CONTINUUM–SITES STEPPING–STONE MODELS, COALESCING EXCHANGEABLE PARTITIONS, AND RANDOM TREES

PETER DONNELLY, STEVEN N. EVANS, KLAUS FLEISCHMANN, THOMAS G. KURTZ, AND XIAOWEN ZHOU

Abstract. Analogues of stepping–stone models are considered where the site–space is continuous, the migration process is a general Markov process, and the type–space is infinite. Such processes were defined in previous work of the second author by specifying a Feller transition semigroup in terms of expectations of suitable functionals for systems of coalescing Markov processes. An alternative representation is obtained here in terms of a limit of interacting particle systems. It is shown that, under a mild condition on the migration process, the continuum–sites stepping–stone process has continuous sample paths. The case when the migration process is Brownian motion on the circle is examined in detail using a duality relation between coalescing and annihilating Brownian motion. This duality relation is also used to show that a random compact metric space that is naturally associated to an infinite family of coalescing Brownian motions on the circle has Hausdorff and packing dimension both almost surely equal to $\frac{1}{2}$ and, moreover, this space is capacity equivalent to the middle–$\frac{1}{2}$ Cantor set (and hence also to the Brownian zero set).

1. Introduction

Stepping–stone models originally arose in population genetics. The simplest version can be described as follows. There is a a finite or countable collection of sites (the *site–space*). At each site there is a finite population. Each population is composed of individuals who can be one of two possible genetic types, say A or B. At each site the genetic composition of the population evolves via a continuous–time *resampling* mechanism. Independently of each other, individuals *migrate* from one site to another according to a continuous–time Markov chain (the *migration chain*) on the site–space.

If the number of individuals at each site becomes large, then, under appropriate conditions, the process describing the proportion of individuals of type A at the various sites converges to a diffusion limit. This limit can be thought of informally as an ensemble of *Fisher–Wright diffusions* (one diffusion at each site) that are coupled together with a drift determined by the jump rates of the migration chain (see, for example, [Shi80]).

1991 *Mathematics Subject Classification.* Primary 60K35; Secondary 60G57, 60J60.

Key words and phrases. coalesce, partition, right process, annihilate, dual, diffusion, exchangeable, vector measure, tree, Hausdorff dimension, packing dimension, capacity equivalence, fractal.

Donnelly supported in part by NSF grant DMS-9505129, and UK EPSRC grants B/AF/1255 and GR/M14197.

Evans supported in part by NSF grant DMS-9703845.

Kurtz supported in part by NSF grant DMS-9626116.

Zhou supported in part by NSF grant DMS-9703845 and a Michel and Liné Loève Fellowship.

A natural refinement of this two–type diffusion model, considered in [Han90, DGV95], is the corresponding *infinitely–many–types* model. Here the Fisher–Wright processes at each site are replaced by mutationless *Fleming–Viot processes* of evolving random probability measures on a suitable uncountable *type–space* (typically the unit interval $[0,1]$).

Much of the research on such interacting Fisher–Wright and Fleming–Viot diffusion models (see, for example, [BCG86, FG94, FG96, Kle95]) has centred on their *clustering* behaviour in the case when the space of sites is either the integer lattice $\mathbb{Z}^d$ or a discrete hierarchical group and the migration chain is a random walk. That is, one asks how regions where "most of the populations are mostly of one type" grow and interact with each other. The primary tool for analysing this behaviour is the *duality* (in the sense of duality of martingale problems) between these models and sytems of delayed coalescing random walks that was first exploited by [Shi80].

One of the factors that lead to interesting clustering is the scaling behaviour of the migration process. However, because random walks on $\mathbb{Z}^d$ in the domain of attraction of a stable law and their analogues on discrete hierarchical group only have approximate scaling, the role that scaling plays is somewhat obscured. In order to make the effect of scaling clearer, related two–type models were considered in [EF96] in the hierarchical group setting. In essence, the processes in [EF96] are the result of taking a further limit in which one "stands back" from the site–space so that the discrete hierarchical group approaches a continuous one and the random walk converges to a "stable" Lévy process on the continuous hierarchical group that does have exact rescaling. These *continuum–sites, two–type stepping–stone models* have as their state–space the collection of measurable functions x from the site–space (that is, the continuous hierarchical group) into $[0,1]$. For a state x and site e, the value $x(e)$ is interpreted intuitively as the proportion of the population at the site that is of type A.

One of the noteworthy feature of [EF96] is that the limit models are defined by specifying moment–like quantities for the associated Feller transition semigroup in terms of systems of *delayed or instantaneously coalescing Lévy processes*, using formulae that are analogues of the duality relations between the discrete–sites models and delayed coalescing random walks mentioned above (see Theorems 3 and 4 of [EF96]). In particular, the limit models are **not** defined infinitesimally via a generator, SDE/SPDE, or martingale problem formulation analogous to that of the discrete–sites models. We note, however, that it should be possible to "stand back" in a similar manner from a discrete–sites model where the migration chain is simple random walk on $\mathbb{Z}$ and obtain the process considered in [MT95]: this process is constructed there as an SPDE on $\mathbb{R}$ but is also dual to delayed coalescing Brownian motions via the same sort of formulae considered in [EF96]. However, the processes in [EF96] that have their semigroups defined in terms of instantaneously coalescing Lévy processes do not appear to have even a very informal interpretation as SPDE–like objects. Rather, a typical value for such a process is a function x such that $x(e) \in \{0,1\}$ for all sites e, and so such processes are more like continuum analogues of particle systems (see Theorem 6 of [EF96]).

The programme of defining continuum–sites models in terms of "duality" formulae using instantaneously coalescing Markov processes was continued in [Eva97] (see Section 4 below for a recapitulation). There the infinitely–many–types case was considered and the migration process (that is, the process used to build the coalescing system) was taken to be a general *Borel right process* subject only to a duality condition (with duality here taken in the sense of the general theory of Markov processes). Now a state of the process, which we denote from now on by X, will be a function x from the site–space into the collection of probability measures on an uncountable type–space. For a state x, a site e and a measurable subset G of the type–space, the value $x(e)(G)$ is interpreted intuitively as the proportion of the population at the site possessing types from G.

We give a more concrete description of the infinitely–many–types, continuum–sites process X in Section 5. Under suitable conditions on the migration process, we show (at the level of convergence of finite–dimensional distributions) that X is the *high–density limit* of a family of *particle systems* with the following description. The particles move about in the Cartesian product of the site–space and the type–space. The site–space–valued components of the particles evolve according to independent copies of the migration process. When particles collide in the site–space a type is chosen at random from the types of the particles participating in the collision and the types of all the participating particles are changed to this randomly chosen type.

One of the open problems left by [Eva97] was to determine conditions on the migration under which the process X (which is again a Feller process) has *continuous* rather than just càdlàg sample paths. In Theorem 7.2, we establish the sufficiency of a mild condition to the effect that the coalescing system doesn't coalesce too rapidly. The condition holds, for example, for all Lévy processes on $\mathbb{R}$.

By the same argument as in the proof of Proposition 5.1 of [Eva97], it is possible to show that if the migration process is a stable process on the *circle* $\mathbb{T}$ that hits points, then, for fixed $t > 0$, there almost surely exists a random countable subset $\{k_1, k_2, \dots\}$ of the type–space such that for Lebesgue almost all $e \in \mathbb{T}$ the probability measure $X_t(e)$ is a point mass at one of the k_i. That is, rather loosely speaking, at each site all individuals in the population have the same type and the total number of types seen across all sites is countable. We improve this result in Theorem 10.2 for the case of *Brownian motion migration* on $\mathbb{T}$ by showing that the total number of types is, in fact, almost surely finite and such a result holds simultaneously at all positive times rather than just for fixed times.

The primary tool used in the proof of Theorem 10.2 is a *duality* relation between systems of *coalescing* and *annihilating Brownian motions* that is developed in Section 9. This relation enables us to perform detailed computations with the system of coalescing Brownian motions that begins with countably many particles independently and uniformly distributed on $\mathbb{T}$.

The latter process is an interesting object in its own right. In particular, it can be used to define a *random metric* on the positive integers by declaring that the distance between i and j is the time until the descendents of the i^{th} and j^{th} particles at time zero coalesce. In Theorem 11.2 we adapt the methods of [Eva98] to show that the completion of the integers in this metric is almost surely *compact*, with *Hausdorff* and *packing dimensions* both equal to $\frac{1}{2}$. Moreover, this space is *capacity equivalent* to the middle–$\frac{1}{2}$ *Cantor set* (and hence also to the Brownian zero set).

Notation 1.1. Write $\mathbb{N} := \{1, 2, \dots\}$. We will adopt the convention throughout that the infimum of a subset of $\mathbb{R}$ or $\mathbb{N}$ is defined to be ∞ when the subset is empty.

2. Coalescing Markov processes and labelled partitions

Suppose that E is a *Lusin space* and that (Z, P^z) is a *Borel right process* on E with semigroup $\{P_t\}_{t\geq 0}$ satisfying $P_t 1 = 1$, $t \geq 0$, so that Z has infinite lifetime (see [Sha88] for a discussion of Lusin spaces and Borel right processes). Suppose that there is another Borel right process $(\hat{Z}, \hat{P}^z)$ with semigroup $\{\hat{P}_t\}_{t\geq 0}$ and a diffuse, Radon measure $m \neq 0$ on $(E, \mathcal{E})$ such that for all non-negative Borel functions on f, g on E we have $\int m(de)\, P_t f(e) g(e) = \int m(de)\, f(e) \hat{P}_t g(e)$ (our definition of Radon measure is that given in Section III.46 of [DM78]). The space E is the *site–space* and $\hat{Z}$ is the *migration process* for the continuum–sites stepping–stone model X, whereas Z will serve as the basic motion in the coalescing systems "dual" to X.

We remark that our assumption on the Markov processes Z and $\hat{Z}$ is not quite the usual notion of *weak duality* with respect to m (see, for example, Section 9 of [GS84]); in order for weak duality to hold we would also require that P^m–a.s. (resp. $\hat{P}^m$–a.s.) the left–limit $Z(t-)$ (resp. $\hat{Z}(t-)$) exists for all $t > 0$.

The following notation will be convenient for us. Given $\mathbf{e} = (e_1, \ldots, e_n) \in E^n$, for some $n \in \mathbb{N}$, with $e_i \neq e_j$ for $i \neq j$, let $\mathbf{Z}^{\mathbf{e}} = (Z^{e_1}, \ldots, Z^{e_n})$ be an E^n–valued process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that Z^{e_i} has the distribution of Z under P^{e_i} and $Z^{e_1}, \ldots, Z^{e_n}$ are independent.

We now define the *system of coalescing Markov processes* $\check{\mathbf{Z}}^{\mathbf{e}}$ associated with $\mathbf{Z}^{\mathbf{e}}$. Adjoin a point $\dagger$, the *cemetery*, to E to form $E^\dagger := E \cup \{\dagger\}$. Construct the $(E^\dagger)^n$–valued process $\check{\mathbf{Z}}^{\mathbf{e}} = (\check{Z}_1^{\mathbf{e}}, \ldots, \check{Z}_n^{\mathbf{e}})$ inductively as follows. Suppose that times $0 =: \tau_0 \leq \ldots \leq \tau_k \leq \infty$ and sets $\{1, \ldots, n\} =: \Theta_0 \supseteq \ldots \supseteq \Theta_k \supseteq \{1\}$ have already been defined and that $\check{\mathbf{Z}}^{\mathbf{e}}$ has already been defined on $[0, \tau_k[$. If $\tau_k = \infty$, then just set $\tau_{k+1} := \infty$ and $\Theta_{k+1} := \Theta_k$. Otherwise, put

$$\tau_{k+1} := \inf\{t > \tau_k : \exists i, j \in \Theta_k,\, i \neq j,\, Z^{e_i}(t) = Z^{e_j}(t)\}, \tag{2.1}$$

$$\Theta_{k+1} := \begin{cases} \Theta_k, & \text{if } \tau_{k+1} = \infty, \\ \{i \in \Theta_k : \nexists j < i,\, j \in \Theta_k,\, Z^{e_i}(\tau_{k+1}) = Z^{e_j}(\tau_{k+1})\}, & \text{otherwise,} \end{cases} \tag{2.2}$$

and

$$\check{Z}_i^{\mathbf{e}}(t) := \begin{cases} Z^{e_i}(t),\ \tau_k \leq t < \tau_{k+1}, & \text{if } i \in \Theta_k, \\ \dagger,\ \tau_k \leq t < \tau_{k+1}, & \text{otherwise.} \end{cases} \tag{2.3}$$

In other words, the coordinate processes of the coalescing Markov process $\check{\mathbf{Z}}^{\mathbf{e}}$ evolve as independent copies of Z until they collide. When two or more coordinate processes collide (which happens at one of the times τ_ℓ with $0 < \tau_\ell < \infty$), the one with the smallest index "lives on" while the other coordinates involved in the collision are sent to the cemetery $\dagger$. The set Θ_k is the set of coordinates that are still alive at time τ_k. As the following lemma shows, for $m^{\otimes n}$–a.e. $\mathbf{e}$ almost surely only one coordinate process of $\mathbf{Z}^{\mathbf{e}}$ is sent to the cemetery at a time in the construction of $\check{\mathbf{Z}}^{\mathbf{e}}$. (Recall that (Z, P^z) is said to be a *Hunt process* if Z has càdlàg sample paths and is also *quasi–left–continuous*; that is, if whenever $T_1 \leq T_2 \leq \ldots$ are stopping times for Z and $T = \sup_n T_n$, then $P^z\{\lim_n Z(T_n) = Z(T),\, T < \infty\} = P^z\{T < \infty\}$ for all $z \in E$.)

Lemma 2.1. *Let Y be an E–valued Markov process on some probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ with the same law as Z under $P^q := \int_E q(dz)\, P^z$, where q is a probability measure on $(E, \mathcal{E})$ that is absolutely continuous with respect to m. Let (T, V) be a $[0, \infty[\times E$–valued random variable that is independent of Y. Then $\bar{\mathbb{P}}\{Y(T) = V\} = 0$. Moreover, if Z is Hunt process, then $\bar{\mathbb{P}}\{Y(T-) = V\} = 0$, also. A similar result holds with Y replaced by a process $\hat{Y}$ with the same law as $\hat{Z}$ under $\hat{P}^q := \int_E q(dz)\, \hat{P}^z$.*

Proof. For fixed $t \geq 0$ and $v \in E$ we have, writing h for the Radon–Nikodym derivative of q with respect to m,

$$\bar{\mathbb{P}}\{Y(t) = v\} = \int_E m(dz)\, h(z) P_t \mathbf{1}_{\{v\}}(z) = \int_E m(dz)\, \mathbf{1}_{\{v\}}(z) \hat{P}_t h(z) = 0 \tag{2.4}$$

by the duality assumption and the assumption that m is diffuse. Moreover, under the Hunt assumption,

$$\bar{\mathbb{P}}\{Y(t) \neq Y(t-)\} = 0. \tag{2.5}$$

The result now follows by Fubini. □

It will be convenient to embellish $\check{\mathbf{Z}}^{\mathbf{e}}$ somewhat and consider an enriched process $\zeta^{\mathbf{e}}$ defined below that keeps track of which particles have collided with each other.

Let Π_n denote the set of *partitions* of $\mathbb{N}_n := \{1, \ldots, n\}$. That is, an element π of Π_n is a collection $\pi = \{A_1, \ldots, A_h\}$ of subsets of $\mathbb{N}_n$ with the property that $\bigcup_i A_i = \mathbb{N}_n$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. The sets $A_1, \ldots A_h$ are the *blocks* of the partition π. Equivalently, we can think of Π_n as the set of equivalence relations on $\mathbb{N}_n$ and write $i \sim_\pi j$ if i and j belong to the same block of $\pi \in \Pi_n$.

An *E–labelled partition* of $\mathbb{N}_n$ is a collection

$$\lambda = \{(A_1, e_{A_1}), \ldots, (A_h, e_{A_h})\}, \tag{2.6}$$

with $\{A_1, \ldots, A_h\} \in \Pi_n$, $\{e_{A_1}, \ldots, e_{A_h}\} \subseteq E$, and $e_{A_i} \neq e_{A_j}$ for $i \neq j$. Let Λ_n denote the set of E–labelled partitions of $\mathbb{N}_n$. Put $\alpha(\lambda) := \{A_1, \ldots, A_h\}$ and $\varepsilon(\lambda) := (e_A)_{A \in \alpha(\lambda)}$.

For $\mathbf{e} \in E^n$ with $e_i \neq e_j$ for $i \neq j$, we wish to define a Λ_n–valued process $\zeta^{\mathbf{e}}$ (the *process of coalescing Markov labelled partitions*) with the following intuitive description. The initial value of $\zeta^{\mathbf{e}}$ is the labelled partition $\{(\{1\}, e_1), \ldots, (\{n\}, e_n)\}$. As t increases, the corresponding partition $\alpha(\zeta^{\mathbf{e}}(t))$ remains unchanged and the labels $\varepsilon(\zeta^{\mathbf{e}}(t))$ evolve as a vector of independent copies of Z until immediately before two (or more) such labels coincide. At the time of such a collision, the blocks of the partition corresponding to the coincident labels are merged into one block (that is, they *coalesce*). This new block is labelled with the common element of E. The evolution then continues in the same way.

More formally, we will take $\zeta^{\mathbf{e}}$ to be defined in terms of $\mathbf{Z}^{\mathbf{e}}$ as follows (using the ingredients τ_k and Θ_k that went into the definition of $\check{\mathbf{Z}}^{\mathbf{e}}$). The corresponding partition–valued process $\xi^{\mathbf{e}} := \alpha(\zeta^{\mathbf{e}})$ is constant on intervals of the form $[\tau_k, \tau_{k+1}[$ and $\xi^{\mathbf{e}}(\tau_0) := \{\{1\}, \ldots, \{n\}\}$. Suppose for $k \geq 0$ that $\xi^{\mathbf{e}}(\tau_0), \ldots, \xi^{\mathbf{e}}(\tau_k)$ have been defined and $\tau_{k+1} < \infty$. Let $\xi^{\mathbf{e}}(\tau_{k+1})$ be the partition that is obtained by merging for each $i \in \Theta_{k+1}$ those blocks of $\xi^{\mathbf{e}}(\tau_k)$ whose least elements j are such that $Z^{e_i}(\tau_{k+1}) = Z^{e_j}(\tau_{k+1})$. Thus each block A of $\xi^{\mathbf{e}}(\tau_{k+1})$ is such that the least element $\min A$ of A is the unique element $i \in A$ for which $\check{Z}^{\mathbf{e}}_i(\tau_{k+1}) \neq \dagger$. The definition of $\zeta^{\mathbf{e}}$ is completed by labelling each block A of the partition $\xi^{\mathbf{e}}(t)$ with $\check{Z}^{\mathbf{e}}_{\min A}(t) = Z^{e_{\min A}}(t)$.

