, one can show that (A.2) has a unique (weak or strong) solution (see [DG93a, DG96, Swa00]). We call $\mathbf{x}$ a system of *hierarchically interacting $\overline{D}$-valued diffusions* with migration constants $(c_k)_{k\geq 0}$ and local diffusion rate $w_{ij} = \sum_k \sigma_{ik}\sigma_{jk}$. Such systems are used to model gene frequencies or population sizes in population biology [SF83].

The long-time behavior of the system in (A.2) depends crucially on the recurrence versus transience of the continuous-time random walk on Ω_N which jumps from a point ξ to a point $\eta \neq \xi$ with rate

$$a(\eta - \xi) := \sum_{k=\|\xi-\eta\|}^{\infty} \frac{c_{k-1}}{N^{2k-1}}. \tag{A.5}$$

This random walk is recurrent if and only if

$$\sum_{k=0}^{\infty} \frac{1}{d_k} = \infty, \qquad \text{where} \quad d_k := \sum_{n=0}^{\infty} \frac{c_{k+n}}{N^n} \tag{A.6}$$

(see [DG93a, Kle96]; a similar problem is treated in [DE68]). Assuming that the law of $\mathbf{x}(t)$ converges weakly as $t \to \infty$ to the law of some $\overline{D}^{\Omega_N}$-valued random variable $\mathbf{x}(\infty)$, one expects that in the recurrent case $\mathbf{x}(\infty)$ must have the following properties:

$$\begin{array}{lll} \text{(i)} \quad \mathbf{x}_\xi(\infty) = \mathbf{x}_\eta(\infty) & \text{a.s.} & \forall \xi, \eta \in \Omega_N, \\ \text{(ii)} \quad \mathbf{x}_\xi(\infty) \in \partial_w D & \text{a.s.} & \forall \xi \in \Omega_N. \end{array} \tag{A.7}$$

Here $\partial_w D$ is the effective boundary of D, defined in (2.3). If $\mathbf{x}(t)$ converges in law to a limit $\mathbf{x}(\infty)$ satisfying (A.7), then we say that $\mathbf{x}$ *clusters*. In the transient case, it is believed that solutions of (A.2) do not cluster. (For compact $\overline{D}$ these facts were proved in [Swa00].)

An important tool in the study of solutions to (A.2) is the so-called *interaction chain*. This is the chain $(\mathbf{x}_0^0(t), \mathbf{x}_0^1(t), \ldots)$ of block-averages around the origin. Heuristic arguments suggest that in the *local mean field limit* $N \to \infty$, the interaction chain converges to a certain well-defined Markov chain.

Conjecture A.1 *Fix $w \in \mathcal{W}$, $\theta \in D$, and positive numbers $(c_k)_{k\geq 0}$ such that for N large enough, $\sum_k c_k/N^k < \infty$. For all N large enough, let $\mathbf{x}^N$ be a solution to (A.2)–(A.3), and assume that t_N are constants such that, for some $n \geq 1$, $\lim_{N\to\infty} N^{-n} t_N = T \in [0, \infty)$. Then*

$$\Big(\mathbf{x}_0^{N,n}(t_N), \ldots, \mathbf{x}_0^{N,0}(t_N)\Big) \underset{N\to\infty}{\Longrightarrow} (I_{-n}^w, \ldots, I_0^w), \tag{A.8}$$

where $(I_{-n}^w, \ldots, I_0^w)$ is a Markov chain with transition laws

$$P[I_{-k}^w \in \mathrm{d}y | I_{-k-1}^w = x] = \nu_x^{c_k, F^{(k)}w}(\mathrm{d}y) \qquad (x \in \overline{D},\ 0 \leq k \leq n-1) \tag{A.9}$$

and initial state

$$I^w_{-n} = \mathbf{y}_T, \quad \textit{where} \quad \mathrm{d}\mathbf{y}_t = c_n(\theta - \mathbf{y}_t)\mathrm{d}t + \sqrt{2}\sigma^{(n)}(\mathbf{y}_t)\mathrm{d}B_t, \quad \mathbf{y}_0 = \theta, \tag{A.10}$$

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

Rigorous versions of conjecture A.1 have been proved for renormalization classes on $\overline{D} = [0, 1]$ and $\overline{D} = [0, \infty)$ in [DG93a, DG93b].

