isid/ms/2004/14
May 13, 2004
http://www.isid.ac.in/~statmath/eprints

Spatial Coupling of Neutral Measure-valued Population Models

Siva Athreya
Anita Winter

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

SPATIAL COUPLING OF NEUTRAL MEASURE-VALUED POPULATION MODELS

SIVA ATHREYA AND ANITA WINTER

Abstract. In this article we discuss spatial couplings for measure-valued population models which have a particle representation. We will show that provided the corresponding genealogical trees are compact the qualitative behavior of a coupling of a particle's individual motion translates into a coupling of the continuous mass measure-valued models. As applications of the above method we present a coupling of diffusions on $\mathbb{R}^n_+$ and a perturbation estimate for a class of semilinear partial differential equations.

1. Introduction

Couplings of Markov processes ξ have been widely studied in the literature and applied to various areas of Probability and Analysis. In this paper we study spatial couplings of measure-valued population models, X, with values in finite measures on a locally compact and separable metric space (E, d). These processes are obtained as weak rescaling limits of branching particle systems where the particles *migrate* or *mutate* according to a Markov process ξ in a geographical respectively type space E, and give birth to particles upon dying by a specified *branching mechanism*.

Let $x, y \in E$. Assume that there is a coupling of two copies of ξ starting at x and y, such that $\lim_{t\to\infty} d(\xi^x_t, \xi^y_t) = 0$, almost surely. Consider two copies X^x_t and X^y_t of the measure-valued process starting at δ_x and δ_y, respectively, having synchronous branching events, and underlying migration(mutation) process given by ξ^x and ξ^y, respectively. We investigate the following question: Does the distance X^x_t and X^y_t go to zero as well, and if so does this distance converge at the same rate?

We begin with the difficulties that arise when one tries to answer the above question. First we need to choose a meaningful measurement of coupling of finite measures. In particular, one has to keep in mind that the supports propagate, so it is not enough to try to consider the Hausdorff distance of the supports. Moreover, since branching models typically (may) die out, a meaningful coupling should also apply to the models conditioned to stay alive for ever. We consider the various possibilities.

Date: June 28, 2004.

2000 *Mathematics Subject Classification.* Primary: 60G57 Secondary: 60H30.

Key words and phrases. Measure-valued, Coupling, Look-down, Neumann Problem.

Given a spatial coupling of the motion process, ξ, there are various ways in which one can construct a coupling of X conditioned on non-extinction. One approach could be to use the immortal particle representation ([13]). In this representation there is one immortal particle that throws of mass as it lives for ever. One could also work with the historical version of these models as introduced in [9] for superprocesses and in [14] for interacting Fisher-Wright diffusions. However, we will follow an approach which will cover both aspects, namely the so-called look-down construction given by Donnelly and Kurtz ([10],[11]). This approach is applicable for a wide class of measure-valued population models.

We rely on the last approach to present a generic way to spatially couple X from the coupling of ξ. In our main result (Theorem 2.1) we show that a spatial coupling of ξ can be lifted to a spatial coupling of X under certain assumptions. A key fact that we make use of in the proof is that for some time t sufficiently close to the extinction time τ^{ext}, all particles alive at time t share a common ancestor at time 0, i.e., the corresponding genealogical tree is compact.

We apply Theorem 2.1 to different situations. A class of examples we study are super-reflected Brownian motions on a planar domain D. Using synchronous couplings of reflected Brownian motions on certain planar domains D [[2], [3], [6], [7]] we are able to provide successful couplings of these measure valued processes (see Corollary 3.1). As an application we study the corresponding Neumann problem, i.e.,

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{1}{2}\Delta u - \ell u + \frac{\gamma}{2}u^{1+\beta}, \quad x \text{ in } D \\
\frac{\partial u}{\partial n} &\equiv 0, \quad x \text{ on } \partial D \\
u(0,x) &= g(x), \quad x \text{ in } D,
\end{aligned} \tag{1.1}$$

where $\frac{\partial}{\partial n}$ is the normal derivative, $\ell \geq 0$, and $\beta \in]0,1]$. In [1] such a coupling is used to provide monotonicity results for (1.1). Using Theorem 2.1 we are able to provide a specific rate at which $|u(t,x) - u(t,y)| \to 0$, as $t \to \infty$, for a certain class of planar domains D (see Corollary 3.2). Another set of examples are Super Markov chains on a finite set $E = \{1, \ldots, n\}$. These can be identified with solutions of finite dimensional diffusions in $\mathbb{R}^n_+$. Suppose one has a successful coupling of a Markov chain on E, (see for instance [5]), then our main result provides a successful coupling of finite-dimensional diffusions. This is made precise in Corollary 3.3. Various examples of couplings of multi-dimensional diffusions are given in [15]. However, the diffusions we couple cover a different set of examples.

Furthermore in the literature, spatial couplings of measure-valued processes have been considered. In [12] a spatial coupling of Fleming-Viot processes is used to study its ergodic properties. A coupling as in Theorem 2.1 (ii) is obtained for the Fleming-Viot process in [10] (see the proof of Theorem 4.1).

The rest of the paper is organised as follows. In the next section we state the model and the main result (Theorem 2.1). In Section 3 we discuss the examples of couplings and the application to semi-linear parabolic pdes. Finally in Section 4 we give a proof of Theorem 2.1.

Acknowledgments. We had many useful discussions on various aspects of the problem and a draft of the manuscript with Krzysztof Burdzy, Steven Evans, Thomas Kurtz, Lea Popovic and Iljana Zähle. Our many thanks to all of them.

2. The model and main result

We begin with a particle representation of the measure-valued process, followed by the coupling and then by the main result.

2.1. **The model and its particle representation.** Let $(A,\mathcal{D}(A))$ be the generator of a Markov process $\xi := (\xi_t)_{t\geq 0}$ on E modeling the motion of a *single particle*, and $P := (P_t)_{t\geq 0}$ be a stochastic process taking values in $[0,\infty[$ describing the *total mass* of the model. Let

$$\tau^{\text{ext}} := \inf\{t \geq 0 :\, P_t = 0\}. \tag{2.1}$$

Assumption on P We assume that P_t satisfies the following condition:

(P)

$$\int_0^{\tau^{\text{ext}}} \mathrm{d}s\, \frac{1}{P_s} = \infty.$$

Consider the $E^{\mathbb{N}}$-valued Markov process $(\xi_t^1,\ldots,\xi_t^n,\ldots)$ which evolves until τ^{ext} as follows:

- **(Migration/Mutation)** For each $k\in\mathbb{N}$, the k^{th}-coordinate process, $\xi^k := (\xi_t^k)_{t\geq 0}$, also called the k^{th}-level process, performs, independently of all the other level processes, a Markov-process on E with generator $(A,\mathcal{D}(A))$.
- **(Branching/Resampling)** For any $(i,j)\in\mathbb{N}\times\mathbb{N}$ with $i<j$, at rate $1/P_t$, a particle of "type" (respectively at "position") ξ_t^i is inserted one level above the j^{th} level, while all other particle retain their order. That is, after the jump we end up with the vector $(\xi_t^1,...,\xi_t^{i-1},\xi_t^i,\xi_t^{i+1},...,\xi_t^{j-1},\xi_{t-}^i,\xi_t^j,\xi_t^{j+1},...)$.