For $1 \le i \le n$, put $\boldsymbol{\gamma}^{\mathrm{e}} = (\gamma_1^{\mathrm{e}}, \dots, \gamma_n^{\mathrm{e}})$, where

$$\gamma_i^{\mathrm{e}}(t) := \min\{j : j \sim_{\xi^{\mathrm{e}}(t)} i\}, \tag{2.7}$$

and write

$$\Gamma^{\mathrm{e}}(t) := \{\gamma_i^{\mathrm{e}}(t) : 1 \le i \le n\} = \{j : \check{Z}_j^{\mathrm{e}}(t) \ne \dagger\} \tag{2.8}$$

for the set of surviving indices at time t. Note that $\Gamma^{\mathrm{e}}(\tau_k) = \Theta_k$.

3. The state-space Ξ of the stepping–stone process

We need some elementary ideas from the theory of vector measures. A good reference is [DU77]. Recall the measure space $(E, \mathcal{E}, m)$ introduced in Section 2, and let B be a Banach space with norm $\|\cdot\|$. We say that a function $\phi : E \to B$ is *simple* if $\phi = \sum_{i=1}^k x_i 1_{E_i}$ for $x_1, \dots, x_k \in B$ and $E_1, \dots, E_k \in \mathcal{E}$ for some $k \in \mathbb{N}$. We say that a function $\phi : E \to B$ is *m–measurable* if there exists a sequence $\{\phi_n\}_{n \in \mathbb{N}}$ of simple functions such that $\lim_{n\to\infty} \|\phi_n(e) - \phi(e)\| = 0$ for m-a.e. $e \in E$.

Write K for the compact, metrisable coin-tossing space $\{0,1\}^{\mathbb{N}}$ equipped with the product topology, and let $\mathcal{K}$ denote the corresponding Borel σ-field. Equivalently, $\mathcal{K}$ is the σ-field generated by the cylinder sets.

Write $M(K)$ for the Banach space of finite signed measures on $(K, \mathcal{K})$ equipped with the total variation norm $\|\cdot\|_{M(K)}$. Let $L^\infty(m, M(K))$ denote the space of (equivalence classes of) m-measurable maps $\mu : E \to M(K)$ such that $\operatorname{ess\,sup}\{\|\mu(e)\|_{M(K)} : e \in E\} < \infty$, and equip $L^\infty(m, M(K))$ with the obvious norm to make it a Banach space.

Write $C(K)$ for the Banach space of continuous functions on K equipped with the usual supremum norm $\|\cdot\|_{C(K)}$. Let $L^1(m, C(K))$, denote the Banach space of (equivalence classes of) m-measurable maps $\mu : E \to C(K)$ such that $\int m(de)\, \|\mu(e)\|_{C(K)} < \infty$, and equip $L^1(m, C(K))$ with the obvious norm to make it a Banach space.

From the discussion at the beginning of §IV.1 in [DU77] and the fact that $M(K)$ is isometric to the dual space of $C(K)$ under the pairing $(\nu, y) \mapsto \langle \nu, y\rangle = \int \nu(dk)\, y(k)$, $\nu \in M(K)$, $y \in C(K)$, we see that $L^\infty(m, M(K))$ is isometric to a closed subspace of the dual of $L^1(m, C(K))$ under the pairing $(\mu, x) \mapsto \int m(de)\, \langle \mu(e), x(e)\rangle$, $\mu \in L^\infty(m, M(K))$, $x \in L^1(m, C(K))$. Write $M_1(K)$ for the closed subset of $M(K)$ consisting of probability measures, and let Ξ denote the closed subset of $L^\infty(m, M(K))$ consisting of (equivalence classes of) maps with values in $M_1(K)$. From Corollary V.4.3 and Theorem V.5.1 of [DS58] we see that, as $L^1(m, C(K))$ is separable, Ξ equipped with the relative weak* topology is a compact, metrisable space. From now on, we always take Ξ to be equipped with the relative weak* topology.

We think of the set K as the space of possible *types* in the infinitely–many–types, continuum–sites, stepping–stone model X we will define in Section 4. As we remarked in Section 1, the type–space for infinitely–many–types models is usually taken to be $[0,1]$. However, from a modelling perspective any uncountable set is equally suitable, and, as pointed out in [Eva97], the set K is technically easier to work with. The set E is the corresponding space of *sites*. The intuitive interpretation is that $\mu \in \Xi$ describes an ensemble of populations at the various sites: $\mu(e)(L)$ is the "proportion of the population at site $e \in E$ that has a type belonging to the set $L \in \mathcal{K}$".

Remark 3.1. One can think of Ξ as a subset of the space of Radon measures on $E \times K$ by identifying $\mu \in \Xi$ with the measure that assigns mass $\int_A m(de)\,\mu(e)(B)$ to the set $A \times B$, where $A \in \mathcal{E}$ and $B \in \mathcal{K}$. The topology we are using on Ξ is not the same as the trace of the usual topology of vague convergence of Radon measures. However, the corresponding Borel σ–fields do coincide. In particular, we can think of Ξ–valued random variables as random Radon measures on $E \times K$.

For $n \in \mathbb{N}$ let $M(K^n)$ (respectively, $C(K^n)$) denote the Banach space of finite signed measures (respectively, continuous functions) on the Cartesian product K^n with the usual norm $\|\cdot\|_{M(K^n)}$ (respectively, $\|\cdot\|_{C(K^n)}$). With a slight abuse of notation, write $\langle\,\cdot\,,\,\cdot\,\rangle$ for the pairing between these two spaces.

Definition 3.2. Given $\phi \in L^1(m^{\otimes n}, C(K^n))$, define $I_n(\,\cdot\,;\phi) \in C(\Xi)$ (:= the space of continuous real–valued functions on Ξ) by

$$
\begin{aligned}
I_n(\mu;\phi) &:= \int_{E^n} m^{\otimes n}(d\mathbf{e})\,\langle \bigotimes_{i=1}^n \mu(e_i), \phi(\mathbf{e})\rangle \\
&= \int_{E^n} m^{\otimes n}(d\mathbf{e}) \int_{K^n} \bigotimes_{i=1}^n \mu(e_i)(dk_i)\phi(\mathbf{e})(\mathbf{k}), \quad \mu \in \Xi.
\end{aligned} \tag{3.1}
$$

Write I for I_1.

4. Definition of the stepping–stone process X

Theorem 4.1 below is Theorem 4.1 of [Eva97]. As discussed in Section 1, it is motivated by the characterisation of infinitely–many–types, discrete–sites stepping–stone processes via duality with systems of delayed coalescing continuous–time Markov chains (see [DGV95] and [Han90]). Recall that $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which the processes $\mathbf{Z}^{\mathbf{e}}$, $\check{\mathbf{Z}}^{\mathbf{e}}$, $\zeta^{\mathbf{e}}$, $\xi^{\mathbf{e}}$, *et cetera* are defined.

Theorem 4.1. *There exists a unique, Feller, Markov semigroup $\{Q_t\}_{t\ge 0}$ on Ξ such that for all $t \ge 0$, $\mu \in \Xi$, $\phi \in L^1(m^{\otimes n}, C(K^n))$, $n \in \mathbb{N}$, we have*

$$
\begin{aligned}
&\int Q_t(\mu, d\nu) I_n(\nu;\phi) \\
&= \int_{E^n} m^{\otimes n}(d\mathbf{e})\mathbb{P}\left[\int \bigotimes_{j\in\Gamma^{\mathbf{e}}(t)} \mu(Z_j^{\mathbf{e}}(t))(dk_j)\phi(\mathbf{e})(k_{\gamma_1^{\mathbf{e}}(t)}, \dots, k_{\gamma_n^{\mathbf{e}}(t)})\right].
\end{aligned} \tag{4.1}
$$

Consequently, there is a Hunt process, $(X, \mathbb{Q}^\mu)$, with state-space Ξ and transition semigroup $\{Q_t\}_{t\ge 0}$.

Remark 4.2. The integrand $\mathbb{P}[\cdots]$ in (4.1) should be interpreted as 0 on the $m^{\otimes n}$–null set of $\mathbf{e}$ such that $e_i = e_j$ for some pair (i,j). The integral inside the $[\ ]$ is over a Cartesian product of copies of K, with the copies indexed by the elements of $\Gamma^{\mathbf{e}}(t)$.

Remark 4.3. The following equivalent formulation of Theorem 4.1 will be useful. For $n \in \mathbb{N}$ let $\mathbf{Z}^{[n]} = (Z_1^{[n]}, \dots, Z_n^{[n]})$ be an E^n–valued process defined on a σ-finite measure space $(\Omega^{[n]}, \mathcal{F}^{[n]}, \mathbb{P}^{[n]})$, with

$$
\mathbb{P}^{[n]}\{\mathbf{Z}^{[n]} \in A\} := \int m^{\otimes n}(d\mathbf{e})\mathbb{P}\{\mathbf{Z}^{\mathbf{e}} \in A\}. \tag{4.2}
$$

Define $\check{\mathbf{Z}}^{[n]}$, $\xi^{[n]}$, $\boldsymbol{\gamma}^{[n]}$ and $\Gamma^{[n]}$ from $\mathbf{Z}^{[n]}$ in the same manner that $\check{\mathbf{Z}}^{\mathbf{e}}$, $\xi^{\mathbf{e}}$, $\boldsymbol{\gamma}^{\mathbf{e}}$ and $\Gamma^{\mathbf{e}}$ were defined from $\mathbf{Z}^{\mathbf{e}}$. The right–hand side of (4.1) is just

$$\mathbb{P}^{[n]}\left[\int \bigotimes_{j\in\Gamma^{[n]}(t)} \mu(Z_j^{[n]}(t))(dk_j)\phi(\mathbf{Z}^{[n]}(0))(k_{\gamma_1^{[n]}(t)},\ldots,k_{\gamma_n^{[n]}(t)})\right]. \tag{4.3}$$

Remark 4.4. As we noted in Remark 3.1, we can think of the process X as taking values in the space of Radon measure on $E\times K$ by identifying X_t with the random measure that assigns mass $\int_A m(de)X_t(e)(B)$ to the set $A\times B$, where $A\in\mathcal{E}$ and $B\in\mathcal{K}$. A standard monotone class argument shows that if ψ is any non-negative Borel function on $E^n\times K^n$, then

$$\begin{aligned}
&\mathbb{Q}^\mu\left[\int_{E^n} m^{\otimes n}(d\mathbf{e})\int_{K^n}\bigotimes_{i=1}^n X_t(e_i)(dk_i)\,\psi(\mathbf{e},\mathbf{k})\right]\\
&=\int_{E^n} m^{\otimes n}(d\mathbf{e})\mathbb{P}\left[\int\bigotimes_{j\in\Gamma^{\mathbf{e}}(t)}\mu(Z_j^{\mathbf{e}}(t))(dk_j)\psi(\mathbf{e},k_{\gamma_1^{\mathbf{e}}(t)},\ldots,k_{\gamma_n^{\mathbf{e}}(t)})\right]\\
&=\mathbb{P}^{[n]}\left[\int\bigotimes_{j\in\Gamma^{[n]}(t)}\mu(Z_j^{[n]}(t))(dk_j)\psi(\mathbf{Z}^{[n]}(0),k_{\gamma_1^{[n]}(t)},\ldots,k_{\gamma_n^{[n]}(t)})\right].
\end{aligned}\tag{4.4}$$

5. A particle construction for the stepping–stone model X

In this section we first construct a finite particle model in which particles move through $E\times K$, where we recall that E is our *site–space* and K is our *type–space*. The E–valued components of the particles move independently according to the dynamics of the migration process $\hat{Z}$. The particles interact only when they are located at the same site in E, and the interaction that occurs is that the type of one of the particles is replaced by the type of the other. The particle whose type "wins" is chosen at random from the two particles, with both outcomes equally likely. For our purposes here, we assume that the types are constant except for these replacement interactions, although we could allow "mutation" of the types between the replacement interactions.

Under suitable conditions on the migration process, we then pass to a high–density limit and obtain a process taking values in the space of Radon measures ρ on $E\times K$ with the property that $\rho(A\times K)=m(A)$ for $A\in\mathcal{E}$. Recalling Remark 3.1, we can think of the limit model as a Ξ–valued process, and we establish that as such it has the same finite–dimensional distributions as the continuum–sites stepping–stone process X.

Throughout this section we will work on a probability space $(\hat{\Omega},\hat{\mathcal{F}},\hat{\mathbb{P}})$ and we will assume the following hypothesis (the definition of a Hunt process is recalled in Section 2).

Assumption 5.1. The processes Z and $\hat{Z}$ are Hunt.

For completeness, we recall the following definition and some of its consequences.

Definition 5.2. Let $(S,\mathcal{S})$ be a measurable space, and let ν be a σ-finite measure on $\mathcal{S}$. Say that a map N from $\hat{\Omega}$ into the collection of measures on $(S,\mathcal{S})$ is a *Poisson random measure* with *mean measure* ν if

a) For each $A \in \mathcal{S}$, $N(A)$ is a $\{0, 1, \dots, \infty\}$–valued random variable.
b) For each $A \in \mathcal{S}$ with $\nu(A) < \infty$, the random variable $N(A)$ is Poisson distributed with parameter $\nu(A)$.
c) For $A_1, A_2, \dots \in \mathcal{S}$ disjoint, the random variables $N(A_1), N(A_2), \dots$ are independent.

Remark 5.3. Assume that ν is diffuse. Then for $x \in S$, $N(\{x\})$ must be zero or one, and so we can identify N with its support. We will write $x \in N$ if $N(\{x\}) = 1$. Note that

$$\hat{\mathbb{P}}\left[\int_E N(dx) f(x)\right] = \hat{\mathbb{P}}\left[\sum_{x \in N} f(x)\right] = \int_E \nu(dx) f(x), \ f \in L^1(\nu), \tag{5.1}$$

and more generally, for $f \in L^1(\nu^{\otimes n})$,

$$\hat{\mathbb{P}}\left[\sum_{\substack{x_1, \dots, x_n \in N \\ x_i \neq x_j,\, i \neq j}} f(x_1, \dots, x_n)\right] = \int_{E^n} \nu^{\otimes n}(d\mathbf{x}) f(x_1, \dots, x_n). \tag{5.2}$$

5.1. **Finite particle systems.** Fix a non-zero diffuse finite measure ν_0 on E and a probability kernel $\mu : E \times \mathcal{K} \to [0,1]$. Write $D_E[0, \infty[$ for the Skorohod space of càdlàg E–valued paths and let $\widehat{M}$ denote a Poisson random measure on $D_E[0, \infty[\times K$ with mean measure

$$F \times G \mapsto \int \nu_0(dz)\, \hat{P}^z(F)\, \mu(z, G) \tag{5.3}$$

(recall that $\hat{P}^z$ is the law of $\hat{Z}$ starting at $z \in E$). Thus the push–forward of $\widehat{M}$ by the map $(\zeta, k) \mapsto \zeta(0)$ (:= the value of the path ζ at time 0) is a Poisson random measure on E with mean measure ν_0. More generally, the push–forward of $\widehat{M}$ by the map $(\zeta, k) \mapsto \zeta(t)$ is a Poisson random measure on E with mean measure ν_t, where $\nu_t(H) = \int_E \nu_0(dz)\, \hat{P}_t(z, H)$. We assume that ν_t is diffuse for each $t \geq 0$. By our duality assumption, this will certainly be the case if ν_0 is absolutely continuous with respect to m.

Enumerate the atoms of $\widehat{M}$ as $(\hat{Z}_1, \kappa_1^0), \dots, (\hat{Z}_J, \kappa_J^0)$ in such a way that the conditional distribution of this collection given $J = j$ is that of j i.i.d. $D_E[0, \infty[\times K$–valued random variables with common distribution

$$F \times G \mapsto \nu_0(E)^{-1} \int \nu_0(dz)\, \hat{P}^z(F)\, \mu(z, G). \tag{5.4}$$

We wish to define a collection $\kappa_1, \dots, \kappa_J$ of K–valued processes in such way that the collection $(\hat{Z}_1, \kappa_1), \dots, (\hat{Z}_J, \kappa_J)$ has the dynamics described above: that is, we think of $\kappa_i(t)$ as the type of the particle $\hat{Z}_i$ at time t, and, after two or more such particles collide, the particles participating in the collision must be of the same type with the common type selected at random from among the types of the participating particles (with each possible outcome equally likely).

Suppose that on the probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ we also have defined for each $k \in \mathbb{N}$ a collection $\{\theta_{ik},\, i \in \mathbb{N}\}$ of i.i.d. random variables uniformly distributed on $[0, 1]$. We will implement a specific construction of the θ_{ik} below. Define $\kappa_1, \dots, \kappa_J$ and times $\hat{\tau}_0 \leq \hat{\tau}_1 \leq \dots$ (with $\hat{\tau}_k < \hat{\tau}_{k+1}$ when $\hat{\tau}_k < \infty$) as follows.