Note that the iterated kernels $K^{w,(n)}$ defined in (2.4) are the transition probabilities from time $-n$ to time 0 of the interaction chain in the mean-field limit:

$$K^{w,(n)}_x(\mathrm{d}y) = P[I^w_0 \in \mathrm{d}y | I^w_{-n} = x] \qquad (x \in \overline{D},\ n \geq 0). \tag{A.11}$$

Lemma 2.6 expresses the fact that the system $\mathbf{x}^N$ clusters in the local mean-field limit $N \to \infty$. The condition $s_n \to \infty$ in Lemma 2.6 means that $\sum_{k\geq 0} \frac{1}{c_k} = \infty$, which, in a sense, is the $N \to \infty$ limit of condition (A.6).

A.2 The clustering distribution of linearly interacting diffusions

Let $D \subset \mathbb{R}^d$ be open, *bounded*, and convex, and let $\mathcal{W}$ be a renormalization class on $\overline{D}$. Fix migration constants $(c_k)_{k\geq 0}$ and assume that $s_n \to \infty$ and $s_{n+1}/s_n \to 1 + \gamma^*$ for some $\gamma^* \in [0, \infty]$. Recall the definition of the iterated probability kernels $K^{w,(n)}$ in (2.4). Recall Conjecture 2.7. Assuming that the rescaled renormalized diffusion matrices $s_n F^{(n)} w$ converge to a limit w^*, we can make a guess about the limit of the iterated probability kernels $K^{w,(n)}$.

Conjecture A.2 (Limits of iterated probability kernels) *Assume that $s_n F^{(n)} w \to w^*$ as $n \to \infty$. Then, for any $w \in \mathcal{W}$,*

$$K^{w,(n)} \underset{n\to\infty}{\longrightarrow} K^*, \tag{A.12}$$

where K^* has the following description:

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

$$K^*_x = \lim_{n\to\infty} P^x[I^{\gamma^*}_n \in \cdot\,], \tag{A.13}$$

where $(I^{\gamma^*}_n)_{n\geq 0}$ is the Markov chain with transition law $P[I^{\gamma^*}_{n+1} \in \cdot\, | I^{\gamma^*}_n = x] = \nu^{1/\gamma^*, w^*}$.

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

$$K^*_x = \lim_{t\to\infty} P^x[I^0_t \in \cdot\,], \tag{A.14}$$

where $(I^0_s)_{s\geq 0}$ is the diffusion process with generator $\sum_{i,j=1}^d w^*_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j}$.

(iii) *If $\gamma^* = \infty$, then*

$$K^*_x = \lim_{\gamma\to\infty} \nu^{1/\gamma, w^*}_x. \tag{A.15}$$

For each $N \geq 2$, let $\mathbf{x}^N = (\mathbf{x}^N_\xi)_{\xi\in\Omega_N}$ be a system of hierarchically interacting diffusions as in (A.2) and (A.3). If $\gamma^* = 0$, then because of Conjectures A.1 and A.2, we expect[3] that

$$\lim_{n\to\infty} \lim_{N\to\infty} \mathcal{L}(\mathbf{x}^N_0(N^n T)) = K^*_\theta \qquad (T > 0), \tag{A.16}$$

[3]For $\gamma^* > 0$, the situation is more complex. In this case at the right-hand side of (A.16) we expect the law $\int_{\overline{D}} P^\theta[\mathbf{y}_T \in \mathrm{d}x] K^*_x$, where $\mathbf{y}$ solves the SDE $\mathrm{d}\mathbf{y}_t = \frac{1}{\gamma^*}(\theta - \mathbf{y}_t)\mathrm{d}t + \sqrt{2}\sigma^*(\mathbf{y}_t)\mathrm{d}B_t$ and σ^* is a root of the diffusion matrix w^*. Note that in this case the right-hand side of (A.16) depends on T.

where K^* is the kernel in (A.14).