It is easy to check that under Assumption (P) the dynamics is well-defined, and that if $(\xi_0^1,\xi_0^2,...)$ is exchangeable then so is $(\xi_t^1,\xi_t^2,...)$ for any $t\geq 0$. Hence the de Finetti measure

$$\bar{\mu}_t := \lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^m \delta_{\xi_t^k}, \quad a.s. \tag{2.2}$$

exists. Let $\mu \in M_F(E)$, where $M_F(E)$ denotes the set of finite measures on $(E, \mathcal{B}(E))$. The model $X(P,\xi,\mu)$ we are interested in is the process $X := (X_t)_{t\geq 0}$ defined by

$$X_t := P_t \cdot \bar{\mu}_t \, 1_{\{t \leq \tau^{\text{ext}}\}}, \quad t \geq 0. \tag{2.3}$$

The above particle representation of measure-valued process and their applications was first given in [11]. We refer to that article for more details of the construction.

2.2. **The coupling.** We will use the following construction. Let ξ be the motion processes. Denote by $D([0,\infty[, E)$ the set of maps from $[0,\infty[$ to E which are continuous from the right and have limits from the left. Assume that there is a function

$$f : E \times D([0,\infty[, E) \to D([0,\infty[, E) \tag{2.4}$$

such that for each $x \in E$, $\xi^x := f(x,\xi)$ equals in distribution ξ started in x. In the following f is referred to as a *spatial coupling.*

Let $x, y \in E$. Let two copies $X^x \equiv X(P,\xi,\delta_x)$ and $X^y \equiv X(P,\xi,\delta_y)$ are spatially coupled with respect to f. That is, the processes start at $P_0 \cdot \delta_x$ and $P_0 \cdot \delta_y$, are driven by the same total mass process $P := (P_t)_{t\geq 0}$, the branching events occur simultaneously, but the spatial motion of the i^{th} particle started in x and y is given by $f(x,\xi^i)$ and $f(y,\xi^i)$, respectively. For such a coupling, we provide the following three measurements of efficiency.

- (Averaged spatial distance)

$$\bar{z}_t^{x,y} := \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^m d(\xi_t^{i,x}, \xi_t^{i,y}), \quad \text{a.s.} \tag{2.5}$$

- (Minimal/maximal spatial distance)

$$\underline{\mathrm{m}}_t^{x,y} := \inf_{i\in\mathbb{N}} d(\xi_t^{i,x}, \xi_t^{i,y}), \quad \text{and} \quad \bar{m}_t^{x,y} := \sup_{i\in\mathbb{N}} d(\xi_t^{i,x}, \xi_t^{i,y}). \tag{2.6}$$

- (Wasserstein distance of the random measures)

$$d_{\mathrm{W}}(X_t^x, X_t^y) := \sup_{h\in C_b} \left| \int h \mathrm{d}X_t^x - \int h \mathrm{d}X_t^y \right|, \tag{2.7}$$

  where $C_b = C_b(E)$ is the set of continuous functions $h : E \to \mathbb{R}$ such that $|h(x)| \leq 1$ and $|h(x) - h(y)| \leq d(x,y)$ for $x, y \in E$.

Now note that since $(d(\xi_t^{1,x}, \xi_t^{1,y}), d(\xi_t^{2,x}, \xi_t^{2,y}), \ldots)$ is exchangeable the limit in (2.5) and consequently the quantity $\bar{z}_t^{x,y}$ is well-defined. It is elementary to observe that the above efficiencies can be arranged in order. Let $t \geq 0$, then

$$P_t \underline{\mathrm{m}}_t^{x,y} \leq d_{\mathrm{W}}(X_t^x, X_t^y) \leq P_t \bar{z}_t^{x,y} \leq P_t \bar{m}_t^{x,y}. \tag{2.8}$$

Assuming that we may couple the motion processes such that $d(\xi_t^x, \xi_t^y) \to 0$, as $t \to \infty$. Then by exchangeability $\bar{z}_t^{x,y} \to 0$ in probability, as $t \to \infty$. For almost sure convergence one might try to argue as follows:

$$\lim_{t\to\infty} \bar{z}_t^{x,y} = \lim_{m\to\infty} \lim_{t\to\infty} \max_{1\le i\le m} d(\xi_t^{x,i}, \xi_t^{y,i}) = 0, \quad \text{a.s.} \tag{2.9}$$

The above calculation involves an interchange in limits which cannot be justified without a further assumption on the spatial coupling.

Assumption on ξ Let $\xi := (\xi_t)_{t\ge 0}$ be a stochastic process with values in E such that there exists a coupling of two copies ξ^x and ξ^y started in $x, y \in E$ such that one of the following conditions is satisfied.

(S1) There exist a function $g : \mathbb{R}^+ \times E \times E \to \mathbb{R}^+$ with

$$g(t,x,y) d(\xi_t^x, \xi_t^y) \to 0, \quad a.s., \tag{2.10}$$

and an almost surely finite stopping time T such that $g(t,x,y)d(\xi_t^x, \xi_t^y)$ is almost surely non-increasing on $[T, \infty[$.

(S2) There exists an almost surely finite stopping time S with

$$d(\xi_t^x, \xi_t^y) = 0 \tag{2.11}$$

for all $t \ge S$.

We are now ready to state our main result.

Theorem 2.1. *Let X^x and X^y be spatially coupled versions of X as described in the previous section.*

(i) *Assume (S1). On the event $\{\tau^{\text{ext}} = \infty\}$,*

$$\lim_{t\to\infty} g(t,x,y) \bar{m}_t^{x,y} = 0, \quad a.s. \tag{2.12}$$

(ii) *Assume (S2). On the event $\{\tau^{\text{ext}} = \infty\}$,*

$$\tau := \inf\{t \ge 0 : \bar{m}_t^{x,y} = 0\} < \infty, \quad a.s. \tag{2.13}$$

We will prove this theorem in Section 4. The main idea in the proof is to prevent the occurrence of "exceptional" particles. For this we will make use of the fact that the underlying genealogical tree is compact, i.e., for any $t \ge 0$ and $\varepsilon \in]0, t[$, the countably many levels at time t have finitely many ancestor levels at time $t - \varepsilon$. The latter can be derived from Condition (P) for the total mass process. Examples for total mass processes satisfying Condition (P) are discussed in Subsection 3.1.