Put $\kappa_i(0) = \kappa_i^0$ and $\hat{\tau}_0 = 0$. Suppose that $\hat{\tau}_0, \dots, \hat{\tau}_k$ have already been defined and, for $1 \le i \le J$, the processes κ_i has been defined on $[0, \hat{\tau}_k]$ (or $[0, \infty[$ if $\hat{\tau}_k = \infty$). If $\hat{\tau}_k = \infty$, then the definition of κ_i, $1 \le i \le J$, is complete and just define $\hat{\tau}_\ell = \infty$ for $\ell > k$. Suppose, then, that $\hat{\tau}_k < \infty$. Put

$$\hat{\tau}_{k+1} := \inf\{t > \hat{\tau}_k : \hat{Z}_i(t) = \hat{Z}_j(t),\ \kappa_i(\hat{\tau}_k) \neq \kappa_j(\hat{\tau}_k),\ \text{some } i \neq j\}. \tag{5.5}$$

Put $\kappa_i(t) := \kappa_i(\hat{\tau}_k)$ for $\hat{\tau}_k \le t < \hat{\tau}_{k+1}$ and $1 \le i \le J$. If $\hat{\tau}_{k+1} = \infty$, then this completes the definition of κ_i, $1 \le i \le J$. Otherwise, if $\hat{\tau}_{k+1} < \infty$, then define $\kappa_i(\hat{\tau}_{k+1})$, $1 \le i \le J$, as follows. Let $\hat{\Gamma}_i(\hat{\tau}_{k+1}) := \{j : \hat{Z}_j(\hat{\tau}_{k+1}) = \hat{Z}_i(\hat{\tau}_{k+1})\}$, and let $\hat{\gamma}_i(\hat{\tau}_{k+1}) \in \hat{\Gamma}_i(\hat{\tau}_{k+1})$ satisfy $\theta_{\hat{\gamma}_i(\hat{\tau}_{k+1}),k+1} \le \theta_{j,k+1}$ for all $j \in \hat{\Gamma}_i(\hat{\tau}_{k+1})$. We set $\kappa_i(\hat{\tau}_{k+1}) = \kappa_{\hat{\gamma}_i(\hat{\tau}_{k+1})}(\hat{\tau}_k)$. Note that if $\hat{Z}_i(\hat{\tau}_{k+1}) = \hat{Z}_j(\hat{\tau}_{k+1})$, then $\hat{\Gamma}_i(\hat{\tau}_{k+1}) = \hat{\Gamma}_j(\hat{\tau}_{k+1})$, $\hat{\gamma}_i(\hat{\tau}_{k+1}) = \hat{\gamma}_j(\hat{\tau}_{k+1})$, and $\kappa_i(\hat{\tau}_{k+1}) = \kappa_j(\hat{\tau}_{k+1})$.

Our requirement that the types of colliding particles be changed to a type independently and uniformly selected from those of the participants in a collision will be met if for $k \in \mathbb{N}$, the collection $\{\theta_{ik},\, i \in \mathbb{N}\}$ is independent of $\mathcal{F}^{\hat{Z}}_{\hat{\tau}_k} \vee \mathcal{F}^{\kappa}_{\hat{\tau}_{k-1}}$, where $\{\mathcal{F}^{\hat{Z}}_t\}_{t \ge 0}$ is the filtration generated by $(\hat{Z}_1, \dots \hat{Z}_J)$ and $\{\mathcal{F}^{\kappa}_t\}_{t \ge 0}$ is the filtration generated by $(\kappa_1, \dots, \kappa_J)$. In particular, the distribution of the process $(\hat{Z}, \kappa)$ will be the same regardless of how we define the $\{\theta_{ik}\}$ as long as for each k, the conditional distribution of $\{\theta_{ik}\}$ given $\mathcal{F}^{\hat{Z}}_{\hat{\tau}_k} \vee \mathcal{F}^{\kappa}_{\hat{\tau}_{k-1}}$ is i.i.d. uniform on $[0, 1]$.

We note that $\hat{\mathbb{P}}$-a.s. there exists $\ell \in \mathbb{N}$ such that $\hat{\tau}_\ell = \infty$, so that the above construction does indeed lead to a value of $\kappa_i(t)$, $1 \le i \le J$, for all $t \ge 0$. To see this, let $R_h(t) = \{1 \le j \le J : \kappa_j(t) = \kappa_h(t)\}$ for $1 \le h \le J$ and $0 \le t < \sup_k \hat{\tau}_k$. Since there are only finitely many particles, $\hat{\mathbb{P}}\{R_h(\hat{\tau}_k) \subseteq R_h(\hat{\tau}_{k+1}) \subset \cdots \mid \mathcal{F}^{\hat{Z}}_{\hat{\tau}_k} \vee \mathcal{F}^{\kappa}_{\hat{\tau}_k}\} \ge 2^{-J} > 0$. Consequently, either there exists $\hat{\tau}_k < \infty$ such that $R_h(\hat{\tau}_k) = \{1, \dots, J\}$ or there exists a time after which $\hat{Z}_h$ does not collide with any particle having a different type.

Now we will give an explicit construction of the $\{\theta_{ik}\}$ which leads to a useful construction of our particle system $(\hat{Z}_1, \kappa_1), \dots, (\hat{Z}_J, \kappa_J)$. We assign to each particle a distinct $[0, 1]$–valued initial *level* U_i^0, $1 \le i \le J$, at time 0 and use these initial levels to define a family of $[0, 1]$–valued processes of levels $\{U_i(t)\}_{t \ge 0}$, $1 \le i \le J$. The $\{\theta_{ik}\}$ will be defined using these level processes. We will assume that the conditional distribution of $\{U_i^0\}$ given $\widehat{M}$ is that of J i.i.d. random variables uniformly distributed on $[0, 1]$. This assumption implies

$$\sum_{i=1}^{J} \delta_{(\hat{Z}_i, \kappa_i^0, U_i^0)} \tag{5.6}$$

is a Poisson random measure with mean measure

$$F \times G \times H \mapsto \int_E \nu_0(dz)\, \hat{P}^z(F)\, \mu(z, G)\, l(H), \tag{5.7}$$

where l denotes Lebesgue measure on $[0, 1]$. We define $\theta_{i1} := U_i^0$. For $0 \le t < \hat{\tau}_1$, set $U_i(t) := U_i(0) := U_i^0$. If $\hat{\tau}_1 < \infty$ and $|\hat{\Gamma}_i(\hat{\tau}_1)| = 1$, then put $U_i(\hat{\tau}_1) := U_i(\hat{\tau}_1-)$. If $\hat{\tau}_1 < \infty$ and $\hat{\Gamma}_i(\hat{\tau}_1) = \{i_1, \dots, i_n\}$, $n > 1$, then put $U_{i_l}(\hat{\tau}_1) := U_{i_{\sigma_l}}(\hat{\tau}_1-)$ where $\sigma_1, \dots, \sigma_n$ is a uniform random permutation of $1, \dots, n$ selected independently of all other quantities. Observe that $U_1(\hat{\tau}_1), \dots, U_J(\hat{\tau}_1)$ are conditionally i.i.d. uniform on $[0, 1]$ given $\mathcal{F}^{\hat{Z}}_{\hat{\tau}_2} \vee \mathcal{F}^{\kappa}_{\hat{\tau}_1}$. Define $\theta_{i2} := U_i(\hat{\tau}_1)$. Put $U_i(t) := U_i(\hat{\tau}_1)$, $\hat{\tau}_1 < t < \hat{\tau}_2$.

We continue inductively, at each time $\hat{\tau}_k < \infty$ randomly permuting the levels with indices in each $\hat{\Gamma}_i(\hat{\tau}_k)$ and defining $\theta_{i,k+1} = U_i(\hat{\tau}_k)$.

Although the level assigned to a particle may change at the time of a collision, since these changes only involve the permuation of the assignment of the levels, the set of levels is the fixed random set $\mathcal{U} := \{U_i^0\}$. Consequently, we could index the particles and their types by their corresponding level; that is, for $u \in \mathcal{U}$, define $\hat{Z}_u(t) = \hat{Z}_i(t)$ and $\kappa_u(t) = \kappa_i(t)$ if and only if $U_i(t) = u$. Since the particle assigned to level u changes only when the newly assigned particle is at the same location as the previously assigned particle, the strong Markov property implies that the processes $\{\hat{Z}_u,\, u \in \mathcal{U}\}$ are conditionally independent given $\mathcal{U}$ and $\{\hat{Z}_u(0),\, u \in \mathcal{U}\}$, and conditionally each $\hat{Z}_u$ is a Markov process with transition semigroup $\{\hat{P}_t\}$. Note that

$$\hat{\tau}_{k+1} = \inf\{t > \hat{\tau}_k : \hat{Z}_u(t) = \hat{Z}_v(t),\ \kappa_u(t-) \neq \kappa_v(t-),\ \text{some } u \neq v\}, \tag{5.8}$$

and if we define $\hat{\Gamma}_u(\hat{\tau}_k) := \{v \in \mathcal{U} : \hat{Z}_v(\hat{\tau}_k) = \hat{Z}_u(\hat{\tau}_k)\}$ and $\hat{\gamma}_u(\hat{\tau}_k) := \min(\hat{\Gamma}_u(\hat{\tau}_k))$, then $\kappa_u(\hat{\tau}_k) = \kappa_{\hat{\gamma}_u(\hat{\tau}_k)}(\hat{\tau}_k-)$. That is, if two or more particles collide, the particles involved in the collision "look down" to the lowest level particle at the same location, and change types to the type of that particle. (We note in passing that this construction is reminiscent of the "look down" construction of the Moran model in [DK96].) Consequently, if we start with a Poisson random measure on $D_E[0,\infty[\times K \times [0,1]$

$$\sum_{u\in\mathcal{U}} \delta_{(\hat{Z}_u,\kappa_u^0,u)} \tag{5.9}$$

with mean measure specified by (5.7), then a particle model

$$\Psi_t = \sum_{u\in\mathcal{U}} \delta_{(\hat{Z}_u(t),\kappa_u(t),u)} \tag{5.10}$$

is completely determined by the requirement that whenever two or more particles "collide" the types of the higher level particles involved in the collision switches to the type of the lowest level particle in the collision. This observation allows us to extend the construction to systems with infinitely many particles with mild additional assumptions.

5.2. **Particle systems with stationary location processes.** We now want to extend the construction of the previous section to arrive at a model in which the distribution of locations of particles is stationary in time, and we want to allow for the possibility of there being infinitely many particles. We emphasize that Assumption 5.1 is in force throughout this section.

Consider a Poisson random measure

$$\sum_{u\in\mathcal{U}} \delta_{(\hat{Z}_u,\kappa_u^0,u)} \tag{5.11}$$

on $D_E[0,\infty[\times K \times [0,1]$ with mean measure specified by (5.7) with $\nu_0 = m$. By the assumption that m is Radon, there exist open sets $E_1 \subseteq E_2 \subseteq \cdots$ such that $m(E_n) < \infty$ for all $n \in \mathbb{N}$ and $E = \bigcup_n E_n$ (of course, if $m(E)$ is finite we can take $E_n = E$ for all $n \in \mathbb{N}$). Put $\mathcal{U}_n = \{u \in \mathcal{U} : \hat{Z}_u(0) \in E_n\}$, and note that

$$\sum_{u\in\mathcal{U}_n} \delta_{(\hat{Z}_u,\kappa_u^0,u)} \tag{5.12}$$

is a Poisson random measure on $D_E[0,\infty[\times K \times [0,1]$ with finite mean measure specified by (5.7) with $\nu_0 = m(\cdot \cap E_n)$. As in the previous subsection, we can construct a corresponding finite particle model

$$\Psi^{[n]}(t) = \sum_{u \in \mathcal{U}_n} \delta_{(\hat{Z}_u(t), \kappa_u^{[n]}(t), u)} \tag{5.13}$$

and times $\hat{\tau}_0^{[n]} \leq \hat{\tau}_1^{[n]} \leq \ldots$. We would like to define $\Psi_t = \lim_{n\to\infty} \Psi^{[n]}(t)$; however, the type processes $\kappa_u^{[n]}$ may not converge without some additional assumptions regarding the behavior of the migration processes $\{\hat{Z}_u,\, u \in \mathcal{U}\}$.

Henceforth, we will also assume the following, which will ensure that for all $n \in N$ and $t \geq 0$ the expectation $\hat{\mathbb{P}}[\,|\{u \in \mathcal{U} : \hat{Z}_u(s) \in E_n \text{ for some } 0 \leq s \leq t\}|\,]$ is finite.

Assumption 5.4. The sequence $\{E_n\}$ of open sets can be chosen so that

$$\hat{P}^m\{\sigma_{E_n} \leq t\} < \infty, \text{ for all } n \in \mathbb{N} \text{ and } t > 0,$$

where

$$\sigma_A := \inf\{t \geq 0 : \hat{Z}(t) \in A\},\ A \in \mathcal{E}.$$

Remark 5.5. By our duality assumption, the measure m is stationary for $\hat{Z}$. It follows easily that if for $A \in \mathcal{E}$ the condition $\hat{P}^m\{\sigma_A \leq t\} < \infty$ holds for some $t > 0$, then it holds for all $t > 0$. Furthermore, the condition $\hat{P}^m\{\sigma_A \leq t\} < \infty$ for all $t > 0$ is also equivalent to $\hat{P}^m[\exp(-\lambda\sigma_A)] < \infty$ for all (equivalently, some) $\lambda > 0$. Using this equivalence, the question of whether or not Asssumption 5.4 is satisfied becomes a standard question in capacity theory. Under our duality assumption and the Assumption 5.1 that $Z, \hat{Z}$ are Hunt, Assumption 5.4 will certainly hold (with $\{E_n\}$ any increasing sequence of relatively compact open sets such that $\bigcup_n E_n = E$) if the Lusin space E is locally compact and λ-excessive functions for both semigroups $\{P_t\}$ and $\{\hat{P}_t\}$ are lower semi–continuous (see, for example, Remark 2.10 of [Get84]). In particular, Assumption 5.4 holds if E is locally compact and Z and $\hat{Z}$ have strong Feller λ–resolvent operators (see Exercise II.2.16 of [BG68]). Also, Assumption 5.4 holds when Z and $\hat{Z}$ are Lévy processes on $\mathbb{R}^d$ and m is Lebesgue measure (see Lemma II.6 of [Ber96]).

Fix $t > 0$ and $u \in \mathcal{U}_n$. Let

$$\alpha_n(u,t) := 0 \vee \sup\{0 < s < t : \hat{Z}_u(s) = \hat{Z}_v(s) \text{ some } v < u,\ v \in \mathcal{U}_n\}, \tag{5.14}$$

and let $\beta_n(u,t)$ be the corresponding value of $v \in \mathcal{U}_n$, with $\beta_n(u,t) := u$ if $\alpha_n(u,t) = 0$. Lemma 2.1 implies that $\beta_n(u,t)$ is well-defined. In general, $\alpha_n(u,t)$ will not be one of the times $\{\hat{\tau}_k^{[n]}\}$, but we will have

$$\kappa_u^{[n]}(t) = \kappa_u^{[n]}(\alpha_n(u,t)) = \kappa_{\beta_n(u,t)}^{[n]}(\alpha_n(u,t)). \tag{5.15}$$

Define $\beta_{n,u}^t(s) := u$ for $\alpha_n(u,t) < s \leq t$ and $\beta_{n,u}^t(s) := \beta_n(u,t)$ for $\alpha_n(\beta_n(u,t), \alpha_n(u,t)) < s \leq \alpha_n(u,t)$. This definition extends iteratively to determine $\beta_{n,u}^t(s)$ on the interval $0 \leq s \leq t$ with the property that

$$\kappa_u^{[n]}(t) = \kappa_{\beta_{n,u}^t(s)}^{[n]}(s), \tag{5.16}$$

so, in particular, $\kappa_u^{[n]}(t) = \kappa_{\beta_{n,u}^t(0)}^0$. Consequently, convergence of $\kappa_u^{[n]}(t)$ is equivalent to convergence of $\beta_{n,u}^t$.

For $t > 0$ and $u \in \mathcal{U}$ set

$$\alpha(u,t) := 0 \vee \sup\{0 < s < t : \hat{Z}_u(s) = \hat{Z}_v(s) \text{ some } v < u,\, v \in \mathcal{U}\}. \tag{5.17}$$

Let U be a $[0,1]$–valued random variable that is $\sigma(\mathcal{U})$–measurable and takes values in the random set $\mathcal{U}$ (that is, U is a $\sigma(\mathcal{U})$–measurable selection from $\mathcal{U}$). By the duality assumption and the Hunt hypothesis Assumption 5.1 (cf. Proposition 15.7 of [GS84]), $t - \alpha(U,t)$ has the same distribution as

$$\inf\{s > 0 : Z_U(s) = Z_v(s) \text{ some } v < U,\, v \in \mathcal{U}\} \wedge t, \tag{5.18}$$

where $\sum_{u\in\mathcal{U}} \delta_{(Z_u,u)}$ is any Poisson random measure with mean measure

$$F \times H \mapsto \int_E m(dz)\, P^z(F)\, l(H) \tag{5.19}$$

constructed from $\mathcal{U}$ using suitable further randomisation. Let $\beta(u,t)$ be the corresponding value of $v \in \mathcal{U}$ in (5.17), with $\beta(u,t) := u$ if $\alpha(u,t) = 0$. Define $\beta_u^t(s) := u$ for $\alpha(u,t) < s \le t$ and $\beta_u^t(s) := \beta(u,t)$ for $\alpha(\beta(u,t),\alpha(u,t)) < s \le \alpha(u,t)$. Extending this definition iteratively, either we determine $\beta_u^t(s)$ on the interval $0 \le s \le t$ and there are only finitely many levels in the range of β_u^t or there exists $T_u^t \ge 0$ such that $\lim_{s\downarrow T_u^t} \beta_u^t(s) = 0$. We show that this latter possibility cannot occur.

Suppose that the latter possibility does occur. As above, let U be a $\sigma(\mathcal{U})$–measurable random variable taking values in $\mathcal{U}$. Define T_U^t by analogy with T_u^t, with the convention that $T_U^t := t$ if there are only finitely many levels in the range of β_U^t. Set

$$Z_U^t(r) := \lim_{r'\downarrow r} \hat{Z}_{\beta_U^t(t-r')}(t-r'),\ 0 \le r < t - T_U^t. \tag{5.20}$$

Then by the strong Markov property, the duality assumption and Assumption 5.1, $\{Z_U^t(r),\, 0 \le r < t - T_U^t\}$ is a càdlàg Markov process with transition semigroup $\{P_r\}$. In particular, the range of this process is almost surely relatively compact and is contained in one of the E_n for n sufficiently large. Now, by Assumption 5.4, the cardinality of the set

$$\{v \in \mathcal{U} : v < U,\, \hat{Z}_v(s) \in E_n \text{ some } 0 \le s \le t\} \tag{5.21}$$

is $\hat{\mathbb{P}}$-a.s. finite for all $n \in \mathbb{N}$. Consequently, $\hat{\mathbb{P}}$-a.s. there are indeed only finitely many levels in the range of β_U^t and hence only finitely many levels in the range of β_u^t for all $u \in \mathcal{U}$. It follows that $\hat{\mathbb{P}}$-a.s. we have $\lim_{n\to\infty} \beta_{n,u}^t = \beta_u^t$ and hence

$$\lim_{n\to\infty} \kappa_u^{[n]}(t) = \kappa^0_{\beta_u^t(0)} =: \kappa_u(t) \tag{5.22}$$

for all $u \in \mathcal{U}$.