In particular, consider the case that the migration constants $(c_k)_{k\geq 0}$ are of the form $c_k = r^k$ for some $r > 0$. In this case, $s_{n+1}/s_n \to \frac{1}{r} \vee 1$, and $s_n \to \infty$ if and only if $r \leq 1$. One can check (see (A.6)) that for fixed $N \geq 2$, the random walk with the kernel a in (A.5) is recurrent if and only if $r \leq 1$. The critical case $r = 1$ corresponds to a *critically recurrent* random walk. For a precise definition of critical recurrence, see [Kle96, formula (1.15)]. For $r = 1$, we expect that the double limit in (A.16) can be replaced by a single limit. More precisely, for each fixed $N \geq 2$, we expect that

$$\lim_{t\to\infty} \mathcal{L}(\mathbf{x}_0^N(t)) = K_\theta^*. \tag{A.17}$$

In this case, we call K_θ^* the *clustering distribution* of $\mathbf{x}^N$. The clustering distribution of linearly interacting isotropic diffusions was studied in [Swa00]. We expect (A.17) to hold, even more generally, for all systems of linearly interacting diffusions with a critically recurrent migration mechanism. In particular, we expect (A.17) to hold for symmetric nearest-neighbor interaction on $\mathbb{Z}^d$ in the critical dimension $d = 2$. If one is ready to make this enormous leap of faith, then combining Conjectures 2.7 and A.2, one arrives at the following conjecture.

Conjecture A.3 (Critical clustering) *Let $D \subset \mathbb{R}^d$ be open, bounded, and convex, and let $\mathcal{W}$ be a renormalization class on $\overline{D}$. Assume that the asymptotic fixed point equation (2.16) (ii) has a unique solution w^* in $\mathcal{W}$. Let σ be a continuous root of a diffusion matrix $w \in \mathcal{W}$. Let $\mathbf{x} = (\mathbf{x}_\xi)_{\xi\in\mathbb{Z}^2}$ be a $\overline{D}^{\mathbb{Z}^2}$-valued process, solving the system of SDE's*

$$\mathrm{d}\mathbf{x}_\xi(t) = \sum_{\eta:\,|\eta-\xi|=1} \big(\mathbf{x}_\eta(t) - \mathbf{x}_\xi(t)\big)\,\mathrm{d}t + \sigma(\mathbf{x}_\xi(t))\mathrm{d}B_\xi(t), \tag{A.18}$$

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

$$\mathbf{x}_\xi(t) \underset{t\to\infty}{\Longrightarrow} I_\infty^\theta \qquad (\xi \in \mathbb{Z}^2), \tag{A.19}$$

where $(I_s^\theta)_{s\geq 0}$ is the diffusion with generator $\sum_{i,j} w_{ij}^*(y)\frac{\partial^2}{\partial y_i \partial y_j}$ and initial condition $I_0^\theta = \theta$.

Acknowledgements

We thank Anton Wakolbinger and Martin Möhle for pointing out reference [Ewe04] and the fact that the distribution in (5.17) is a β-distribution.

References

[BCGdH95] J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. I. The compact case. *Canad. J. Math.* 47(1): 3–27, 1995.

[BCGdH97] J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander. On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. II. The non-compact case. *J. Funct. Anal.* 146: 236–298, 1997.

[CDG04] J.T. Cox, D.A. Dawson, and A. Greven. Mutually catalytic super branching random walks: Large finite systems and renormalization analysis. *Mem. Am. Math. Soc.* 809, 2004.