Remark 1.

(1) We have chosen $\bar{m}_t^{x,y}$ to work with as this clearly illustrates the coupling of the supports of X_t^x and X_t^y, i.e. a spatial coupling. This is also the case with $\bar{z}_t^{x,y}$. This is not so with the Wasserstein distance. If one were to work with this distance, if P_t dies out then $d_W(X_t^x, X_t^y)$ approaches zero trivially. If P_t does not die out, then in order to ensure that $d_W(X_t^x, X_t^y)$ approaches zero, one would require $\bar{m}_t^{x,y}$ to approach zero at a certain rate. In the next section we will discuss various examples of P_t and in certain cases it will be immediate to conclude that $d_W(X_t^x, X_t^y)$ approaches zero given that $\bar{m}_t^{x,y}$ approaches zero. However, in Section 3 we provide an example of a particle system (Remark 2 (iii)) where the Wasserstein distance will not approach zero even though $\bar{m}_t^{x,y}$ does go to zero.

(2) In certain cases ([7]) one can show that $d(\xi_t^{x,i}, \xi_t^{y,i}) \leq c(x,y)f(t)$ for all $i \in \mathbb{N}$ and $t \in \mathbb{R}^+$ where $f(t) \to 0$ as $t \to \infty$. Such a uniform deterministic bound immediately yields

$$\bar{m}_t^{x,y} \leq c(x,y)f(t). \tag{2.14}$$

We will discuss examples in Subsection 3.2 to illustrate this.

Assuming that we have a uniform deterministic bound on $\bar{m}_t^{x,y}$ a perturbation estimate for the log-Laplace equation is immediate. Let $h : E \to \mathbb{R}$ be a Lipschitz function, i.e., there exists a constant $L_h > 0$ with $|h(x) - h(y)| \leq L_h d(x,y)$ for all $x, y \in E$. Let $X := (X_t)_{t \geq 0}$ be given by (2.3). Define $u : \mathbb{R}^+ \times E$ as

$$u(t,x) := -\log \mathbf{E}[e^{-<X_t^x, h>}]. \tag{2.15}$$

Lemma 2.1. *Assume that there exists a function $\varphi : \mathbb{R}^+ \times E^2 \to \mathbb{R}^+$ which is symmetric in the second and third coordinate, and such that $\bar{m}_t^{x,y} \leq \varphi(t,x,y)$. Then $t \geq 0$, and $x, y \in E$,*

$$\big|u(t,x) - u(t,y)\big| \leq -\log \mathbf{E}\big[e^{-L_h \varphi(t,x,y) P_t}\big]. \tag{2.16}$$

Proof. By (2.3) and the assumption of the l emma, we have

$$\begin{aligned} < X_t^y, h > &= P_t \cdot \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n h(\xi_t^{y,i}) \\ &\leq P_t \cdot \big[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n h(\xi_t^{x,i}) + L_h \bar{m}_t^{x,y}\big] \\ &\leq < X_t^x, h > + < X_t^x, L_h \varphi(t,x,y) > . \end{aligned} \tag{2.17}$$

In the last equation we have used the fact that $P_t =< X_t^x, 1 >=< X_t^y, 1 >$. Now, since $(\xi_t^1, ...)$ are exchangeable, for two non-negative functions h_1 and h_2, $< X_t, h_1 >$ and $< X_t, h_2 >$ are non-negatively correlated. Hence

$$\mathbf{E}[e^{-<X_t, h_1 + h_2>}] \geq \mathbf{E}[e^{-<X_t, h_1>}]\mathbf{E}[e^{-<X_t, h_2>}]. \tag{2.18}$$

So, we have

$$
\begin{aligned}
u(t,y) - u(t,x) &= \log \frac{\mathbf{E}[e^{-<X_t^x,h>}]}{\mathbf{E}[e^{-<X_t^y,h>}]} \\
&\le \log \frac{\mathbf{E}[e^{-<X_t^x,h>}]}{\mathbf{E}[e^{-<X_t^x,h>-<X_t^x,L_h\varphi(t,x,y)>}]} \\
&\le \log \frac{\mathbf{E}[e^{-<X_t^x,h>}]}{\mathbf{E}[e^{-<X_t^x,h>}]\mathbf{E}[e^{-<X_t^x,L_h\varphi(t,x,y)>}]} \\
&= -\log \mathbf{E}[e^{-L_h\varphi(t,x,y)P_t}].
\end{aligned} \tag{2.19}
$$

By symmetry in x and y, also $u(t,x) - u(t,y) \le -\log \mathbf{E}[e^{-L_h\varphi(t,x,y)P_t}]$. Hence we obtain the result. □

3. Applications and Examples

In this section we describe applications and examples. We begin with examples, from the literature, of processes that satisfy (P) and couplings that satisfy (S1) or (S2).

3.1. Examples of total mass processes.

In this subsection we discuss examples of P_t satisfying (P). These are part of the folklore in the literature. As we could not find references for all the examples, we present them with short proofs.

(Sub)-critical, β-stable, super-$(A,\mathcal{D}(A))$-process Let A be the generator of a Markov process ξ on E. Let the total mass, $P^{(\ell,\beta,\gamma)}$ be the (sub-)critical, β-stable super branching process with sub-criticality $\ell \ge 0$ and branching rate γ. That is, its Laplace transform is for $P_0 \ge 0, \gamma > 0$, $t > 0$, $0 < \beta \le 1$, and $\ell \ge 0$, given by

$$
-\log \mathbf{E}[e^{-\lambda P_t^{(\ell,\beta,\gamma)}}] = \begin{cases} \frac{e^{-\ell t}P_0}{(\lambda^{-\beta}+\frac{\gamma}{2\ell}(1-e^{-\ell\beta t}))^{\frac{1}{\beta}}}, & \text{if } \ell \ne 0, \\ \frac{P_0}{(\lambda^{-\beta}+\frac{\gamma\beta t}{2})^{\frac{1}{\beta}}}, & \text{if } \ell = 0 \end{cases} \tag{3.1}
$$

In this case X is the (sub-)critical, β-stable super-$(A,\mathcal{D}(A))$-process. The next lemma states that Condition (P) is satisfied for $P^{(\ell,\beta,\gamma)}$.