If $\beta_{u_1}^t(s) = \beta_{u_2}^t(s)$ for some $0 \le s \le t$, then $\beta_{u_1}^t(s') = \beta_{u_2}^t(s')$ for all $0 \le s' \le s$. Moreover, if we define

$$Z_u^t(r) = \lim_{r'\downarrow r} \hat{Z}_{\beta_u^t(t-r')}(t-r'),\ 0 \le r \le t, \tag{5.23}$$

for each $u \in \mathcal{U}$, then conditional on $\mathcal{U}$ each Z_u^t is a Markov process with transition semigroup $\{P_r\}$. In particular, $\{Z_u^t,\, u \in \mathcal{U}\}$ form a coalescing system of Markov processes, and for $0 \le r \le t$ the equivalence relation defined by $u \sim v$ if and only if $\beta_u^t(t-r) = \beta_v^t(t-r)$ determines a partition $\{\mathcal{U}_k^t(r)\}$ of the set of levels $\mathcal{U}$. For definiteness, assume that $\mathcal{U}$ is ordered $\mathcal{U} = \{U_1, U_2, \ldots\}$ where the U_i are $\mathcal{U}$–measurable random variables, and let $\mathcal{U}_1^t(r)$ be the equivalence class containing U_1, let $\mathcal{U}_2^t(r)$ be the equivalence class containing the U_i with smallest index not contained in $\mathcal{U}_1^t(r)$, etc. For each k, $Z_u^t(r)$ has the same value for all $u \in \mathcal{U}_k^t(r)$, which we denote by $Z_k^t(r)$. Then $((Z_1^t, \mathcal{U}_1^t), (Z_2^t, \mathcal{U}_2^t), \dots)$ forms a coalescing Markov labelled partition of $\mathcal{U}$.

Since the initial particle types $\{\kappa_u^0\}$ are conditionally independent given $\{\hat{Z}_u,\, u \in \mathcal{U}\}$ and $\mathcal{U}$, and

$$\begin{aligned}\hat{\mathbb{P}}\left[g(\kappa_{u_1}^0, \dots, \kappa_{u_n}^0) \,|\, \{\hat{Z}_u,\, u \in \mathcal{U}\}, \mathcal{U}\right] \\ = \int_{K^n} \mu(\hat{Z}_{u_1}(0), dk_1) \cdots \mu(\hat{Z}_{u_n}(0), dk_n) g(k_1, \dots, k_n),\end{aligned} \tag{5.24}$$

for $u_1, \dots, u_n \in \mathcal{U}$, we have

$$\begin{aligned}\hat{\mathbb{P}}\left[f(\hat{Z}_{u_1}(t), \kappa_{u_1}(t), \dots, \hat{Z}_{u_n}(t), \kappa_{u_n}(t)) \,\middle|\, \{\hat{Z}_u(t)\}, \mathcal{U}\right] \\ = \mathbf{H}_t f(\hat{Z}_{u_1}(t), \dots, \hat{Z}_{u_n}(t)),\end{aligned} \tag{5.25}$$

where, in the notation of Section 2,

$$\mathbf{H}_t f(e_1, \dots, e_n) := \hat{\mathbb{P}}\left[\int \bigotimes_{j \in \Gamma^{\mathbf{e}}(t)} \mu(Z_j^{\mathbf{e}}(t), dk_j) f(e_1, k_{\gamma_1^{\mathbf{e}}(t)}, \dots, e_n, k_{\gamma_n^{\mathbf{e}}(t)})\right] \tag{5.26}$$

for $e_1, \dots, e_n \in E$ with $e_i \neq e_j$, $i \neq j$.

By (5.2) and (5.25) we have

$$\begin{aligned}\hat{\mathbb{P}}\left[\sum_{\substack{u_1, \dots, u_n \in \mathcal{U} \\ u_i \neq u_j,\, i \neq j}} f(\hat{Z}_{u_1}(t), \kappa_{u_1}(t), \dots, \hat{Z}_{u_n}(t), \kappa_{u_n}(t)\right] \\ = \int_{E^n} m^{\otimes n}(d\mathbf{e})\, \mathbf{H}_t f(\mathbf{e}),\end{aligned} \tag{5.27}$$

for f a bounded measurable function on $(E \times K)^n$. This identity gives a duality in the sense of (4.4.36) of [EK86] between the discrete–particle, continuum–sites model and the corresponding coalescing Markov labelled partition process.

Write

$$\Psi_t^1 = \sum_{u \in \mathcal{U}} \delta_{(\hat{Z}_u(t), \kappa_u(t), u)} \tag{5.28}$$

and

$$X_t^1 = \sum_{u \in \mathcal{U}} \delta_{(\hat{Z}_u(t), \kappa_u(t))}. \tag{5.29}$$

Set $\mathcal{F}_t^{X^1} = \sigma(X_s^1 : s \leq t)$. The levels of the particles are independent of $\mathcal{F}_t^{X^1}$, so if $f(e, k, u)$ satisfies $\int_{E \times [0,1]} m(de)\, l(du)\, \sup_{k \in K} |f(e, k, u)| < \infty$, then

$$\hat{\mathbb{P}}\left[\int_{E \times K \times [0,1]} d\Psi_t^1 f \,\middle|\, \mathcal{F}_t^{X^1}\right] = \int_{E \times K \times [0,1]} X_t^1(de \times dk)\, l(du)\, f(e, k, u). \tag{5.30}$$

Moreover, if $f(e,k,u)$ satisfies $\int_{E\times[0,1]} m(de)\,l(du)\,(\exp(\sup_{k\in K} f(e,k,u))-1) < \infty$, then

$$\begin{aligned}\hat{\mathbb{P}}\left[\exp\left(\int_{E\times K\times[0,1]} d\Psi_t^1 f\right) \Bigm| \mathcal{F}_t^{X^1}\right] \\ = \exp\left(\int_{E\times K} X_t^1(de\times dk)\,\log\int_{[0,1]} l(du)\,\exp(f(e,k,u))\right).\end{aligned} \tag{5.31}$$

5.3. **Measure–valued, continuum–sites, stepping–stone model.** We emphasize that Assumptions 5.1 and 5.4 are still in force. Consider $\lambda > 1$. We increase the "local density" of particles in the above construction by replacing m by λm and select the levels $\mathcal{U}^\lambda$ to be i.i.d. uniform on $[0,\lambda]$ rather than $[0,1]$. Following the construction of Ψ^1 and X^1 above, define

$$\Psi_t^\lambda = \sum_{u\in\mathcal{U}^\lambda} \delta_{(\hat{Z}_u(t),\kappa_u(t),u)} \tag{5.32}$$

and

$$X_t^\lambda = \frac{1}{\lambda}\sum_{u\in\mathcal{U}^\lambda} \delta_{(\hat{Z}_u(t),\kappa_u(t))}. \tag{5.33}$$

Note that κ_u only depends on locations and types of particles at levels $v \leq u$ and we can construct $\Psi^{\lambda'}$ from Ψ^λ simultaneously for all $1 \leq \lambda' \leq \lambda$ by taking $\Psi_t^{\lambda'}$ to be the restriction of Ψ_t^λ to the particles with levels in $[0,\lambda']$; that is,

$$\Psi_t^{\lambda'} = \sum_{u\in\mathcal{U}^\lambda,\, u\leq\lambda'} \delta_{(\hat{Z}_u(t),\kappa_u(t),u)}. \tag{5.34}$$

Consequently, we may carry out the obvious construction to build Ψ_t^∞ and $\mathcal{U}^\infty$ with levels in $[0,\infty[$. The initial locations and the levels are such that

$$\Psi_0^\infty = \sum_{u\in\mathcal{U}^\infty} \delta_{(\hat{Z}_u(0),\kappa_u(0),u)} \tag{5.35}$$

is a Poisson random measure with mean measure given by

$$A\times B\times C \mapsto \int_A m(dz)\,\mu(z,B)\,l(C), \tag{5.36}$$

where l is now Lebesgue measure on $[0,\infty[$, and for $1 \leq \lambda < \infty$ each of the Ψ_t^λ can now be defined via (5.32) with $\mathcal{U}^\lambda := \{u\in\mathcal{U}^\infty : u \leq \lambda\}$. The analogue of (5.27) becomes

$$\begin{aligned}\hat{\mathbb{P}}\left[\sum_{\substack{u_1,\dots,u_n\in\mathcal{U}^\lambda\\ u_i\neq u_j,\, i\neq j}} f(\hat{Z}_{u_1}(t),\kappa_{u_1}(t),\dots,\hat{Z}_{u_n}(t),\kappa_{u_n}(t))\right] \\ = \lambda^n\int_{E^n} m^{\otimes n}(d\mathbf{e})\,\mathbf{H}_t f(\mathbf{e}),\end{aligned} \tag{5.37}$$

where $\mathbf{H}_t$ is defined as in (5.26).

Define $\mathcal{F}_t^\lambda = \sigma(X_s^\lambda, \Psi_s^\infty - \Psi_s^\lambda : s \le t)$, and note that, as in (5.30) and (5.31), we have

$$\hat{\mathbb{P}}\left[\int_{E\times K\times[0,\lambda]} d\Psi_t^\lambda f \,\Big|\, \mathcal{F}_t^\lambda\right] = \int_{E\times K\times[0,\lambda]} X_t^\lambda(de\times dk)\, l(du)\, f(e,k,u) \tag{5.38}$$

and

(5.39)

$$\begin{aligned}&\hat{\mathbb{P}}\left[\exp\left(\int_{E\times K\times[0,\lambda]} d\Psi_t^\lambda f\right) \Big|\, \mathcal{F}_t^\lambda\right] \\ &\quad = \exp\left(\int_{E\times K} X_t^\lambda(de\times dk)\,\lambda\log\left(1+\frac{1}{\lambda}\int_{[0,\lambda]} l(du)\ (\exp(f(e,k,u))-1)\right)\right).\end{aligned}$$

Suppose that $f(e,k,u) = 0$ for $u > \lambda_0$. Then for $\lambda > \lambda_0$, the random variables on the left of (5.38) and (5.39) do not depend on λ. Since for fixed $t \ge 0$ the σ–fields $\mathcal{F}_t^\lambda$ are decreasing in λ, the left sides of (5.38) and (5.39) are positive, reverse martingales, and hence converge $\hat{\mathbb{P}}$-a.s. as $\lambda \uparrow \infty$. It follows that X_t^λ converges $\hat{\mathbb{P}}$-a.s. to a random measure X_t^∞ satisfying

$$\begin{aligned}&\hat{\mathbb{P}}\left[\int_{E\times K\times[0,\infty[} d\Psi_t^\infty f \,\Big|\, \mathcal{F}_t^{X^\infty}\right] \\ &\quad = \int_{E\times K\times[0,\infty[} X_t^\infty(de\times dk)\, l(du)\, f(e,k,u),\end{aligned} \tag{5.40}$$

and

$$\begin{aligned}&\hat{\mathbb{P}}\left[\exp\left(\int_{E\times K\times[0,\infty[} d\Psi_t^\infty f\right) \Big|\, \mathcal{F}_t^{X^\infty}\right] \\ &\quad = \exp\left(\int_{E\times K\times[0,\infty[} X_t^\infty(de\times dk)\, l(du)\ (\exp f(e,k,u)-1)\right).\end{aligned} \tag{5.41}$$

In particular, by (5.41), for each $t \ge 0$, Ψ_t^∞ is a doubly stochastic Poisson process (that is, a Cox process) with random mean measure given by $X_t^\infty \otimes l$.

Dividing both sides of (5.37) by λ^n and letting $\lambda \to \infty$, we have

$$\begin{aligned}&\hat{\mathbb{P}}\left[\int_{(E\times K)^n} X_t^\infty(de_1\times dk_1)\cdots X_t^\infty(de_n\times dk_n)\, f(e_1,k_1,\ldots,e_n,k_n)\right] \\ &\quad = \int_{E^n} m^{\otimes n}(d\mathbf{e})\, \mathbf{H}_t f(\mathbf{e}).\end{aligned} \tag{5.42}$$

Note also that

$$X_t^\infty(\cdot\times K) = m. \tag{5.43}$$

By Remark 3.1 we can regard the measure $m(de)\mu(e,dk)$ as an element of Ξ (which we will also denote by μ) and the random measure X_t^∞ as a Ξ–valued random variable. By Theorem 4.1, X_t^∞ has the same law as X_t under $\mathbb{Q}^\mu$ for each $t \ge 0$. In fact, it is not difficult to show that $(X_t^\infty, t \ge 0)$ is a Markov process with the same finite–dimensional distributions as X under $\mathbb{Q}^\mu$. We stress, however, that we have only constructed X_t^∞ as an almost sure limit for each fixed $t \ge 0$ rather than as an almost sure limit in some space of càdlàg paths.

6. Dissimilarity for the stepping–stone process X

Suppose in this section that the reference measure m is finite. Without loss of generality, we can take m to be a probability measure.

Definition 6.1. Consider $\nu \in \Xi$. For $n = 2, 3, \ldots$ define the n^{th}*–order dissimilarity* of ν to be the quantity

$$D_n(\nu) := \int m^{\otimes n}(d\mathbf{e}) \bigotimes_{i=1}^{n} \nu(e_i)\left(\{\mathbf{k} \in K^n : k_j \neq k_\ell, \text{ for all } j \neq \ell\}\right). \tag{6.1}$$

Note that $1 \geq D_2(\nu) \geq D_3(\nu) \geq \cdots \geq 0$. Write $\breve{D}(\nu) := \sup\{n : D_n(\nu) > 0\}$ for the *maximal dissimilarity* of ν, where we set $\sup \emptyset = 1$.

As we remarked in the Introduction, it is possible, by exactly the same argument used in Proposition 5.1 of [Eva97], to show that if Z (and hence also $\hat{Z}$) is a symmetric α–stable process on the circle $\mathbb{T}$ with index $1 < \alpha \leq 2$, then for fixed $t > 0$ there $\mathbb{Q}^\mu$–a.s. exists a random countable subset $\{k_1, k_2, \ldots\}$ of the type–space K such that for Lebesgue almost all $e \in \mathbb{T}$ the probability measure $X_t(e)$ is a point mass at one of the k_i. Indeed, under suitable hypotheses a similar argument should extend to certain other processes for which points are regular. It is clear that if X_t also has finite maximal dissimilarity $\mathbb{Q}^\mu$–a.s., then the set $\{k_1, k_2, \ldots\}$ is, in fact, finite $\mathbb{Q}^\mu$–a.s. Theorem 6.4 below provides a sufficient condition for the maximal dissimilarity $\breve{D}(X_t)$ to be finite. We are able to verify this condition when Z (and hence also $\hat{Z}$) is Brownian motion on $\mathbb{T}$ (see Corollary 9.3 and the beginning of the proof of Theorem 10.2). We suspect that the condition is also true for Lévy processes on $\mathbb{T}$ for which points are not essentially polar (see [EP98] for an indication that this might be so).

Definition 6.2. Observe that if $n' > n$, then

$$\left((Z_1^{[n']}, \ldots, Z_n^{[n']}), (\gamma_1^{[n']}, \ldots, \gamma_n^{[n']}), \xi_{|\mathbb{N}_n}^{[n']}\right) \tag{6.2}$$

has the same distribution as $(\mathbf{Z}^{[n]}, \boldsymbol{\gamma}^{[n]}, \xi^{[n]})$, where we write $\xi_{|\mathbb{N}_n}^{[n']}(t)$ for the restriction of the partition $\xi^{[n']}(t)$ to $\mathbb{N}_n$. Consequently, on some probability space $(\Omega^{[\infty]}, \mathcal{F}^{[\infty]}, \mathbb{P}^{[\infty]})$ there is an E^∞–valued process $\mathbf{Z}$, an $\mathbb{N}^\infty$–valued process $\boldsymbol{\gamma}$, and a process ξ taking values in the space of partitions of $\mathbb{N}$ such that, in an obvious notation, $((Z_1, \ldots, Z_n), (\gamma_1, \ldots, \gamma_n), \xi_{|\mathbb{N}_n})$ has the same distribution as $(\mathbf{Z}^{[n]}, \boldsymbol{\gamma}^{[n]}, \xi^{[n]})$.

Remark 6.3. Recall the definition of $\mathbf{Z}^{[n]}$, $\breve{\mathbf{Z}}^{[n]}$ and $\xi^{[n]}$ from Remark 4.3. Let $\breve{\mathbf{Z}}^{[n]\ddagger}$ and $\xi^{[n]\ddagger}$ be defined from $\mathbf{Z}^{[n]}$ in a similar manner to $\breve{\mathbf{Z}}^{[n]}$ and $\xi^{[n]}$, with the difference that when two coordinate processes of $\mathbf{Z}^{[n]}$ collide, rather than the one with the higher index being killed, a colliding particle is killed at random independently of the past (with both possibilities equally likely). It is immediate from the strong Markov property that $(\mathbf{Z}^{[n]}, \xi^{[n]\ddagger})$ has the same distribution as $(\mathbf{Z}^{[n]}, \xi^{[n]})$ for all $n \in \mathbb{N}$. Consider $t \geq 0$ and a bijection $\beta : \mathbb{N} \to \mathbb{N}$ and define $\xi^{(\beta)}(t)$, a random partition of $\mathbb{N}$, by $i \sim_{\xi^{(\beta)}(t)} j$ if and only if $\beta^{-1}(i) \sim_{\xi(t)} \beta^{-1}(j)$. Then $((Z_{\beta(i)}(0))_{i \in \mathbb{N}}, \xi^{(\beta)})$ has the same distribution as $((Z_i(0))_{i \in \mathbb{N}}, \xi)$. In particular, for each $t \geq 0$ the random partition $\xi(t)$ is *exchangeable* in the sense of Kingman's definition of exchangeable random partitions (see Section 11 of [Ald85]).