[DE68] D.A. Darling and P. Erdős. On the recurrence of a certain chain. *Proc. Am. Math. Soc.* 19(1): 336-338, 1968.

[DG93a] D.A. Dawson and A. Greven. Hierarchical models of interacting diffusions: Multiple time scale phenomena, phase transition and pattern of cluster-formation. *Probab. Theory Related Fields* 96(4): 435–473, 1993.

[DG93b] D.A. Dawson and A. Greven. Multiple time scale analysis of interacting diffusions. *Probab. Theory Related Fields* 95(4): 467–508, 1993.

[DG96] D.A. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.*, 1: no. 14, approx. 84 pp. (electronic), 1996.

[DGV95] D.A. Dawson, A. Greven and J. Vaillancourt. Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.* 347(7): 2277–2360, 1995.

[dHS98] F. den Hollander and J.M. Swart. Renormalization of hierarchically interacting isotropic diffusions. *J. Stat. Phys.* 93: 243–291, 1998.

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

[ER91] N. El Karoui and S. Roelly. Propriétés de martingales, explosion et représentation de Lévy- Khintchine d'une classe de processus de branchement à valeurs mesures. *Stoch. Proc. Appl.* 38(2): 239–266, 1991.

[Ewe04] W.J. Ewens. *Mathematical Population Genetics. I: Theoretical Introduction. 2nd ed.* Interdisciplinary Mathematics 27. Springer, New York, 2004.

[Fit88] P.J. Fitzsimmons. Construction and regularity of measure-valued branching processes. *Isr. J. Math.* 64(3): 337–361, 1988.

[FS03] K. Fleischmann and J.M. Swart. Extinction versus exponential growth in a supercritical super-Wright-Fischer diffusion. *Stoch. Proc. Appl.* 106(1): 141–165, 2003.

[FS04] K. Fleischmann and J.M. Swart. Trimmed trees and embedded particle systems. *Ann. Probab.* 32(3a): 2179–2221, 2004.

[GKW01] A. Greven, A. Klenke, and A. Wakolbinger. Interacting Fisher-Wright diffusions in a catalytic medium. *Probab. Theory Related Fields* 120(1): 85–117, 2001.

[Jir64] M. Jiřina. Branching processes with measure-valued states. In *Trans. Third Prague Conf. Information Theory, Statist. Decision Functions, Random Processes (Liblice, 1962)*, pages 333–357, Czech. Acad. Sci., Prague, 1964.

[Kal76] O. Kallenberg. *Random Measures.* Akademie-Verlag, Berlin, 1976.

[Kle96] A. Klenke. Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.* 24(2): 660–697, 1996.

[Lie81] A. Liemant. Kritische Verzweigungsprozesse mit allgemeinem Phasenraum. IV. *Math. Nachr.* 102: 235–254, 1981.

[Loe63] M. Loève. *Probability Theory* 3rd ed. Van Nostrand, Princeton, 1963.

[Loe78] M. Loève. *Probability Theory II* 4th ed. Graduate Texts in Mathematics 46. Springer, New York, 1978.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York, 1983.

[RW87] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales, Volume 2: Ito Calculus.* Wiley, Chichester, 1987.

[Sch98] F. Schiller. *Application of the Multiple Space-Time Scale Analysis on a System of R-valued, Hierarchically Interacting, Stochastic Differential Equations.* Master thesis, Universtity Erlangen-Nürnberg, 1998.

[SF83] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.* 20: 1–10, 1983.

[Shi80] T. Shiga. An interacting system in population genetics. *J. Math. Kyoto Univ.* 20: 213–242, 1980.

[Swa99] J.M. Swart. *Large Space-Time Scale Behavior of Linearly Interacting Diffusions.* PhD thesis, Katholieke Universiteit Nijmegen, 1999. http://helikon.ubn.kun.nl/mono/s/swart_/largspscb.pdf.

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

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