Lemma 3.1. *Let $P^{(\ell,\beta,\gamma)}$ be the total mass process of the (sub-)-critical, β-stable super-process X with sub-criticality $\ell \ge 0$ and branching rate γ. Then $P^{(\ell,\beta,\gamma)}$ satisfies (P).*

Proof. To see this, let us first assume that $\ell = 0$. Let $\mathcal{L}^a[P_t]$ be the law of P_t starting at a. In this case, we may use the following scaling:

$$
\mathcal{L}^\theta[P_t] = \mathcal{L}^{\alpha^{\frac{1}{\beta}}\theta}[\alpha^{-\frac{1}{\beta}}P_{\alpha t}], \quad \theta > 0, \alpha > 0. \tag{3.2}
$$

Let $\alpha \in]0,1[$ and define $\tau_\alpha := \inf\{t \geq 0 : \alpha^{-\frac{1}{\beta}} P_{\alpha t} = 0\}$.

$$\int_0^{\tau_1} \mathrm{d}s\, P_s^{-1} = \alpha^{1-1/\beta} \int_0^{\tau_\alpha} \mathrm{d}u\, \frac{\alpha^{1/\beta}}{P_{\alpha u}} \geq \int_0^{\tau_\alpha} \mathrm{d}u\, \frac{\alpha^{1/\beta}}{P_{\alpha u}}. \tag{3.3}$$

Let X is distributed like the left hand side of (3.3), where P is the critical, β-stable total mass process started in $P_0 = \theta$, while Y is distributed like $\int_0^{\inf\{t\geq 0:\, P_t=\theta\}} \mathrm{d}s\, \frac{1}{P_s}$ where P is the critical, β-stable total mass process started $P_0 = \alpha^{-1/\beta}\theta$. By the strong Markov property of P, (3.3) says therefore that $X \geq X + Y$ in distribution, i.e, for all $t \geq 0$, $\mathbf{P}\{X \leq t\} \geq \mathbf{P}\{X + Y \leq t\}$. For $\theta > 0$, we know that $Y > 0$ almost surely. Hence $\mathbf{P}\{X = \infty\} = 1$.

Now let $\ell > 0$. Then a simple calculation shows that $\hat{P}_t := e^{\ell t} P_t^{\ell,\beta,\gamma}$ is the critical, β-stable superprocess with time-inhomogeneous branching rate $\gamma e^{\ell t}$. Hence for each time $T \geq 0$,

$$(\ell T + 1)\hat{P}_{\ell^{-1}\log(\ell T+1)} \stackrel{d}{=} P_T^{0,\beta,\gamma}. \tag{3.4}$$

Therefore

$$\begin{aligned} \int_0^{\tau_1} \mathrm{d}u\, \big(P_u^{\ell,\beta,\gamma}\big)^{-1} &\geq \int_0^{\inf\{t\geq 0:\, \hat{P}_t=0\}} \mathrm{d}u\, \hat{P}_u^{-1} \\ &= \int_0^{\inf\{t\geq 0:\, \hat{P}_{\ell^{-1}\log(\ell t+1)}=0\}} \mathrm{d}s\, (\ell s+1)^{-1} \hat{P}^{-1}_{\ell^{-1}\log(\ell s+1)} \\ &= \int_0^{\inf\{t\geq 0:\, P^{0,\beta,\gamma}=0\}} \mathrm{d}u\, \big(P_u^{0,\beta,\gamma}\big)^{-1} = \infty. \end{aligned} \tag{3.5}$$

□

The supercritical β-stable superprocess. If P_t is the total mass process of a superprocess which does not die out with positive probability, then it is easy to see that the corresponding genealogical tree is not compact anymore. Indeed, if the Laplace transform of P satisfies (3.1) for a $\ell < 0$, then $e^{-\ell t/\beta} P_t$ is a martingale which converges to a nontrivial random variable Z, and $\int_0^\infty \mathrm{d}s P_s^{-1} < \infty$ on $Z \neq 0$. However note that, conditioned on the event $\{\tau^{\text{ext}} < \infty\}$, P_t has the same distribution as the sub-critical branching diffusion with $-\ell > 0$, and hence will satisfy (P).

Fleming-Viot process. If we condition the superprocess on having constant total mass, then the process $\bar{\mu}^{(x,y)}$ is a Fleming-Viot process with type space E. In this case (P) is trivially fulfilled.

Size-biased Feller diffusion. Conditioning the superprocess on non-extinction is equivalent to conditioning the total mass process not to hit zero. When P_t satisfies (3.1) with $\ell = 0$ and $\beta = 1$ it is well known that P_t is a solution of Feller's branching diffusion. Namely,

$$\mathrm{d}P_t = \sqrt{\gamma P_t}\,\mathrm{d}B_t, \quad t \geq 0, \tag{3.6}$$

where B_t is a Brownian motion. Likewise the conditioned total mass process $\hat{P}$ satisfies

$$d\hat{P}_t = \gamma dt + \sqrt{\gamma \hat{P}_t} dB_t, \quad t \geq 0, \tag{3.7}$$

i.e., its law equals the size-biased law of Feller's branching diffusion. In this case $\tau^{\text{ext}} = \infty$, and the following lemma ensures that $\hat{P} := (\hat{P}_t)_{t\geq 0}$ satisfies (P).

Lemma 3.2. *Let $\hat{P} := (\hat{P}_t)_{t\geq 0}$ be a version of the unique strong solution of (3.7). Then there exists a random variable Z with $\mathbf{P}\{Z > 0\} = 1$ and such that*

$$\liminf_{t\to\infty} \frac{1}{\log t} \int_0^t ds\, \hat{P}_s^{-1} \geq Z, \quad a.s. \tag{3.8}$$

In particular, $\hat{P}$ satisfies Condition (P).