Theorem 6.4. *For any $\mu \in \Xi$ and $t \geq 0$, the maximal dissimilarity $\breve{D}(X_t)$ under $\mathbb{Q}^\mu$ is stochastically dominated by the number of blocks in the partition $\xi(t)$. In particular, if for some $t > 0$ the partition $\xi(t)$ has finitely many blocks $\mathbb{P}^{[\infty]}$–a.s., then $\breve{D}(X_t) < \infty$, $\mathbb{Q}^\mu$-a.s.*

Proof. Fix a diffuse probability measure κ on K (for example, κ could be fair coin–tossing measure). Given another probability measure ρ on K, let $\rho^\rightarrow$ denote the push–forward of the measure $\rho \otimes \kappa$ on $K \times K$ by the mapping $(k, h) \mapsto (k_1, h_1, k_2, h_2, k_3, \ldots)$ from $K \times K$ into K. Let $\rho^\leftarrow$ denote the push–forward of ρ by the mapping $k \mapsto (k_1, k_3, k_5, \ldots)$ from K into K. Thus the operations $\rho \mapsto \rho^\rightarrow$ and $\rho \mapsto \rho^\leftarrow$ are one–sided inverses of each other: we have $(\rho^\rightarrow)^\leftarrow = \rho$. Given $\mu \in \Xi$, define $\mu^\rightarrow, \mu^\leftarrow \in \Xi$ by $\mu^\rightarrow(e) := \mu(e)^\rightarrow$ and $\mu^\leftarrow(e) := \mu(e)^\leftarrow$. Of course, the operations $\mu \mapsto \mu^\rightarrow$ and $\mu \mapsto \mu^\leftarrow$ are also one–sided inverses of each other. Note for any $\mu \in \Xi$ that $\mu^\rightarrow(e)$ is diffuse for all $e \in E$.

Fix $t \geq 0$. It is straighforward to check from the definition in Theorem 4.1 that the distribution of $X_t^\leftarrow$ under $\mathbb{Q}^{\mu^\rightarrow}$ coincides with the distribution of X_t under $\mathbb{Q}^\mu$. (This is, of course, what we expect from the stepping–stone model interpretation: a model that keeps track of the types for one trait should look the same as a model that keeps track of the types for two traits if we don't look at one of the traits.) Clearly, $D_n(\nu^\leftarrow) \leq D_n(\nu)$ for any n and $\nu \in \Xi$, so $\breve{D}(\nu^\leftarrow) \leq \breve{D}(\nu)$. Consequently, $\breve{D}(X_t)$ under $\mathbb{Q}^\mu$ is stochastically dominated by $\breve{D}(X_t)$ under $\mathbb{Q}^{\mu^\rightarrow}$.

We can use Remark 4.4 to compute multivariate moments of the form

$$\mathbb{Q}^\nu[\{D_{n_1}(X_t)\}^{a_1} \ldots \{D_{n_\ell}(X_t)\}^{a_\ell}],\ n_i \in \{2, 3, \ldots\},\ a_i \in \mathbb{N},\ 1 \leq i \leq \ell,\ \ell \in \mathbb{N},$$

and discover that they are independent of ν within the class of $\nu \in \Xi$ with the property that $\nu(e)$ is diffuse for all $e \in E$. Because $0 \leq D_k(X_t) \leq 1$ for all $k \geq 2$, the multivariate moment problem for each of the vectors $(D_{n_1}(X_t), \ldots, D_{n_\ell}(X_t))$ is well–posed and hence the joint distribution of $(D_2(X_t), D_3(X_t), \ldots)$ under $\mathbb{Q}^\nu$ is the same for all such ν. Consequently, the distribution of $\breve{D}(X_t)$ under $\mathbb{Q}^\nu$ is also the same for all such ν. In particular, if $\lambda \in \Xi$ is defined by $\lambda(e) := \kappa$ for all $e \in E$, then the distributions of $\breve{D}(X_t)$ under $\mathbb{Q}^{\mu^\rightarrow}$ and $\mathbb{Q}^\lambda$ are the same.

Putting the above observations together, we see that it suffices to show that $\breve{D}(X_t)$ under $\mathbb{Q}^\lambda$ is stochastically dominated by the number of blocks of $\xi(t)$.

Let $(\tilde{L}_i)_{i \in \mathbb{N}}$ be an i.i.d. sequence of K–valued random variables (which we suppose are also defined on $(\Omega^{[\infty]}, \mathcal{F}^{[\infty]}, \mathcal{P}^{[\infty]})$) that is independent of $\mathbf{Z}$ with $\tilde{L}_i$ having distribution κ. Put $L_i = \tilde{L}_{\gamma_i(t)}$, so that $L_i = L_j$ if and only if $i \sim_{\xi(t)} j$, $\mathbb{P}^{[\infty]}$–a.s. It follows from Remark 6.3 that the sequence $((Z_i(0), L_i))_{i \in \mathbb{N}}$ of $E \times K$–valued random variables is exchangeable.

Let Δ_i denote the point mass at $(Z_i(0), L_i)$. By an extension of the standard reverse martingale proof of de Finetti's theorem, as $n \to \infty$ the sequence of random probability measures $Y_n := n^{-1} \sum_{i=1}^n \Delta_i$ converges $\mathbb{P}^{[\infty]}$–a.s. in the weak topology to a random probability measure Y on $E \times K$. Moreover, if we let $\mathcal{I}$ denote the permutation invariant σ-field corresponding to $((Z_i(0), L_i))_{i \in \mathbb{N}}$ (that is, $\mathcal{I} := \bigcap_n \sigma\{Y_n, Y_{n+1}, \ldots\}$), then we have

$$\mathbb{P}^{[\infty]}[\phi((Z_1(0), L_1), \ldots, (Z_n(0), L_n)) \mid \mathcal{I}] = \int dY^{\otimes n} \phi \tag{6.3}$$

for any bounded Borel function ϕ on $(E \times K)^n$. (See the proof of Theorem 2.4 of [DK96] for the details of this sort of argument.)

We claim that Y has the same distribution as X_t under $\mathbb{Q}^\lambda$ (recall from Remark 3.1 that we can identify $\nu \in \Xi$ with the probability measure $m(de)\nu(e)(dk)$ on $E \times K$ and that Ξ–valued random variables become random probability measures on $E \times K$ when thought of in this way). If ϕ is a bounded Borel function on $(E \times K)^n$ for some $n \in \mathbb{N}$, then, by (6.3),

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\left[\int dY^{\otimes n}\phi\right] \\
&\quad = \mathbb{P}^{[\infty]}\left[\phi\left((Z_1(0), L_1), \ldots, (Z_n(0), L_n)\right)\right] \\
&\quad = \mathbb{P}^{[n]}\left[\int \bigotimes_{j \in \Gamma^{[n]}(t)} \kappa(dk_j)\phi\left((Z_1^{[n]}(0), k_{\gamma_1^{[n]}(t)}), \ldots, (Z_n^{[n]}(0), k_{\gamma_n^{[n]}(t)})\right)\right] \\
&\quad = \mathbb{P}^{[n]}\left[\int \bigotimes_{j \in \Gamma^{[n]}(t)} \lambda(Z_j^{[n]}(t))(dk_j)\phi\left((Z_1^{[n]}(0), k_{\gamma_1^{[n]}(t)}), \ldots, (Z_n^{[n]}(0), k_{\gamma_n^{[n]}(t)})\right)\right].
\end{aligned}$$

Comparing this with the equivalent definition of $(X, \mathbb{Q}^\mu)$ in Remark 4.3 shows that Y does indeed have the same distribution as X_t under $\mathbb{Q}^\lambda$.

Finally, by (6.3) we have

$$\begin{aligned}
D_n(Y) &= \mathbb{P}^{[\infty]}\{L_i \neq L_j,\ 1 \le i < j \le n \mid \mathcal{I}\} \\
&\le \mathbb{P}^{[\infty]}\{\exists \ell_1, \ldots, \ell_n : L_{\ell_i} \neq L_{\ell_j},\ 1 \le i < j \le n \mid \mathcal{I}\} \\
&= \mathbf{1}\{\exists \ell_1, \ldots, \ell_n : L_{\ell_i} \neq L_{\ell_j},\ 1 \le i < j \le n\} \\
&= \mathbf{1}\{\xi(t) \text{ has at least } n \text{ blocks}\}.
\end{aligned} \tag{6.4}$$

It is thus certainly the case that $\check{D}(X_t)$ under $\mathbb{Q}^\lambda$ is stochastically dominated by the number of blocks of $\xi(t)$. □

Remark 6.5. If Z and $\hat{Z}$ are both Hunt processes (that is, if Assumption 5.1 holds), then the particle representation of Section 5 can be used to give a somewhat more direct proof of Theorem 6.4 (note that Assumption 5.4 holds because m is a probability measure). We can sketch the proof as follows. The set of levels $\mathcal{U}^\infty$ in the construction of Section 5 is the set of points of a Poisson random measure on $[0, \infty[$ with Lebesgue intensity and hence $\mathcal{U}^\infty$ is discrete. The dissimilarity $D_n(X_t^\infty)$ is just the conditional probability (conditioning on X_t^∞) that the particles with the n lowest levels are all of different types. The argument that lead to (5.27) establishes that the total number of types exhibited by all particles is just the number of blocks in the corresponding coalescing Markov labelled partition.

Remark 6.6. In the spirit of the previous remark, it is easy to see that $\check{D}(X_t)$ is almost surely finite for all $t > 0$ when Z (and hence also $\hat{Z}$) is Brownian motion on the circle $\mathbb{T}$ and m is normalised Lebesgue measure. Once again, we just sketch the argument as a more quantitative result will follow from Corollary 9.3 below. Almost surely, there will exist two particles, say with the m^{th} and n^{th} lowest levels, $m < n$, such that by time t these two particles have collided and after the collision the first particle moved around the circle clockwise while the second particle moves around anti–clockwise until they collided again. The total number of types exhibited by all particles at time t is then at most $n-1$.

7. Sample path continuity of the stepping–stone process X

Our aim in this Section is to present a sufficient condition for X to have continuous sample paths (Theorem 7.2) and use it to establish that if the migration Markov process is a Lévy process or a "nice" diffusion, then X has continuous sample paths (Corollary 7.3, Corollary 7.4 and Remark 7.5). The proof of Theorem 7.2 is postponed to the next section. We emphasize that we are no longer assuming that the reference measure m is finite.

Definition 7.1. For $\mathbf{e} = (e_1, e_2) \in E^2$ with $e_1 \neq e_2$, let $T^{\mathbf{e}} := \inf\{t \geq 0 : Z^{e_1}(t) = Z^{e_2}(t)\}$ denote the first time that Z^{e_1} and Z^{e_2} collide.

Theorem 7.2. *Suppose there exists $\varepsilon > 0$ such that for all non-negative $\psi \in L^1(m) \cap L^\infty(m)$,*

$$\limsup_{t \downarrow 0} t^{-\varepsilon} \int m^{\otimes 2}(d\mathbf{e}) \psi^{\otimes 2}(\mathbf{e}) \mathbb{P}\{T^{\mathbf{e}} \leq t\} < \infty.$$

Then X has continuous sample paths $\mathbb{Q}^\mu$-a.s. for all $\mu \in \Xi$.

Corollary 7.3. *Suppose that Z is a Lévy process on $\mathbb{R}^d$ or the torus $\mathbb{T}^d$ for some $d \in \mathbb{N}$, and m is Lebesgue measure. Then X has continuous sample paths $\mathbb{Q}^\mu$-a.s. for all $\mu \in \Xi$.*

Proof. For $d \geq 2$ we have that $T^{\mathbf{e}} = \infty$, $\mathbb{P}$–a.s. for $m^{\otimes 2}$–a.e. $\mathbf{e}$, and so Theorem 7.2 certainly gives the result. In fact, it follows from the remarks at the beginning of Section 5 in [Eva97] that X evolves deterministically and continuously in this case.

Now consider the case where Z is $\mathbb{R}$-valued. The $\mathbb{T}$-valued case is similar and is left to the reader.

Write $(\bar{Z}, \bar{P}^z)$ for the Lévy process that is the symmetrisation of Z. That is, the distribution of $\bar{Z}$ starting at 0 is the same as that of $Z' - Z''$, where Z', Z'' are two independent copies of Z both started at 0. Put

$$\bar{T}^0 := \inf\{t \geq 0 : \bar{Z}(t) = 0\}. \tag{7.1}$$

Then for non-negative $\psi \in L^1(m) \cap L^\infty(m)$,

$$\begin{aligned} \int m^{\otimes 2}(d\mathbf{e}) \psi^{\otimes 2}(\mathbf{e}) P\{T^{\mathbf{e}} \leq t\} &= \int m^{\otimes 2}(d\mathbf{e}) \psi^{\otimes 2}(\mathbf{e}) \bar{P}^{e_1 - e_2}\{\bar{T}^0 \leq t\} \\ &= \int m(dx) \bar{\psi}(x) \bar{P}^x\{\bar{T}^0 \leq t\}, \end{aligned} \tag{7.2}$$

where $\bar{\psi}(x) := \int m(dy)\psi(x+y)\psi(y) \in L^1(m) \cap L^\infty(m)$ and we are, of course, using the shift invariance of m.

For $\alpha > 0$ write $\bar{C}^\alpha$, $\bar{U}^\alpha$ and $\bar{e}^\alpha$ for the α-capacity, α-resolvent and α-energy corresponding to $\bar{Z}$ (see Sections I.2, II.3 and II.4 of [Ber96] for definitions). Because $\bar{Z}$ is symmetric, these coincide with the corresponding dual objects. Write Ψ for the characteristic exponent of $\bar{Z}$ (see Section I.1 of [Ber96]). Note that Ψ is real–valued and non-negative.

Using the convention $\frac{1}{\infty} = 0$, we have from Theorems II.7 and II.13 of [Ber96] that

$$\int m(dx)\bar{\psi}(x)\bar{P}^x[\exp(-\alpha\bar{T}^0)] = \bar{C}^\alpha(\{0\})\bar{U}^\alpha\bar{\psi}(0) \le \frac{\bar{U}^\alpha\bar{\psi}(0)}{\bar{e}^\alpha(\{0\})} \le \frac{\|\bar{\psi}\|_\infty}{\alpha\bar{e}^\alpha(\{0\})}. \tag{7.3}$$

By Proposition I.2 of [Ber96],

$$\Psi(z) \le cz^2, \quad |z| \ge 1, \tag{7.4}$$

for a suitable constant c, and so, for $\alpha \ge 1$,

$$\alpha\bar{e}^\alpha(\{0\}) = \frac{\alpha}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\alpha+\Psi(z)}\,dz \ge \frac{\alpha}{2\pi}\int_{|z|\ge 1}\frac{1}{\alpha+cz^2}\,dz \ge c'\alpha^{\frac{1}{2}} \tag{7.5}$$

for a suitable constant c'.

Use the inequality $\mathbf{1}_{[0,t]}(x) \le e^{-\alpha x} + 1 - e^{-\alpha t} \le e^{-\alpha x} + \alpha t$ and take $\alpha = t^{-\frac{2}{3}}$ to get, for $t \le 1$,

$$\int m(dx)\bar{\psi}(x)\bar{P}^x\{\bar{T}^0 \le t\} \le c''\|\bar{\psi}\|_\infty\alpha^{-\frac{1}{2}} + \|\bar{\psi}\|_1\alpha t \le c^*(\|\bar{\psi}\|_1 + \|\bar{\psi}\|_\infty)t^{\frac{1}{3}}, \tag{7.6}$$

for suitable constants c'' and c^*. Now apply Theorem 7.2 with $\varepsilon = 1/3$. $\square$

Corollary 7.4. *Let d be a metric inducing the topology of the Lusin space E. Write $B(x,r) := \{y \in E : d(x,y) \le r\}$ for the closed ball of radius $r > 0$ centred at $x \in E$ and $S^r := \inf\{t \ge 0 : Z(t) \notin B(Z(0),r)\}$ for the time taken by Z to first travel distance r from its starting point. Suppose that there are constants $\alpha, \beta, \gamma > 0$ such that*

$$\limsup_{r\downarrow 0} r^{-\alpha}\sup_{x\in E} m(B(x,r)) < \infty$$

and

$$\limsup_{r\downarrow 0} r^{-\gamma}\sup_{x\in E} P^x\{S^{r^\beta} \le r\} < \infty.$$

Then X has continuous sample paths $\mathbb{Q}^\mu$-a.s. for all $\mu \in \Xi$.

Proof. For non-negative $\psi \in L^1(m)\cap L^\infty(m)$ and $\delta > 0$ we have

$$\begin{aligned}
&\int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbb{P}\{T^{\mathbf{e}} \le t\} \\
&= \int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbf{1}\{d(e_1,e_2) \le \delta\}\mathbb{P}\{T^{\mathbf{e}} \le t\} \\
&\quad + \int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbf{1}\{d(e_1,e_2) > \delta\}\mathbb{P}\{T^{\mathbf{e}} \le t\} \\
&\le \|\psi\|_1\|\psi\|_\infty\sup_{x\in E} m(B(x,\delta)) + 2\|\psi\|_1^2\sup_{x\in E} P^x\{S^{\delta/2} \le t\}.
\end{aligned} \tag{7.7}$$

Take $\delta = t^\beta$ to get that the hypothesis of Theorem 7.2 holds with $\varepsilon = (\alpha\beta)\wedge\gamma$. $\square$

Remark 7.5. The above result can be applied to the case where Z is a regular diffusion on $\mathbb{R}$ in natural scale. In this case m is the speed measure and $\hat{Z} = Z$. If $m(dx) = a(x)dx$ with a bounded away from 0 and ∞ and d is the usual Euclidean metric on $\mathbb{R}$, then it is not difficult to see that the conditions of the corollary hold for $\alpha = 1$, $\beta < 1/2$, and any $\gamma > 0$. We leave the details to the reader.

8. Proof of Theorem 7.2

The proof of Theorem 7.2 will involve checking *Kolmogorov's criterion* for the sample path continuity of real–valued processes of the form $(I(X_t;\phi))_{t\geq 0}$ for suitable $\phi \in C(K)$, where $I(\cdot;\cdot)$ is defined in Definition 3.2. The proof will be via several lemmas.