Proof. Let $\mathcal{L}^\theta[P]$ and $\mathcal{L}^\theta[\hat{P}]$ denote the laws of a Feller diffusion and a size-biased Feller diffusion (compare (3.6) and (3.7)), respectively, started in $\theta \geq 0$. Then $\mathcal{L}^\theta[\hat{P}] = \mathcal{L}^\theta[P] * \mathcal{L}^0[\hat{P}]$, where $*$ denotes convolution. Therefore w.l.o.g. we can assume $\hat{P}_0 = 0$. It is easy to check that $\mathbf{E^0}[\hat{P}_t^n] = (n+1)!(\gamma t/2)^n$. Hence

$$\mathbf{P^0}\Big\{\frac{\hat{P}_t}{t} \geq x\Big\} = \int_x^\infty dy\, \big(\frac{\gamma}{2}\big)^2 y\, e^{-\gamma y}. \tag{3.9}$$

Since $\hat{P}_t$ is a size-biased martingale, $\hat{P}_t^{-1} := (1/\hat{P}_t)_{t\geq 0}$ is a martingale, and $(\frac{t}{\hat{P}_t})_{t\geq 0}$ is a non-negative submartingale, which therefore converges almost surely to a non-trivial random variable Z. The distribution of Z is given by the inverse of a Gamma$(2, \gamma/2)$-variable (compare the right hand side of (3.9)), and hence $Z > 0$, a.s.

Fix $\varepsilon \in]0,1[$. Then

$$\begin{aligned} \liminf_{t\to\infty} \frac{1}{\log t} \int_0^t ds\, \hat{P}_s^{-1} &\geq \liminf_{t\to\infty} \frac{1}{\log t} \int_{t^\varepsilon}^t ds\, \hat{P}_s^{-1} \\ &= \liminf_{t\to\infty} \int_\varepsilon^1 da\; t^a \cdot \hat{P}_{t^a}^{-1} = Z(1-\varepsilon), \quad \text{a.s.} \end{aligned} \tag{3.10}$$

Since ε was chosen arbitrarily, (3.8) follows clearly. □

3.2. **Examples of couplings.** In this subsection we give applications of the main theorem based on couplings of two spatial Markov processes. These are a reflecting Brownian motion in a convex domain $D \subset \mathbb{R}^2$ and a finite state Markov chain.

Super Reflected Brownian motion and Assumption (S1). Let $D \subset \mathbb{R}^2$ be a convex domain, and (B^1, B^2) be a planar Brownian motion with $B_0 \in \bar{D}$. Starting from a realization of B, by Theorem 2.1 in [16] there exists a uniquely determined solution (ξ, L) of

$$\xi_t = B_t + L_t \tag{3.11}$$

such that $L_t : \mathbb{R}^+ \to \mathbb{R}^2$ is a right continuous function with left limits of bounded variation with $L_0 = 0$. $\xi := (\xi_t)_{t\geq 0}$ is a strong Markov process with continuous paths living in $\bar{D}$ called *reflected Brownian motion*, and can be written as $\xi = g(B)$, for a measurable function g. Given g, a *synchronous coupling* between two copies of reflected Brownian motion is defined by letting $f(x, g(B)) := g(B + x)$, $x \in D$ (recall the definition of a coupling f from (2.4)).

Consider the super-reflected Brownian motion on D. That is, in our set up, $E = D$, ξ is the reflected Brownian motion on D and P_t satisfies (3.1).

Corollary 3.1. *Let $D \subset \mathbb{R}^2$, ξ^x and ξ^y be two synchronously coupled reflected Brownian motions on D, and X^x and X^y be two spatially coupled population models satisfying Condition (P) with motion processes ξ^x and ξ^y.*

(i) *If D is a convex domain with a C^2-boundary, and the curvature $K(x)$ of the boundary ∂D is bounded to below by a positive constant K then there exist constants $c, \mu > 0$ such that $\bar{m}_t^{x,y} \leq ce^{-\mu t}$.*
(ii) *If D is a polygonal domain or Lip domain then $\bar{m}_t^{x,y} \to 0$, a.s.*

Proof. (i) In [7] the authors make the following observation under the above assumption. For $x, y \in D$,

$$d(\xi_t^x, \xi_t^y) \leq d(x, y)\exp(-c\phi(t)), \tag{3.12}$$

where

$$\lim_{t\to\infty} \frac{\phi(t)}{t} = \frac{1}{2}\big[K^2 + \frac{1}{m(D)}\int_{\partial D} \sigma(\mathrm{d}y)K(y)\big] \tag{3.13}$$

with m denoting the Lebesgue measure. This was first observed in the PhD-thesis of Weerasinghe for the unit disc D. With the above and Remark 1 (2), part (i) follows.

(ii) In [3], Theorem 1.1 and its proof, it is shown that in such domains D, (S1) holds. Hence part (ii) now follows from Theorem 2.1. □

Semilinear Partial Differential Equation. Let X be a super-reflected Brownian motion in D as above with $X_0 = \delta_x$. It is well known that the log Laplace functional of X, $u : \mathbb{R}_+ \times D \to \mathbb{R}$ defined by (2.15), is a unique solution of the initial value problem with Neumann boundary conditions given by (1.1). We now present a convergence result for these solutions with Lipschitz initial conditions.

Corollary 3.2. *Let D be a convex planar domain with a C^2-boundary, such that the curvature $K(x)$ of the boundary ∂D is bounded to below by a positive constant K, $g : D \to \mathbb{R}$ be Lipschitz, and u be a solution to (1.1). Then there exists a constant C such that*

$$\big|u(t,x) - u(t,y)\big| \leq L_g \cdot C \cdot e^{-(\mu+\ell)t}, \tag{3.14}$$

where $\mu = \frac{1}{2}\big[K^2 + \frac{1}{m(D)}\int_{\partial D}\sigma(\mathrm{d}y)K(y)\big]$.

Proof. Consider X_t^x and X_t^y, the two spatially coupled super reflected Brownian motions as in the previous corollary. Let $X_0^x = \delta_x$ and $X_0^y = \delta_y$. Let $u(t,\cdot)$ be as in (2.15). From Corollary 3.1, we know that the assumptions of Lemma 2.1 are satisfied with the above μ. Using Lemma 2.1 we have that,

$$\big|u(t,x) - u(t,y)\big| \le -\log \mathbf{E}\big[e^{-L_g c e^{-\mu t} P_t}\big], \tag{3.15}$$

where the distribution of P_t is given by (3.1). Assume $\ell \ge 0$. An Itô calculation implies that $\mathbf{E}\big[e^{-L_g c e^{-\mu t} P_t}\big] = e^{-V(t)}$ where $V(t)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} V = -(\mu+\ell)V - e^{-\mu t}\frac{\gamma}{2} V^{1+\beta}, \quad V(0) = cL_g. \tag{3.16}$$

Solving this ordinary differential equation one obtains

$$\begin{aligned} V(t) &= \frac{cL_g e^{-(\mu+\ell)t}}{\big(1 + \frac{\gamma (cL_g)^\beta}{2(\mu\frac{1+\beta}{\beta}+\ell)}(1 - e^{-(\mu\frac{1+\beta}{\beta}+\ell)\beta t})\big)^{1/\beta}} \\ &\le cL_g e^{-(\mu+\ell)t}. \end{aligned} \tag{3.17}$$

Using the above bound and substituting into (3.15) we have the result. □

Super Markov Chains and Assumption (S2). A class of couplings called "Efficient Markovian Couplings" is considered in [5]. Here the Markov chains on a finite state space along with various couplings of reflected Brownian motions satisfying (S1) and (S2) are discussed. Using these results we are able to present a coupling result on a class of finite-dimensional diffusions on $\mathbb{R}_n^+$. For $x \in \mathbb{R}_n^+$, let $\| x \|_n$ be the usual Euclidean norm. Let $m \in \{1,2,\ldots,n\}$ and $\gamma, \{q_{ml}\}_{l,m=1}^n$ be non-negative constants. Let X be a solution of

$$\mathrm{d}X_t^m = \sum_{l=1}^n q_{ml} X_t^l \mathrm{d}t + \gamma \mathrm{d}t + \sqrt{\gamma X_t^m}\,\mathrm{d}B_t^m, \quad X_0^m = x^m \ge 0, \tag{3.18}$$

where B_t^m are independent Brownian motions.

Corollary 3.3. *Let X_t and Y_t be a solution of (3.18) starting at $x = (x^1, \ldots x^n)$ and $y = (y^1, \ldots, y^n)$, respectively, with $x^l, y^l \ge 0$. Assume*

1. $P_0 := \sum_{l=1}^n x^l = \sum_{l=1}^n y^l$, *and*
2. $q := \{q_{ml}\}_{m,l=1\ldots n}$ *are transition rates of a Markov chain on $\{1,\ldots,n\}$ satisfying Assumption (S2).*

Then there exits a coupling of X_t and Y_t such that

$$\eta := \inf\{t \ge 0 : \| X_t - Y_t \|_n = 0\} < \infty, \quad a.s. \tag{3.19}$$

Proof. First if $P_0 = 0$ then $\eta = 0$, and we are done. Now assume that $P_0 = 1$. Set $E := \{1,2,\ldots,n\}$. Then $\mu_x := \sum_{l=1}^n x^l \delta_l$ is a probability measure on E. Consider ξ to be a continuous Markov chain with state space E with transition rates given by q_{ml}. Consider the total mass process P satisfying (3.7) with $\beta = 1$. Finally construct X as in (2.3) with $X_0 = \mu$. Using the identification of $M_F(E) = \mathbb{R}^+_n$, it is easy to identify X as a solution of (3.18) with $X_0 = x$. Similarly, we can construct a solution Y of (3.18) with $Y_0 = y$. For the proof of this corollary, we will view these diffusions as finite measures on E. Assume that the rates q_{ml} satisfy Assumption (S2). Notice that two particles starting at any l and m, respectively, performing jumps according to the Markov chain with transition rates q_{ml} will land at the same position in finite time, almost surely.

Since $X_0 = \mu_x$ and $Y_0 = \mu_y$ we shall start our exchangeable motion process with starting points in E sampled according to μ_x and μ_y, respectively. We shall refer to the motion process governing X and Y as ξ^{μ_x} and ξ^{μ_y}, respectively. Note that not all of the particles in each level will have the same starting point anymore. As we have finitely many starting points and the Markov chain satisfy Assumption (S2), we can say that

$$S' := \max_{(l,m)\in E\times E} \inf\{t \geq 0 : d(\xi^{\mu_x}_t, \xi^{\mu_y}_t) = 0, \xi^{\mu_x}_0 = l, \xi^{\mu_y}_0 = m\} < \infty, \quad \text{a.s.} \tag{3.20}$$

So we have a uniform finite stopping time over all starting points in E such that the position of the two coupled Markov chains are the same. Therefore now proceed to apply Theorem 2.1 with Assumption (S2) holding with S' (as opposed to S) to see that $\eta < \infty$, a.s.

So far we have assumed that $P_0 = 1$. Suppose now that $P_0 > 0$. It is easy to see that if q satisfies Assumption (S2) then so will the chain with transition rates $\frac{q}{P_0}$. Hence we can do a scale change by $\frac{1}{P_0}$, repeat the above argument to conclude the proof of the corollary. □

We conclude this section with some remarks.

Remark 2

(i) In [2], couplings of reflected Brownian motions were used to analyse the "hot spots" conjecture. A key fact about the couplings is that the geometry of the initial starting location is preserved. It is easy to see from the proof of Theorem 2.1 that this translates to the measure-valued setting as well. This fact was established earlier in [1] using historical processes.

(ii) Corollary 3.3 presents a new way to couple diffusions in $\mathbb{R}^n_+$ provided they have a particle representation. Even though Corollary 3.3 has been stated with Assumption (S2), this can be easily adapted to the situation when Assumption (S1) holds instead of Assumption (S2).

If P_t satisfied (3.1) with $\beta = 1$ and $\ell \geq 0$ then the diffusions considered in the corollary would be a solution to

$$\mathrm{d}X^m_t = \sum_{l=1}^{n} q_{ml} X^l_t \mathrm{d}t - \ell X^m_t \mathrm{d}t + \gamma \mathrm{d}t + \sqrt{\gamma X^m_t}\,\mathrm{d}B^m_t, \quad X^m_0 = x^m \geq 0. \tag{3.21}$$

The coupling result would hold for these diffusions as well.

Notice that our examples are not covered by those obtained in [15] for multidimensional diffusions by reflecting the increments of the driving Brownian motion. Their result applies to diffusions whose diffusion matrix are perturbations of a constant matrix and whose drift is such that the ordinary differential equation given by ignoring the noise ensures that the distance between two solutions decreases in time.