Remark 8.1. In performing the necessary moment computations for Kolmogorov's criterion we will need to consider the various orders in which particles can coalesce in the coalescing system that define these moments and estimate the contribution of each possible sequence of collisions. We will repeatedly use the fact that if, for fixed $i \neq j$, we "swap" $Z^{e_i}(t)$ and $Z^{e_j}(t)$ immediately after a stopping time S for $\mathbf{Z}^{\mathbf{e}}$ at which $Z^{e_i}(S) = Z^{e_j}(S)$ to form a new process $\tilde{\mathbf{Z}}^{\mathbf{e}}$, then $\tilde{\mathbf{Z}}^{\mathbf{e}}$ has the same distribution as $\mathbf{Z}^{\mathbf{e}}$. More precisely, if we define

$$\tilde{Z}_i^{\mathbf{e}}(t) := \begin{cases} Z_i^{\mathbf{e}}(t), & \text{for } t \leq S, \\ Z_j^{\mathbf{e}}(t), & \text{for } t > S, \end{cases} \tag{8.1}$$

$$\tilde{Z}_j^{\mathbf{e}}(t) := \begin{cases} Z_j^{\mathbf{e}}(t), & \text{for } t \leq S, \\ Z_i^{\mathbf{e}}(t), & \text{for } t > S, \end{cases} \tag{8.2}$$

and

$$\tilde{Z}_h^{\mathbf{e}} := Z_h^{\mathbf{e}},\ h \notin \{i,j\}, \tag{8.3}$$

then, by the strong Markov property, $\tilde{\mathbf{Z}}^{\mathbf{e}}$ has the same distribution as $\mathbf{Z}^{\mathbf{e}}$.

Definition 8.2. For $n', n'' \in \mathbb{N}$ consider $\mathbf{e}' \in E^{n'}$ and $\mathbf{e}'' \in E^{n''}$ with $e'_1, \ldots, e'_{n'}, e''_1, \ldots e''_{n''}$ distinct. Define a process $\Phi^{\mathbf{e}'|\mathbf{e}''}$ taking values in the collection of finite sequences of two element subsets of $\{e'_1, \ldots, e'_{n'}, e''_1, \ldots e''_{n''}\}$ and stopping times $0 = T_0^{\mathbf{e}'|\mathbf{e}''} \leq T_1^{\mathbf{e}'|\mathbf{e}''} \leq \ldots$ as follows. For $e'_i \in \{e'_1, \ldots, e'_{n'}\}$, write

$$\begin{aligned} S_{e'_i}^{\mathbf{e}'|\mathbf{e}''} := \inf\{t \geq 0 : & Z^{e'_i}(t) = Z^{e'_j}(t) \,\text{for some}\ j \neq i \\ & \text{or}\ Z^{e'_i}(t) = Z^{e''_k}(t) \,\text{for some}\ k \text{ such that } \check{Z}_k^{\mathbf{e}''}(s) \neq \dagger \,\text{for all}\, s < t\}. \end{aligned}$$

For $e''_i \in \{e''_1, \ldots, e''_{n''}\}$ write

$$\begin{aligned} S_{e''_i}^{\mathbf{e}'|\mathbf{e}''} := \inf\{t \geq 0 : & Z^{e''_i}(t) = Z^{e'_j}(t) \,\text{for some}\ j \\ & \text{or}\ Z^{e''_i}(t) = Z^{e''_k}(t) \,\text{for some}\ k \neq i \text{ such that } \check{Z}_k^{\mathbf{e}''}(s) \neq \dagger \,\text{for all}\, s < t\}. \end{aligned}$$

Loosely put, we are thinking of the particles starting at coordinates of $\mathbf{e}'$ as evolving freely without coalescence whereas the particles starting at coordinates of $\mathbf{e}''$ are undergoing coalescence among themselves. Moreover, $S_e^{\mathbf{e}'|\mathbf{e}''}$ (if it is finite) is the first time that the particle starting at e (where e is either a coordinate of $\mathbf{e}'$ or $\mathbf{e}''$) collides with a "living" particle starting at one of the other coordinates.

Let $R^{\mathbf{e}'|\mathbf{e}''} < n' + n''$ denote the cardinality of the random set of time points

$$\left\{S_e^{\mathbf{e}'|\mathbf{e}''} : e \in \{e'_1, \ldots, e'_{n'}, e''_1, \ldots e''_{n''}\} \text{ and } S_e^{\mathbf{e}'|\mathbf{e}''} < \infty\right\} \tag{8.4}$$

and, if $R^{\mathbf{e}'|\mathbf{e}''} > 0$, write $T_1^{\mathbf{e}'|\mathbf{e}''} < \ldots < T_{R^{\mathbf{e}'|\mathbf{e}''}}^{\mathbf{e}'|\mathbf{e}''}$ for an ordered listing of this set. Put $T_0^{\mathbf{e}'|\mathbf{e}''} := 0$ and $T_\ell^{\mathbf{e}'|\mathbf{e}''} := \infty$ for $\ell > R^{\mathbf{e}'|\mathbf{e}''}$.

Set $\Phi^{\mathbf{e}'|\mathbf{e}''}(T_0^{\mathbf{e}'|\mathbf{e}''}) := \emptyset$. For $1 \le k \le R^{\mathbf{e}'|\mathbf{e}''}$ write

$$\{x_k, y_k\} \subseteq \{e'_1, \dots, e'_{n'}, e''_1, \dots, e''_{n''}\}, \tag{8.5}$$

for the $\mathbb{P}$–a.s. unique unordered pair such that $Z^{x_k}(T_k^{\mathbf{e}'|\mathbf{e}''}) = Z^{y_k}(T_k^{\mathbf{e}'|\mathbf{e}''})$. By definition, at most one of x_k and y_k belong to $\{x_1, y_1, \dots, x_{k-1}, y_{k-1}\}$. Put $\Phi^{\mathbf{e}'|\mathbf{e}''}(T_k^{\mathbf{e}'|\mathbf{e}''}) := (\{x_1, y_1\}, \dots, \{x_k, y_k\})$. Complete the definition of $\Phi^{\mathbf{e}'|\mathbf{e}''}$ by setting $\Phi^{\mathbf{e}'|\mathbf{e}''}(t) := \Phi^{\mathbf{e}'|\mathbf{e}''}(T_\ell^{\mathbf{e}'|\mathbf{e}''})$, where $\ell \ge 0$ is such that $T_\ell^{\mathbf{e}'|\mathbf{e}''} \le t < T_{\ell+1}^{\mathbf{e}'|\mathbf{e}''}$.

Definition 8.3. Given $\mathbf{e} \in E^n$, $n \in \mathbb{N}$, with $e_1, \dots, e_n$ distinct, define a process $\Phi^{\mathbf{e}}$ taking values in the collection of finite sequences of two element subsets of $\{e_1, \dots, e_n\}$ and stopping times $0 = T_0^{\mathbf{e}} \le T_1^{\mathbf{e}} \le \dots$ by (with a slight abuse) re-using the definitions of $\Phi^{\mathbf{e}'|\mathbf{e}''}$ and $T_k^{\mathbf{e}'|\mathbf{e}''}$ with $\mathbf{e}' = \mathbf{e}$ and $\mathbf{e}''$ the null vector. That is, all particles evolve freely with none of them killed due to coalescence. Note that if $n = 2$, then $T_1^{\mathbf{e}} = T^{\mathbf{e}}$, where $T^{\mathbf{e}}$ is the first collision time from Definition 7.1.

Notation 8.4. Given $\mathbf{e}' \in E^{n'}$ and $\mathbf{e}'' \in E^{n''}$, write $\mathbf{e}' : \mathbf{e}''$ for the *concatenation* of these two vectors. That is, $\mathbf{e}' : \mathbf{e}' := (e'_1, \dots, e'_{n'}, e''_1, \dots e''_{n''}) \in E^{n'+n''}$.

Notation 8.5. For $x \in \mathbb{R}$ write $\lfloor x \rfloor$ for the greatest integer less than or equal to x.

Lemma 8.6. *For non-negative $\psi \in L^1(m) \cap L^\infty(m)$, $t \ge 0$, and $q, n', n'' \in \mathbb{N}$ we have*

$$\begin{aligned} &\int m^{\otimes n'} \otimes m^{\otimes n''}(d\mathbf{e}' \otimes d\mathbf{e}'')\psi^{\otimes n'}(\mathbf{e}')\psi^{\otimes n''}(\mathbf{e}'')\mathbb{P}\{T_q^{\mathbf{e}'|\mathbf{e}''} \le t\} \\ &\le \int m^{\otimes n'} \otimes m^{\otimes n''}(d\mathbf{e}' \otimes d\mathbf{e}'')\psi^{\otimes n'}(\mathbf{e}')\psi^{\otimes n''}(\mathbf{e}'')\mathbb{P}\{T_q^{\mathbf{e}':\mathbf{e}''} \le t\} \\ &\le c(n'+n'', q, \psi)\left(\int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbb{P}\{T^{\mathbf{e}} \le t\}\right)^{\lfloor \frac{q}{3} \rfloor} \end{aligned} \tag{8.6}$$

for some constant $c(n'+n'', q, \psi)$ that depends only on $n'+n''$, q and ψ.

Proof. By definition, $S_e^{\mathbf{e}':\mathbf{e}''} \le S_e^{\mathbf{e}'|\mathbf{e}''}$ for any $e \in \{e'_1, \dots, e'_{n'}, e''_1, \dots e''_{n''}\}$, and so $T_q^{\mathbf{e}':\mathbf{e}''} \le T_q^{\mathbf{e}'|\mathbf{e}''}$ for all q. It therefore suffices to show that for $q, n \in \mathbb{N}$

$$\begin{aligned} &\int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e})\mathbb{P}\{T_q^{\mathbf{e}} \le t\} \\ &\le c(n, q, \psi)\left(\int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbb{P}\{T^{\mathbf{e}} \le t\}\right)^{\lfloor \frac{q}{3} \rfloor}. \end{aligned} \tag{8.7}$$

We begin with some notation. For any sequence of pairs

$$H = (\{x_1, y_1\}, \dots, \{x_\ell, y_\ell\}), \text{with } x_i, y_i \in E, \text{ and } x_i \ne y_i \text{ for } 1 \le i \le |H| := \ell,$$

and $t > 0$, define an event

$$A_t^H := \{T^{(x_1, y_1)} \le T^{(x_2, y_2)} \le \dots \le T^{(x_\ell, y_\ell)} \le t\}. \tag{8.8}$$

It is easy to see that $A_t^{H'} \supseteq A_t^H$ for any subsequence H' of H. Put $\overline{H} := \bigcup_{i=1}^{\ell}\{x_i, y_i\} \subseteq E$. For $z \in \overline{H}$, define $\iota(z, H) = \{1 \le i \le |H| : z \in \{x_i, y_i\}\}$ to be the set of indices of the pairs in which z appears.

Now fix $G = (\{x_1, y_1\}, \dots, \{x_q, y_q\})$, with $x_i, y_i \in E$, $x_i \ne y_i$, $1 \le i \le q$, and $\overline{G} := \bigcup_{i=1}^{q}\{x_i, y_i\} \subseteq \{e_1, \dots, e_n\}$. We wish to estimate $\mathbb{P}(A_t^G)$.

Let $(i_1, \ldots, i_h)$ be the subsequence of $(1, 2, \ldots, q)$ obtained by listing the elements of $\{\max \iota(z, G),\, z \in \overline{G}\}$ in increasing order. Define a subsequence G_* of G by

$$G_* := (\{x_{i_1}, y_{i_1}\}, \ldots, \{x_{i_h}, y_{i_h}\}) =: (\{x_{0,1}, y_{0,1}\}, \ldots, \{x_{0,|G_*|}, y_{0,|G_*|}\}) \tag{8.9}$$

(the reason for the alternative indexing will become clear as we proceed). Note that $|\overline{G}| = |\overline{G_*}|$ because for all $z \in \overline{G}$, $z \in \{x_{\max \iota(z,G)}, y_{\max \iota(z,G)}\} \subseteq \overline{G_*}$. By definition, for $1 \le j \le |G_*|$ the inequalities $\max \iota(x_{i_j}, G_*) \ge j$ and $\max \iota(y_{i_j}, G_*) \ge j$ hold, and at least one of these inequalities is an equality. In other words,

$$\min\{\max \iota(x_{0,j}, G_*), \max \iota(y_{0,j}, G_*)\} = j, \quad 1 \le j \le |G_*|. \tag{8.10}$$

Without loss of generality we can assume that $i_1 = \max \iota(x_{i_1}, G) \le \max \iota(y_{i_1}, G)$. Then $x_{i_1} \notin \{x_r, y_r\}$ for $i_1 < r \le q$, and, *a fortiori*, $x_{i_1} \notin \{x_{i_p}, y_{i_p}\}$ for $1 < p \le h$. Hence, $\iota(x_{0,1}, G_*) = \iota(x_{i_1}, G_*) = \{1\}$ and we are now in one of the following three cases:

Case I: $|\iota(y_{0,1}, G_*)| = 1$. Let G_1 be the subsequence of G_* obtained by deleting $\{x_{0,1}, y_{0,1}\}$. Then $\overline{G}_1 \cap \{x_{0,1}, y_{0,1}\} = \emptyset$,

$$\begin{aligned} \mathbb{P}(A_t^G) \le \mathbb{P}(A_t^{G_*}) &\le \mathbb{P}\left(\{T^{(x_{0,1}, y_{0,1})} \le t\} \cap A_t^{G_1}\right) \\ &= P\{T^{(x_{0,1}, y_{0,1})} \le t\}\mathbb{P}(A_t^{G_1}), \end{aligned} \tag{8.11}$$

and $|\overline{G}_1| = |\overline{G_*}| - 2 = |\overline{G}| - 2$.

Case II: $|\iota(y_{0,1}, G_*)| = 2$. Write $\iota(y_{0,1}, G_*) = \{1, j_2\}$. Define a subsequence G_1 of G_* by deleting $\{x_{0,1}, y_{0,1}\}$ and $\{x_{0,j_2}, y_{0,j_2}\})$ from G_*. Then $\overline{G}_1 \cap \{x_{0,1}, y_{0,1}\} = \emptyset$,

$$\begin{aligned} \mathbb{P}(A_t^G) \le \mathbb{P}(A_t^{G_*}) &\le \mathbb{P}\left(\{T^{(x_{0,1}, y_{0,1})} \le t\} \cap A_t^{G_1}\right) \\ &= \mathbb{P}\{T^{(x_{0,1}, y_{0,1})} \le t\}\mathbb{P}(A_t^{G_1}), \end{aligned} \tag{8.12}$$

and $|\overline{G}_1| \ge |\overline{G_*}| - 3 = |\overline{G}| - 3$.

Case III: $|\iota(y_{0,1}, G_*)| > 2$. Write $\iota(y_{0,1}, G_*) = \{1, j_2, \ldots, j_p\}$ where $1 < j_2 < \ldots < j_p$ and $y_{0,1} = y_{0,j_2} = \ldots = y_{0,j_p}$. Then $\max \iota(x_{0,j_2}, G_*) = j_2$ because $\max \iota(y_{0,j_2}, G_*) = j_p > j_2$. Let

$$G_{**} := (\{x_{0,1}, y'_{0,1}\}, \ldots, \{x_{0,|G_*|}, y'_{0,|G_*|}\}), \tag{8.13}$$

where

$$y'_{0,j} := \begin{cases} y_{0,j}, & \text{if } j \le j_2 \text{ or } y_{0,j} \ne y_{0,1}, \\ x_{0,j_2}, & \text{if } j > j_2 \text{ and } y_{0,j} = y_{0,1}. \end{cases} \tag{8.14}$$

We then have $\overline{G}_* = \overline{G}_{**}$ and by, switching $Z^{x_{0,j_2}}$ and $Z^{y_{0,j_2}}$ at time $T^{(x_{0,j_2}, y_{0,j_2})}$ in the manner described in Remark 8.1, we also have $\mathbb{P}(A_t^{G_*}) = \mathbb{P}(A_t^{G_{**}})$. Moreover, $\iota(x_{0,1}, G_{**}) = 1$ because $x_{0,1} \notin \bigcup_{j=j_2}^q \{x_{0,j}, y_{0,j}\} = \bigcup_{j=j_2}^q \{x_{0,j}, y'_{0,j}\}$. Now $|\iota(y_{0,1}, G_{**})| = 2$ and we are in Case II with G_* replaced by G_{**}. From the discussion in Case II we know that there exists a subsequence G_1 of G_{**} such that $\overline{G}_1 \cap \{x_{0,1}, y_{0,1}\} = \emptyset$,

$$\begin{aligned} \mathbb{P}(A_t^G) \le \mathbb{P}(A_t^{G_*}) = \mathbb{P}(A_t^{G_{**}}) &\le \mathbb{P}\left(\{T^{(x_{0,1}, y_{0,1})} \le t\} \cap A_t^{G_1}\right) \\ &= \mathbb{P}\{T^{(x_{0,1}, y_{0,1})} \le t\}\mathbb{P}(A_t^{G_1}), \end{aligned} \tag{8.15}$$

and $|\overline{G}_1| \geq |\overline{G}_{**}| - 3 = |\overline{G}_*| - 3 = |\overline{G}| - 3$.

The reduction procedure that transformed G into G_1 can be repeated at least $(\lfloor|\overline{G}|/3\rfloor - 1)^+$ more times. That is, for $0 \leq \ell \leq \lfloor|\overline{G}|/3\rfloor$, there exist sequences $G_\ell = (\{x_{\ell,1}, y_{\ell,1}\}, \ldots, \{x_{\ell,|G_\ell|}, y_{\ell,|G_\ell|}\}$ such that $G_0 = G$ and for $0 \leq \ell \leq \lfloor|\overline{G}|/3\rfloor - 1$

$$\mathbb{P}(A_t^{G_\ell}) \leq \mathbb{P}\left(\{T^{(x_{\ell,1},y_{\ell,1})} \leq t\} \cap A_t^{G_{\ell+1}}\right) = \mathbb{P}\{T^{(x_{l,1},y_{l,1})} \leq t\}\mathbb{P}(A_t^{G_{l+1}}), \tag{8.16}$$

with $\overline{G}_{\ell+1} \cap \{x_{\ell,1}, y_{\ell,1}\} = \emptyset$, $\overline{G}_{l+1} \subseteq \overline{G}_\ell$, and $|\overline{G}_{l+1}| \geq |\overline{G}_\ell| - 3$.