(iii) We present an example to show that the Wasserstein distance may not be the most appropriate distance to consider. In [4], a synchronous coupling of two reflected Brownian motions in a smooth domain $D \subset \mathbb{R}^2$ is considered. Let ξ^x (and ξ^y, respectively,) be the synchronously coupled reflected Brownian motions in D. They show that for certain planar domains D, there exists $\mu > 0$ such that for all x, y,

$$\lim_{t\to\infty} \frac{\log d(\xi^x_t, \xi^y_t)}{t} = -\mu.$$

Consider the following branching reflected Brownian motions. Each particle lives an exponential time (with mean 1) and then dies. Upon dying it is replaced by K particles at the site of its death. These K particles perform independent reflected Brownian motions in D. Assume that we start with one particle at $x \in D$. For $t \geq 0$, let I_t be the index of particles alive at time t. Let $\{\xi^{x,i}_t : i \in I_t\}$ represent the particle positions. Define $X^x_t = \sum_{i=1}^{|I_t|} \delta_{\xi^{x,i}_t}$. Let X^y be another branching reflected Brownian motion such that the branching events are exactly the same as X^x, while the reflected Brownian motion $\xi^{y,\cdot}$ are synchronously coupled with $\xi^{x,\cdot}$. Let $P_t =< X^x_t, 1 >=< X^y_t, 1 >$ be the total mass. Note that by the martingale convergence theorem there exists a non-trivial random variable Z such that

$$\lim_{t\to\infty} e^{-t\log K} P_t = Z, \quad \text{a.s.,} \tag{3.22}$$

where Z is a non-negative random variable. Now for $\epsilon > 0$, for t large enough,

$$\begin{aligned} d_{\mathrm{W}}(X^x_t, X^y_t) &\geq P_t \min_{i\in I_t}\{d(\xi^{x,i}_t, \xi^{y,i}_t)\} \\ &\geq P_t e^{-(\mu+\epsilon)t} \\ &\geq (Z-\epsilon)e^{t\log K} e^{-(\mu+\epsilon)t}. \end{aligned} \tag{3.23}$$

Now choosing K suitably large one sees that $d_{\mathrm{W}}(X^x_t, X^y_t)$ does not go to zero almost surely even though clearly $\sup_{i\in I_t} d(\xi^{x,i}_t, \xi^{y,i}_t) \to 0$, a.s.

4. Proof of Theorem 2.1

Proof. (i) Fix a sequence $(t_n) \uparrow \infty$, as $n \to \infty$, and $\varepsilon > 0$. Given the total population process P, choose a sequence $(\delta_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} (t_n - \delta_n) = \infty, \tag{4.1}$$

and

$$\lim_{n\to\infty}\int_{t_n-\delta_n}^{t_n}\mathrm{d}s\,\frac{1}{P_s}=\infty. \tag{4.2}$$

According to Assumption (S1) or (S2) there exists a coupling $f: E\times D([0,\infty[,E)\to D([0,\infty[,E)$ (recall from (2.4)) such that for all $x,y\in E$ there is a finite stopping time satisfying Assumption (S1) or (S2), respectively.

Let $\{\xi^i : i\in\mathbb{N}\}$ be i.i.d. copies of the motion process ξ, $\hat{V}:=\{\hat{V}^{i,j} : 1\le i<j<\infty\}$ be a family of independent unit rate Poisson processes, and P be a total mass process satisfying Assumption (P). For $i,j\in\mathbb{N}$, let $V^{i,j}$ be the counting process given by the following relation:

$$V^{i,j}\{t\}:=\begin{cases}1, & \text{if} \quad \hat{V}^{i,j}\{\int_0^t \mathrm{d}sP_s^{-1}\}=1,\\ 0, & \qquad\text{else.}\end{cases} \tag{4.3}$$

For $i\in\mathbb{N}$ and $0\le s\le t<\infty$, let $A^i_{t,s}$ denote the "ancestor" at time s of level i at time t. That is,

$$A^i_{t,s}:=j\in\mathbb{N}, \qquad \text{iff there exist a "path from } (j,s) \text{ to } (i,t)\text{"}, \tag{4.4}$$

where for each $s\le t$ and $1\le j\le i<\infty$ we say there is a path from (j,s) to (i,t) if there exist $s:=s_0<s_1<...<s_n=:t$ and $j:=j_0<j_1<...<j_n=:i$ such that $V^{j_k,j_{k+1}}\{s_{j_k}\}=1$, and $\sum_{j<j_{k+1}}V^{j,j_k}]s_{j_k},s_{j_{k+1}}[=0$. (Note that there is always exactly one path joining (j,s) and (i,t).)

Notice that the process

$$X_t:=P_t\cdot\big(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\delta_{\xi_t^{A^i_{t,0}}}\big)1_{\{\tau^{\mathrm{ext}}<t\}}$$

is a version of the process defined in (2.2) and (2.3).

Denote by

$$\Gamma_{t,s}:=\{\{j\in\mathbb{N}: A^j_{t,s}=A^i_{t,s}\} : i\in\mathbb{N}\}, \tag{4.5}$$

the partition of $\mathbb{N}$ into the family patches of individuals at time t sharing a common ancestor at time s. Condition (P) ensures that for all $0\le s<t<\infty$, $\#\Gamma_{t,s}<\infty$, almost surely. In particular, by Theorem 5.1 in [11], $(\Gamma_{t,s\wedge t})_{s\ge0}=(K_{s\wedge t})_{s\ge0}$ in distribution, where $K:=(K_t)_{t\ge0}$ is Kingman's coalescent.

We will show that Assumption (S1) yields that $(g(t_n-\delta_n,x,y)\bar{m}^{x,y}_{t_n-\delta_n})_{n\in\mathbb{N}}$ is eventually (depending on $x,y\in E$) smaller than a given ε, and therefore $g(t,x,y)\bar{m}^{x,y}_t\to0$, almost surely, as $n\to\infty$.

Let T^i be a stopping time such that $d(\xi^{i,x}_t,\xi^{i,y}_t)_{t\ge0}$ is non-increasing on $[T^i,\infty[$. We consider the "bad events" that at time $t_n-\delta_n$ we can not predict that $g(t_n,x,y)\bar{m}^{x,y}_{t_n}\le\varepsilon$ by looking back a time δ_n. We therefore set

$$B_n := \cup_{i\in\mathbb{N}} B_{n,i}, \tag{4.6}$$

with

$$B_{n,i} := \big\{t_n - \delta_n < T^{A^i_{t_n,t_n-\delta_n}}\big\} \cup \big\{g(t_n,x,y)\, d(\xi^{A^i_{t_n,t_n-\delta_n},x}_{t_n-\delta_n}, \xi^{A^i_{t_n,t_n-\delta_n},y}_{t_n-\delta_n}) \geq \varepsilon\big\}. \tag{4.7}$$

Then

$$\begin{aligned} \mathbf{P}(B_n) &= \mathbf{E}\big[\mathbf{P}(\cup_{\pi\in\Gamma_{t_n,t_n-\delta_n}} \cup_{i\in\pi} B_{n,i} | \Gamma_{t_n,t_n-\delta_n})\big] \\ &= \mathbf{E}\big[\mathbf{P}(\cup_{\pi\in\Gamma_{t_n,t_n-\delta_n}} B_{n,\min\pi} | \Gamma_{t_n,t_n-\delta_n})\big], \end{aligned} \tag{4.8}$$

where we have used that given $\Gamma_{t_n,t_n-\delta_n}$, $B_{n,i} = B_{n,j}$ for all $i,j \in \pi \in \Gamma_{t_n,t_n-\delta_n}$. Moreover, given $\Gamma_{t_n,t_n-\delta_n}$, the events $\{B_{n,\min\pi};\, \pi \in \Gamma_{t_n,t_n-\delta_n}\}$ are all independent, and have the same probabilities. Hence

$$\mathbf{P}(B_n) = \mathbf{E}\big[1 - \mathbf{P}(B^c_{n,1})^{\#\Gamma_{t_n,t_n-\delta_n}}\big]. \tag{4.9}$$

In order to be in a position where we may apply the Borel-Cantelli lemma, we would like to have that $\mathbf{P}(B_n)$ is summable along a subsequence. Indeed, by Assumption (S1), $\mathbf{P}(B^c_{n,1}) \to 1$ as $n \to \infty$, and hence

$$\lim_{n\to\infty} \mathbf{P}(B^c_{n,1})^{\#\Gamma_{t_n,t_n-\delta_n}} = 1, \quad \text{a.s.} \tag{4.10}$$

In particular, we may choose a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ such that $\mathbf{P}(B_{n_k})$ is summable over $k \in \mathbb{N}$. Then $T^i_{t_{n_k}} \leq t_{n_k} - \delta_{n_k}$ and $g(t_{n_k},x,y)d(\xi^{i,x}_{t_{n_k}/2}, \xi^{i,y}_{t_{n_k}/2}) < \varepsilon$, for all $i \in \mathbb{N}$ and k sufficiently large. In particular, for all sufficiently large $k \in \mathbb{N}$,

$$g(t_{n_k},x,y)\bar{m}^{x,y}_{t_{n_k}} < \varepsilon, \tag{4.11}$$

and therefore $g(t_{n_k},x,y)\bar{m}_{t_{n_k}} \to 0$ almost surely, as $k \to \infty$. Since any subsequence of $(t_n)_{n\in\mathbb{N}}$ contains a subsequence along which $(g(t,x,y)\,\bar{m}^{x,y}_t)_{t\geq 0}$ tends to zero, and since $(t_n)_{n\in\mathbb{N}}$ was chosen arbitrarily, we have shown that $g(t,x,y)\bar{m}^{x,y}_t \to 0$, almost surely, as $t \to \infty$.

(ii) Under Assumption (S2), for each $i \in \mathbb{N}$, $S^i := \inf\{t \geq 0 :\, \xi^{i,x}_t = \xi^{i,y}_t\} < \infty$, almost surely. Then the "bad events" are

$$C_n := \cup_{i\in\mathbb{N}}\{S^i > t_n - \delta_n\}. \tag{4.12}$$

Similar to (4.8), we have

$$\mathbf{P}(C_n) = \mathbf{E}\big[1 - \mathbf{P}(S^1 > t_n - \delta_n)^{\#\Gamma_{t_n,t_n-\delta_n}}\big], \tag{4.13}$$

which can be made summable along a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ under (S2), and hence there exists a K such that for all $k \geq K$, $\bar{m}^{x,y}_{t_{n_k}} = 0$. Since the couplings are successful, zero is a trap for $\bar{m}^{x,y}_{t_{n_k}}$, and therefore $\bar{m}^{x,y}_t = 0$ for all $t \geq t_{n_K}$. □

References

[1] **Athreya, S. (1999)** Monotonicity property for a class of semilinear differential equations, Seminar de's probabilite's XXXIV (2000) .

[2] **Bañuelos, R., and Burdzy, K. (1999)** On the 'hot spots' conjecture of J. Rauch, J. Funct. Anal., **164**, 1-33.

[3] **Burdzy, K., and Chen, Z.Q (2002)** Coalescence of synchronous couplings, Prob. Theory and Related Fields, **123**, 553-578.

[4] **Burdzy, K., Chen, Z.Q, and Jones, P. (2004)** Synchronous couplings of reflected Brownian motion in smooth domains, Paper in preparation.

[5] **Burdzy, K., and Kendall, W.S. (2000)** Efficient Markovian couplings: examples and counterexamples, Ann. Appl. Probab., **10**, 362-409.

[6] **Cranston, M., and Le Jan, Y. (1989)** On the non-coalescence of a two point Brownian motion reflecting on a circle, Ann. Inst. Henri Poincaré, **25(2)**, 99-107.

[7] **Cranston, M., and Le Jan, Y. (1990)** Noncoalescence for the Skorochod equation in a convex domain of $\mathbb{R}^2$. Probab. Th. Rel. Fields, **87**, 241-252.

[8] **Dawson, D.A. (1993)** Measure-valued Markov processes, École d'Été de Probabilités de Saint Flour XXI – 1991, Lect. Notes in Math. **1541**, 1-260, Springer-Verlag.

[9] **Dawson, D.A. and Perkins, E.A. (1991)** Historical processes, Memoirs of the American Mathematical Society 93, **454**.

[10] **Donnelly, P. and Kurtz, T.G. (1996)** Particle representation of the Fleming-Viot measure-valued diffusion, Ann. Probab., **24(2)**, 698-742.

[11] **Donnelly, P. and Kurtz, T.G. (1999)** Particle representations for measure-valued population models, Ann. Probab., **27(1)**, 166-205.

[12] **Ethier, S.N., and Kurtz, T.G. (1998)** Coupling and ergodic theorems for Fleming-Viot processes, Ann. Probab., **26(2)**, 533-561.

[13] **Evans, S.N. (1993)** Two representations of a conditioned superprocess, Proc Roy Soc Ed, Sec. A, **123**, 959-971.

[14] **Greven, A., Limic, V., and Winter, A. (2004)** Representation theorems for interacting Moran models, interacting Fisher-Wright models and applications, *Preprint*.

[15] **Lindvall, T., and Rogers, L.C.G. (1986)** Coupling of multidimensional diffusions by reflection, Ann. Probab., **11(3)**, 860-872.

[16] **Tanaka, H. (1979)** Stochastic differential equation with reflecting boundary condition in convex regions, Hiroshima Math. J., **9**, 405-443.

Siva Athreya, 7, SJS Sansanwal Marg, Indian Statistical Institute, New Delhi 110016, India.

E-mail address: athreya@isid.ac.in

Anita Winter, Mathematisches Institut, Universität Erlangen–Nürnberg, Bismarckstrasse $1\frac{1}{2}$, 91054 Erlangen, Germany.

E-mail address: winter@mi.uni-erlangen.de