It follows that

$$\mathbb{P}(A_t^G) \leq \mathbb{P}\{T^{(x_{0,1},y_{0,1})} \leq t\}\mathbb{P}(A_t^{G_1}) \leq \ldots \leq \prod_{i=0}^{\lfloor\frac{|\overline{G}|}{3}\rfloor - 1} \mathbb{P}\{T^{(x_{i,1},y_{i,1})} \leq t\}, \tag{8.17}$$

where the sets $\{x_{i,1}, y_{i,1}\}$, $i = 0, 1, \ldots, \lfloor|\overline{G}|/3\rfloor - 1$, are pairwise disjoint.

Write $\mathbb{G}_t^{\mathbf{e}}$ for the set of possible values of $\Phi^{\mathbf{e}}(T_q^{\mathbf{e}})$ on the event $T_q^{\mathbf{e}} \leq t$. Note that $|\overline{G}| > q$ for any $G \in \mathbb{G}^{\mathbf{e}}$ and so the rightmost product in (8.17) has at least $\lfloor q/3\rfloor$ terms. Therefore, if we let $c(n, q, \psi)$ denote a constant that only depends on n, q, ψ (but not t), we have from (8.17) that

$$\begin{aligned}
&\int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e})\mathbb{P}\{T_q^{\mathbf{e}} \leq t\} \\
&= \int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e}) \sum_{G \in \mathbb{G}_t^{\mathbf{e}}} \mathbb{P}\{\Phi^{\mathbf{e}}(T_q^{\mathbf{e}}) = G,\, T_q^{\mathbf{e}} \leq t\} \\
&\leq \int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e}) \sum_{G \in \mathbb{G}_t^{\mathbf{e}}} \mathbb{P}(A_t^G) \\
&\leq c(n, q, \psi)\left(\int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbb{P}\{T^{\mathbf{e}} \leq t\}\right)^{\lfloor\frac{q}{3}\rfloor}.
\end{aligned} \tag{8.18}$$

□

Notation 8.7. Given $\mu \in \Xi$, define $\mu_t \in \Xi$, $t \geq 0$, by

$$\begin{aligned}
I(\mu_t; \phi) &:= \int m(de)\mathbb{P}\left[\int \mu(Z^e(t))(dk)\phi(e)(k)\right] \\
&= \int m(de)\mathbb{P}\left[\int \mu(e)(dk)\phi(\hat{Z}^e(t))(k)\right],
\end{aligned} \tag{8.19}$$

so that $I(\mu_t; \phi) = \mathbb{Q}^\mu[I(X_t; \phi)]$.

Lemma 8.8. *Consider $\mu \in \Xi$, $t \geq 0$, $q \in \mathbb{N}$, and $\phi = \psi \otimes \chi$ with non-negative $\psi \in L^1(m) \cap L^\infty(m)$ and non-negative $\chi \in C(K)$. There exists a constant $c(\phi, q)$ that only depends on ϕ and q (and not μ or t) such that*

$$\mathbb{Q}^\mu\left[\{I(X_t; \phi) - I(\mu_t; \phi)\}^{2q}\right] \leq c(\phi, q)\left(\int m^{\otimes 2}(d\mathbf{e})\psi^{\otimes 2}(\mathbf{e})\mathbb{P}\{T^{\mathbf{e}} \leq t\}\right)^{\lfloor\frac{q}{3}\rfloor}.$$

Proof. Given $n', n'' \in \mathbb{N}$ and vectors $\mathbf{f}' \in E^{n'}$ and $\mathbf{f}'' \in E^{n''}$ with $f_1', \ldots, f_{n'}'$, $f_1'', \ldots f_{n''}''$ distinct, write

$$
\begin{aligned}
L^{\mathbf{f}'|\mathbf{f}''}(t) := \int \bigotimes_{i'=1}^{n'} \mu(Z^{f'_{i'}}(t))(dk'_{i'}) \otimes \bigotimes_{i''=1}^{n''} \mu(Z^{f''_{i''}}(t))(dk''_{i'}) \\
\times \chi^{\otimes n'}(k'_1, \ldots, k'_{n'})\, \chi^{\otimes n''}\left(k''_{\gamma_1^{\mathbf{f}''}(t)}, \ldots, k''_{\gamma_{n''}^{\mathbf{f}''}(t)}\right).
\end{aligned} \tag{8.20}
$$

Then, by definition,

$$
\begin{aligned}
&\mathbb{Q}^\mu\left[\{I(X_t;\phi) - I(\mu_t;\phi)\}^{2q}\right] = \sum_{i=0}^{2q}(-1)^i \binom{2q}{i} I^{2q-i}(\mu_t;\phi)\mathbb{Q}^\mu[I^i(X_t;\phi)] \\
&= \sum_{i=0}^{2q}(-1)^i \binom{2q}{i} \int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'')\mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\right].
\end{aligned}
$$

Therefore, by Lemma 8.6, it suffices to show that

$$
\begin{aligned}
&\sum_{i=0}^{2q}(-1)^i \binom{2q}{i} \int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'') \\
&\quad \times \mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\mathbf{1}\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''}\right\}\right] \\
&= 0
\end{aligned} \tag{8.21}
$$

for $0 \le p \le q-1$.

For $\mathbf{e}' \in E^{2q-i}$ and $\mathbf{e}'' \in E^i$ with $e'_1, \ldots, e'_{2q-i}, e''_1, \ldots, e''_i$ distinct, write $\mathbb{S}_{j,h}^{\mathbf{e}'|\mathbf{e}''}$, $0 \le j \le 2q-i$, $0 \le h \le i$, for the collection of subsets of $\{e'_1, \ldots, e'_{2q-i}, e''_1, \ldots, e''_i\}$ with exactly j elements from $\{e'_1, \ldots, e'_{2q-i}\}$ and exactly h elements from $\{e''_1, \ldots, e''_i\}$. Put

$$
C_{j,h}^i := |\mathbb{S}_{j,h}^{\mathbf{e}'|\mathbf{e}''}| = \binom{2q-i}{j}\binom{i}{h}, \quad 0 \le j \le 2q-i,\ 0 \le h \le i. \tag{8.22}
$$

It is clear by construction that, recalling the transformation $H \mapsto \overline{H}$ from the proof of Lemma 8.6,

$$
\begin{aligned}
&\int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'')\mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\mathbf{1}\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''}\right\}\right] \\
&= \sum_{j,h} \int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'') \\
&\quad \times \mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\mathbf{1}\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''}\right\} \sum_{S \in \mathbb{S}_{j,h}^{\mathbf{e}'|\mathbf{e}''}} \mathbf{1}\left\{\overline{\mathbf{\Phi}^{\mathbf{e}'|\mathbf{e}''}}(T_p^{\mathbf{e}'|\mathbf{e}''}) = S\right\}\right] \\
&= \sum_{j,h} C_{j,h}^i \int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'') \\
&\quad \times \mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\mathbf{1}\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''},\ \overline{\mathbf{\Phi}^{\mathbf{e}'|\mathbf{e}''}}(T_p^{\mathbf{e}'|\mathbf{e}''}) = \{e'_1, \ldots, e'_j, e''_1, \ldots, e''_h\}\right\}\right].
\end{aligned}
$$

Note that $|\overline{\Phi^{\mathbf{e}'|\mathbf{e}''}}(T_p^{\mathbf{e}'|\mathbf{e}''})| \le 2p$, and so a necessary condition on j, h for a summand in the last term to be non-zero is that $j + h \le 2p < 2q$.

For fixed $\mathbf{e}'$ and $\mathbf{e}''$ write $\mathbf{e}^* := (\mathbf{e}'_1, \dots, \mathbf{e}'_{2q-i}, e''_{h+1}, \dots, e''_i)$ and $\mathbf{e}^{**} := (e''_1, \dots, e''_h)$. Observe that

$$\begin{aligned}&\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''},\ \overline{\Phi^{\mathbf{e}'|\mathbf{e}''}}(T_p^{\mathbf{e}'|\mathbf{e}''}) = \{e'_1, \dots, e'_j, e''_1, \dots, e''_h\}\right\} \\ &\quad = \left\{T_p^{\mathbf{e}^*|\mathbf{e}^{**}} \le t < T_{p+1}^{\mathbf{e}^*|\mathbf{e}^{**}},\ \overline{\Phi^{\mathbf{e}^*|\mathbf{e}^{**}}}(T_p^{\mathbf{e}^*|\mathbf{e}^{**}}) = \{e'_1, \dots, e'_j, e''_1, \dots, e''_h\}\right\}.\end{aligned} \tag{8.23}$$

Moreover, on this event the partition $\xi^{\mathbf{e}^{**}}(t)$ is the restriction of the partition $\xi^{\mathbf{e}''}(t)$ to $\mathbb{N}_h$, and hence $L^{\mathbf{e}'|\mathbf{e}''}(t) = L^{\mathbf{e}^*|\mathbf{e}^{**}}(t)$ on this event. Therefore, the quantity

$$\begin{aligned}&\int m^{\otimes 2q-i}(d\mathbf{e}') \otimes m^{\otimes i}(d\mathbf{e}'')\psi^{\otimes 2q-i}(\mathbf{e}')\psi^{\otimes i}(\mathbf{e}'') \\ &\quad \times \mathbb{P}\left[L^{\mathbf{e}'|\mathbf{e}''}(t)\mathbf{1}\left\{T_p^{\mathbf{e}'|\mathbf{e}''} \le t < T_{p+1}^{\mathbf{e}'|\mathbf{e}''},\ \overline{\Phi^{\mathbf{e}'|\mathbf{e}''}}(T_p^{\mathbf{e}'|\mathbf{e}''}) = \{e'_1, \dots, e'_j, e''_1, \dots, e''_h\}\right\}\right]\end{aligned} \tag{8.24}$$

does not vary as i ranges from h to $2q - j$.

The proof is complete once we note that for fixed h, j with $h < 2q - j$ we have

$$\begin{aligned}&\sum_{i=h}^{2q-j} (-1)^i \binom{2q}{i} C_{j,h}^i \\ &= \frac{(2q)!(-1)^h}{(2q-j-h)!j!h!} \sum_{i=h}^{2q-j} (-1)^{i-h} \frac{(2q-j-h)!}{(i-h)!(2q-i-j)!} \\ &= \frac{(2q)!(-1)^h}{(2q-j-h)!j!h!}(1-1)^{2q-j-h} \\ &= 0.\end{aligned} \tag{8.25}$$

□

Completion of the Proof of Theorem 7.2. Because X (as a Hunt process) has càdlàg paths $\mathbb{Q}^\mu$-a.s. for all $\mu \in \Xi$, it suffices to show that $I(X_\cdot; \phi)$ has continuous sample paths $\mathbb{Q}^\mu$-a.s. for all $\mu \in \Xi$ and all ϕ belonging to some countable subset of $L^1(m, C(K))$ that is separating for Ξ. Moreover, because $I(X_\cdot; \phi)$ already has càdlàg paths, verifying Kolmogorov's criterion establishes that these paths are, in fact, $\mathbb{Q}^\mu$–a.s. continuous. That is, verifying Kolmogorov's criterion does more than just establish the existence of a continuous version of X, it establishes that the version we already have is continuous.

Let $\{\hat{U}^\alpha\}_{\alpha > 0}$ denote the resolvent corresponding to the semigroup $\{\hat{P}_t\}_{t \ge 0}$. Suppose that $\mathcal{S}$ is a countable collection of bounded, m-integrable, continuous, non-negative functions on E with dense linear span in $L^1(m)$ (such a collection can be seen to exist by combining Lemma A.1 of [Eva97] with Proposition 3.4.2 of [EK86]). Note that if $\theta \in \mathcal{S}$, then $\alpha \hat{U}^\alpha \theta$ converges to θ pointwise as $\alpha \to \infty$. Also, $\int m(dx)\alpha\hat{U}^\alpha\theta(x) = \int m(dx)\theta(x) < \infty$ by the duality hypothesis for the pair $Z, \hat{Z}$. By a standard extension of Lebesgue's dominated convergence theorem (see, for example, Proposition 18 in Chapter 11 of [Roy68]), if $g \in L^\infty(m)$, then $\lim_{\alpha\to\infty} \int m(dx)\alpha\hat{U}^\alpha\theta(x)g(x) = \int m(dx)\theta(x)g(x)$.

Write $\mathcal{D} := \{\hat{U}^\alpha \theta : \theta \in \mathcal{S},\, \alpha \,\text{rational}\} \subseteq L^1(m) \cap L^\infty(m)$. It follows easily from what we have just observed that if $\tilde{C}$ is a countable dense subset of $\{\chi \in C(K) : \chi \ge 0\}$, then the countable collection of functions of the form $\psi \otimes \chi$, with $\psi \in \mathcal{D}$ and $\chi \in \tilde{C}$, is separating for Ξ.

Fix $\psi \in \mathcal{D}$ (with $\psi = \hat{U}^\alpha \theta$ for $\theta \in \mathcal{S}$ and α rational), $\chi \in \tilde{C}$, and $q \in \mathbb{N}$ such that $\lfloor q/3 \rfloor \varepsilon > 1$, where $\varepsilon > 0$ is as in the statement of the theorem. In order to show that $I(X_\cdot; \psi \otimes \chi)$ has $\mathbb{Q}^\mu$-a.s. continuous sample paths for all μ, it suffices by the Markov property of X and Kolmogorov's continuity criterion to show for some constants c and δ which depend only on ψ, χ, q that

$$\mathbb{Q}^\mu \left[\{ I(X_t; \psi \otimes \chi) - I(\mu; \psi \otimes \chi) \}^{2q} \right] \le c t^{1+\delta} \tag{8.26}$$

for all $t \ge 0$ and $\mu \in \Xi$. This, however, follows from Lemma 8.8 and the observation that

$$\begin{aligned}
|I(\mu_t; \psi \otimes \chi) - I(\mu; \psi \otimes \chi)| &\le \int m(dx) \left| \hat{P}_t \psi(x) - \psi(x) \right| \\
&= \int m(dx) \left| \int_t^\infty \left(e^{-\alpha(s-t)} - e^{-\alpha s} \right) \hat{P}_s \theta(x)\, ds - \int_0^t e^{-\alpha s} \hat{P}_s \theta(x)\, ds \right| \\
&\le 2\alpha^{-1}(1 - e^{-\alpha t}) \int m(dx) \theta(x) \le 2t \int m(dx) \theta(x),
\end{aligned} \tag{8.27}$$

where we have used the consequence of the duality hypothesis on Z, $\hat{Z}$ that $\int m(dx) \alpha \hat{P}_t \theta(x) = \int m(dx) \theta(x)$. □

9. Coalescing and annihilating circular Brownian motions

In this section we develop a duality relationship between systems of coalescing Brownian motions on $\mathbb{T}$, the circle of circumference 2π, and systems of annihilating Brownian motions on $\mathbb{T}$ (Proposition 9.1). This relation will be used in Section 10 to investigate the properties of the stepping–stone model X when the migration process is Brownian motion on $\mathbb{T}$. It will also be used in Section 11 to study the random tree associated with infinitely many coalescing Brownian motions on $\mathbb{T}$. We mention in passing that coalescing Brownian motion has recently become a topic of renewed interest (see, for example, [TW97] and [Tsi98]).

For the rest of this paper, Z (and hence $\hat{Z}$) will be standard Brownian motion on $\mathbb{T}$, and m will be normalised Lebesgue measure on $\mathbb{T}$.

Given a finite non–empty set $A \subseteq \mathbb{T}$, enumerate A as $\{e_1, \ldots, e_n\}$, put $\mathbf{e} := (e_1, \ldots, e_n)$, and define a process W^A, the *set–valued coalescing circular Brownian motion*, taking values in the collection of non–empty finite subsets of $\mathbb{T}$ by

$$W^A(t) := \{ \check{Z}^{\mathbf{e}}_{\gamma_1^{\mathbf{e}}(t)}(t), \ldots, \check{Z}^{\mathbf{e}}_{\gamma_n^{\mathbf{e}}(t)}(t) \} = \{ Z^{\mathbf{e}}_{\gamma_1^{\mathbf{e}}(t)}(t), \ldots, Z^{\mathbf{e}}_{\gamma_n^{\mathbf{e}}(t)}(t) \}, \quad t \ge 0. \tag{9.1}$$

Equivalently, $W^A(t)$ is the set of labels of the coalescing Markov labelled partition process $\zeta^{\mathbf{e}}(t)$. Of course, different enumerations of A lead to different processes, but all these processes will have the same distribution. In words, W^A describes the evolution of a finite set of indistinguishable Brownian particles with the feature that particles evolve independently between collisions but when two particles collide they coalesce into a single particle.

Write $\mathcal{O}$ for the collection of open subsets of $\mathbb{T}$ that are either empty or consist of a finite union of open intervals with distinct end–points. Given $B \in \mathcal{O}$, define on some probability space $(\Sigma, \mathcal{G}, \mathbb{Q})$ an $\mathcal{O}$–valued process V^B, the *annihilating circular*

Brownian motion as follows. The end–points of the constituent intervals execute independent Brownian motions on $\mathbb{T}$ until they collide, at which point they annihilate each other. If the two colliding end–points are from different intervals, then those two intervals merge into one interval. If the two colliding end–points are from the same interval, then that interval vanishes (unless the interval was arbitrarily close to $\mathbb{T}$ just before the collision, in which case the process takes the value $\mathbb{T}$). The process is stopped when it hits the empty set or $\mathbb{T}$.

We have the following duality relation between W^A and V^B.

Proposition 9.1. *For all finite, non–empty subsets $A \subseteq \mathbb{T}$, all sets $B \in \mathcal{O}$, and all $t \geq 0$,*

$$\mathbb{P}\{W^A(t) \subseteq B\} = \mathbb{Q}\{A \subseteq V^B(t)\}.$$

Proof. For $N \in \mathbb{N}$, let $\mathbb{Z}_N := \{0, 1, \dots N-1\}$ denote the integers modulo N. Let $\mathbb{Z}_N^{\frac{1}{2}} := \{\frac{1}{2}, \frac{3}{2}, \dots, \frac{2N-1}{2}\}$ denote the half–integers modulo N. A non-empty subset D of $\mathbb{Z}_N$ can be (uniquely) decomposed into "intervals": an interval of D is an equivalence class for the equivalence relation on the points of D defined by $x \sim y$ if and only if $x = y$, $\{x, x+1, \dots, y-1, y\} \subseteq D$, or $\{y, y+1, \dots, x-1, x\} \subseteq D$ (with all arithmetic modulo N). Any interval other than $\mathbb{Z}_N$ itself has an associated pair of (distinct) "end–points" in $\mathbb{Z}_N^{\frac{1}{2}}$: if the interval is $\{a, a+1, \dots, b-1, b\}$, then the corresponding end–points are $a - \frac{1}{2}$ and $b + \frac{1}{2}$ (with all arithmetic modulo N). Note that the end–points of different intervals of D are distinct.

For $C \subseteq \mathbb{Z}_N$, let W_N^C be a process on some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ taking values in the collection of non–empty subsets of $\mathbb{Z}_N$ that is defined in the same manner as W^A, with Brownian motion on $\mathbb{T}$ replaced by simple, symmetric (continuous time) random walk on $\mathbb{Z}_N$ (that is, by the continuous time Markov chain on $\mathbb{Z}_N$ that only makes jumps from x to $x+1$ or x to $x-1$ at a common rate $\lambda > 0$ for all $x \in \mathbb{Z}_N$). For $D \subseteq \mathbb{Z}_N$, let V_N^D be a process taking values in the collection of subsets of $\mathbb{Z}_N$ that is defined on some probability space $(\Sigma', \mathcal{G}', \mathbb{Q}')$ in the same manner as V^B, with Brownian motion on $\mathbb{T}$ replaced by simple, symmetric (continuous time) random walk on $\mathbb{Z}_N^{\frac{1}{2}}$ (with the same jump rate λ as in the definition of W_N^C). That is, end–points of intervals evolve as annihilating random walks on $\mathbb{Z}_N^{\frac{1}{2}}$.

The proposition will follow by a straightforward weak limit argument if we can show the following duality relationship between the coalescing "circular" random walk W_N^C and the annihilating "circular" random walk V_N^D:

$$\mathbb{P}'\{W_N^C(t) \subseteq D\} = \mathbb{Q}'\{C \subseteq V_N^D(t)\} \tag{9.2}$$

for all non–empty subsets of $C \subseteq \mathbb{Z}_N$, all subsets of $D \subseteq \mathbb{Z}_N$, and all $t \geq 0$.

It is simple, but somewhat tedious, to establish (9.2) by a generator calculation using the usual generator criterion for duality (see, for example, Corollary 4.4.13 of [EK86]). However, as Tom Liggett pointed out to us, there is an easier route. A little thought shows that V_N^D is nothing other than the (simple, symmetric) voter model on $\mathbb{Z}_N$. The analogous relationship between the annihilating random walk and the voter model on $\mathbb{Z}$ due to [Sch76] is usually called the *border equation* (see Section 2 of [BG80] for a discussion and further references). The relationship (9.2) is then just the analogue of the usual duality between the voter model and coalescing random walk on $\mathbb{Z}$ and it can be established in a similar manner by Harris's graphical method (again see Section 2 of [BG80] for a discussion and references). □

Remark 9.2. We have been unable to find an explicit reference to Proposition 9.1 or its analogue for Brownian motion on $\mathbb{R}$. However, if, in the $\mathbb{R}$–valued analogue, one considers a limit where A approaches a dense subset of $\mathbb{R}$, so that the system of coalescing Brownian motions converges to a coalescing Brownian flow, then the analogous result for the limiting flow can be found on p18 of [Arr79].

Recall $\mathbf{Z}$ and $\boldsymbol{\gamma}$ from Definition 6.2. Define set–valued processes $W^{[n]}$, $n \in \mathbb{N}$, and W by

$$W^{[n]}(t) := \{Z_{\gamma_1(t)}(t), \ldots, Z_{\gamma_n(t)}(t)\} \subseteq \mathbb{T},\ t \geq 0, \tag{9.3}$$

and

$$W(t) := \{Z_{\gamma_1(t)}(t), Z_{\gamma_2(t)}(t), \ldots\} \subseteq \mathbb{T},\ t \geq 0. \tag{9.4}$$

Thus, $W^{[1]}(t) \subseteq W^{[2]}(t) \subseteq \ldots$ and $\bigcup_{n\in\mathbb{N}} W^{[n]}(t) = W(t)$. Recall that $(W^{[n]}(t)_{t\geq 0}$ has the same law as $(\{Z^{[n]}_{\gamma_1^{[n]}(t)}(t), \ldots, Z^{[n]}_{\gamma_n^{[n]}(t)}(t)\})_{t\geq 0}$. Put $N(t) := |W(t)|$, the cardinality of the random set $W(t)$. Note that $N(t)$ is also the number of blocks in the partition $\xi(t)$, which is in turn the cardinality of the random set $\Gamma(t)$. It is clear that $\mathbb{P}^{[\infty]}$-a.s. $N(t)$ is a non–increasing, right–continuous function of t and if $N(t_0) < \infty$ for some $t_0 \geq 0$, then $N(t) - N(t-)$ is either 0 or -1 for all $t > t_0$. By the following corollary, $N(t) < \infty$, $\mathbb{P}^{[\infty]}$-a.s., for all $t > 0$.

Corollary 9.3. *For $t > 0$,*

$$\mathbb{P}^{[\infty]}[\,N(t)\,] = 1 + 2\sum_{n=1}^{\infty} \exp\left(-\left(\frac{n}{2}\right)^2 t\right) < \infty$$

and

$$\lim_{t\downarrow 0} t^{\frac{1}{2}}\mathbb{P}^{[\infty]}[\,N(t)\,] = 2\sqrt{\pi}.$$

Proof. Note that if B is a single open interval (and hence for all $t \geq 0$ the set $V^B(t)$ is either an interval or empty) and we let $L(t)$ denote the length of $V^B(t)$, then L is a Brownian motion on $[0, 2\pi]$ with $\mathrm{Var}L(t) = 2t$ that is stopped at the first time it hits $\{0, 2\pi\}$.

Now, for $M \in \mathbb{N}$ and $0 \leq i \leq M-1$ we have from the translation invariance of Z and Proposition 9.1 that

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\left\{W^{[n]}(t) \cap [2\pi i/M, 2\pi(i+1)/M] \neq \emptyset\right\} \\
&= 1 - \mathbb{P}^{[\infty]}\left\{W^{[n]}(t) \subseteq]0, 2\pi(M-1)/M[\right\} \\
&= 1 - \mathbb{P}^{[\infty]}\left\{W^{[n]}(0) \subseteq V^{]0,2\pi(M-1)/M[}(t)\right\},
\end{aligned} \tag{9.5}$$

where we take the annihilating process $V^{]0,2\pi(M-1)/M[}$ to be defined on the same probability space $(\Omega^{[\infty]}, \mathcal{F}^{[\infty]}, \mathbb{P}^{[\infty]})$ as the process $\mathbf{Z}$ that was used to construct $W^{[n]}$ and W, and we further take the processes $V^{]0,2\pi(M-1)/M[}$ and $\mathbf{Z}$ to be independent. Thus,

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{W(t) \cap [2\pi i/M, 2\pi(i+1)/M] \neq \emptyset\} \\
&= 1 - \mathbb{P}^{[\infty]}\left\{V^{]0,2\pi(M-1)/M[}(t) = \mathbb{T}\right\} \\
&= 1 - \tilde{\mathbb{P}}\left\{\tilde{\tau} \leq 2t,\ \tilde{B}(\tilde{\tau}) = 2\pi \mid \tilde{B}(0) = 2\pi(M-1)/M\right\},
\end{aligned} \tag{9.6}$$

where $\tilde{B}$ is a standard one–dimensional Brownian motion on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and $\tilde{\tau} = \inf\{s \geq 0 : \tilde{B}(s) \in \{0, 2\pi\}\}$.

By Theorem 4.1.1 of [Kni81] we have

$$\begin{aligned}
&\mathbb{P}^{[\infty]}[\,|W(t)|\,] \\
&= \lim_{M\to\infty} \mathbb{P}^{[\infty]}\left[\sum_{i=0}^{M-1} \mathbf{1}\,\{W(t) \cap [2\pi i/M, 2\pi(i+1)/M] \neq \emptyset\}\right] \\
&= \lim_{M\to\infty} M\left(1 - \tilde{\mathbb{P}}\left\{\tilde{\tau} \leq 2t,\ \tilde{B}(\tilde{\tau}) = 2\pi \mid \tilde{B}(0) = 2\pi(M-1)/M\right\}\right) \\
&= 1 - \lim_{M\to\infty} M \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(n\pi\left(\frac{M-1}{M}\right)\right) \exp\left(-\left(\frac{n}{2}\right)^2 t\right) \\
&= 1 + 2\sum_{n=1}^{\infty} \exp\left(-\left(\frac{n}{2}\right)^2 t\right) \\
&= \theta\left(\frac{t}{4\pi}\right) < \infty,
\end{aligned}$$

where

$$\theta(u) := \sum_{n=-\infty}^{\infty} \exp(-\pi n^2 u) \tag{9.7}$$

is the Jacobi theta function (we refer the reader to [BPY98] for a survey of many of the other probabilistic interpretations of the theta function). The proof is completed by recalling that θ satisfies the functional equation $\theta(u) = u^{-\frac{1}{2}}\theta(u^{-1})$ and noting that $\lim_{u\to\infty}\theta(u) = 1$. $\square$

We conjecture that $t^{\frac{1}{2}}N(t) \to 2\sqrt{\pi}$ as $t \downarrow 0$, $\mathbb{P}^{[\infty]}$–a.s. However, we are only able to prove the following weaker result, which will be used in Section 11. The proof will be given at the end of this section after some preliminaries.

Proposition 9.4. *With $\mathbb{P}^{[\infty]}$–probability one,*

$$0 < \liminf_{t\downarrow 0} t^{\frac{1}{2}} N(t) \leq \limsup_{t\downarrow 0} t^{\frac{1}{2}} N(t) < \infty.$$

For $t > 0$ the random partition $\xi(t)$ is, by Remark 6.3 and Corollary 9.3, exchangeable with a finite number of blocks. Let $1 = x_1^t < x_2^t < \ldots < x_{N(t)}^t$ be the list in increasing order of the minimal elements of the blocks of $\xi(t)$ (that is, a list in increasing order of the elements of the set $\Gamma(t)$). Results of Kingman (see Section 11 of [Ald85] for a unified account) and the fact that ξ evolves by pairwise coalescence of blocks give that $\mathbb{P}^{[\infty]}$–a.s. for all $t > 0$ the asymptotic frequencies

$$F_i(t) = \lim_{n\to\infty} n^{-1}|\{j \in \mathbb{N}_n : j \sim_{\xi(t)} x_i^t\}| \tag{9.8}$$

exist for $1 \leq i \leq N(t)$ and $F_1(t) + \cdots + F_{N(t)}(t) = 1$.

Lemma 9.5. *With $\mathbb{P}^{[\infty]}$–probability one,*

$$\lim_{t\downarrow 0} t^{-\frac{1}{2}} \sum_{i=1}^{N(t)} F_i(t)^2 = \frac{2}{\pi^{3/2}}.$$

Proof. Put $T_{ij} := \inf\{t \geq 0 : Z_i(t) = Z_j(t)\}$ for $i \neq j$. Observe that

$$\begin{aligned}\mathbb{P}^{[\infty]}\left[\sum_{i=1}^{N(t)} F_i(t)^2\right] &= \mathbb{P}^{[\infty]}\left[\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{k=1}^{n}\mathbf{1}\left\{j \sim_{\xi(t)} k\right\}\right] \\ &= \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2\} \\ &= \mathbb{P}^{[\infty]}\{T_{12} \leq t\}.\end{aligned}$$

From Theorem 4.1.1 of [Kni81] we have

$$\begin{aligned}&\mathbb{P}^{[\infty]}\{T_{12} \leq t\} \\ &= \frac{1}{2\pi}\int_0^{2\pi} 1 - \frac{4}{\pi}\sum_{n=1}^{\infty}\sin\left(\frac{(2n-1)x}{2}\right)\frac{1}{2n-1}\exp\left(-\left(\frac{2n-1}{2}\right)^2 t\right)\,dx \\ &= \frac{8}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\left\{1-\exp\left(-\left(\frac{2n-1}{2}\right)^2 t\right)\right\} \\ &= \frac{2}{\pi^2}\int_0^t\sum_{n=1}^{\infty}\exp\left(-\left(\frac{2n-1}{2}\right)^2 s\right)\,ds \\ &= \frac{2}{\pi^2}\int_0^t\frac{1}{2}\left\{\sum_{n=-\infty}^{\infty}\exp\left(-n^2\frac{s}{4}\right) - \sum_{n=-\infty}^{\infty}\exp\left(-n^2 s\right)\right\}\,ds \\ &= \frac{1}{\pi^2}\int_0^t\left\{\theta\left(\frac{s}{4\pi}\right)-\theta\left(\frac{s}{\pi}\right)\right\}\,ds,\end{aligned}$$

where θ is again the Jacobi theta function defined in (9.7). By the properties of θ recalled after (9.7),

$$\lim_{t\downarrow 0} t^{-\frac{1}{2}}\mathbb{P}^{[\infty]}\left[\sum_{i=1}^{N(t)} F_i(t)^2\right] = \lim_{t\downarrow 0} t^{-\frac{1}{2}}\mathbb{P}^{[\infty]}\{T_{12} \leq t\} = \frac{2}{\pi^{3/2}}. \tag{9.9}$$

Now

$$\begin{aligned}&\mathbb{P}^{[\infty]}\left[\left(\sum_{i=1}^{N(t)} F_i(t)^2\right)^2\right] \\ &= \mathbb{P}^{[\infty]}\left[\lim_{n\to\infty}\frac{1}{n^4}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\mathbf{1}\left\{i_1 \sim_{\xi(t)} i_2,\, i_3 \sim_{\xi(t)} i_4\right\}\right] \\ &= \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2,\, 3 \sim_{\xi(t)} 4\},\end{aligned}$$

and so

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^{N(t)} F_i(t)^2\right) &= \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2,\, 3 \sim_{\xi(t)} 4\} - \mathbb{P}^{[\infty]}\{T_{12} \leq t\}^2 \\ &= \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2,\, 3 \sim_{\xi(t)} 4\} - \mathbb{P}^{[\infty]}\{T_{12} \leq t,\, T_{23} \leq t\}.\end{aligned} \tag{9.10}$$

Observe that

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t,\, T_{13} > t,\, T_{14} > t,\, T_{23} > t,\, T_{24} > t\} \\
&\quad \le \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2,\, 3 \sim_{\xi(t)} 4,\, \{\{1,2,3,4\}\} \ne \xi^{[4]}(t)\} \\
&\quad \le \mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t\}
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t\} - \mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t,\, T_{13} > t,\, T_{14} > t,\, T_{23} > t,\, T_{24} > t\} \\
&\quad \le \sum_{i=1,2} \sum_{j=3,4} \mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t,\, T_{ij} \le t\}.
\end{aligned}$$

Thus

$$\begin{aligned}
\operatorname{Var}\left(\sum_{i=1}^{N(t)} F_i(t)^2\right) &\le \mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2 \sim_{\xi(t)} 3 \sim_{\xi(t)} 4\} \\
&\quad + \sum_{i=1,2} \sum_{j=3,4} \mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t,\, T_{ij} \le t\}.
\end{aligned} \tag{9.11}$$

Put $D_{ij} := |Z_i(0) - Z_j(0)|$. We have

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2 \sim_{\xi(t)} 3 \sim_{\xi(t)} 4\} \\
&\quad = \mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{13} \wedge T_{23} \le t,\, T_{14} \wedge T_{24} \wedge T_{34} \le t\} \\
&\quad = \mathbb{P}^{[\infty]}\Big(\{T_{12} \le t,\, T_{13} \wedge T_{23} \le t,\, T_{14} \wedge T_{24} \wedge T_{34} \le t\} \\
&\qquad\qquad \setminus \{D_{12} \le t^{\frac{2}{5}},\, (D_{13} \wedge D_{23}) \le t^{\frac{2}{5}},\, (D_{14} \wedge D_{24} \wedge D_{34}) \le t^{\frac{2}{5}}\}\Big) \\
&\qquad + \mathbb{P}^{[\infty]}\{D_{12} \le t^{\frac{2}{5}},\, (D_{13} \wedge D_{23}) \le t^{\frac{2}{5}},\, (D_{14} \wedge D_{24} \wedge D_{34}) \le t^{\frac{2}{5}}\} \\
&\quad \le \sum_{1 \le i < j \le 4} \mathbb{P}^{[\infty]}\{T_{ij} \le t,\, D_{ij} > t^{\frac{2}{5}}\} + \mathbb{P}^{[\infty]}\left\{\max_{1 \le i < j \le 4} D_{ij} \le 3t^{\frac{2}{5}}\right\},
\end{aligned} \tag{9.12}$$

where we have appealed to the triangle inequality in the last step. Because $\frac{2}{5} < \frac{1}{2}$, an application of the reflection principle and Brownian scaling certainly gives that the probability $\mathbb{P}^{[\infty]}\{T_{ij} \le t,\, D_{ij} > t^{\frac{2}{5}}\}$ is $o(t^\alpha)$ as $t \downarrow 0$ for any $\alpha > 0$. Moreover, by the translation invariance of m (the common distribution of the $Z_i(0)$), the second term in the rightmost member of (9.12) is at most

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{|Z_2(0) - Z_1(0)| \le 3t^{\frac{2}{5}},\, |Z_3(0) - Z_1(0)| \le 3t^{\frac{2}{5}},\, |Z_4(0) - Z_1(0)| \le 3t^{\frac{2}{5}}\} \\
&\quad = \mathbb{P}^{[\infty]}\{|Z_2(0)| \le 3t^{\frac{2}{5}},\, |Z_3(0)| \le 3t^{\frac{2}{5}},\, |Z_4(0)| \le 3t^{\frac{2}{5}}\} \\
&\quad = ct^{\frac{6}{5}},
\end{aligned}$$

for a suitable constant c when t is sufficiently small. Therefore,

$$\begin{aligned}
&\mathbb{P}^{[\infty]}\{1 \sim_{\xi(t)} 2 \sim_{\xi(t)} 3 \sim_{\xi(t)} 4\} \\
&\quad = \mathbb{P}^{[\infty]}\{\{T_{12} \le t,\, T_{13} \wedge T_{23} \le t,\, T_{14} \wedge T_{24} \wedge T_{34} \le t\} \\
&\quad = O(t^{\frac{6}{5}}), \quad \text{as } t \downarrow 0.
\end{aligned} \tag{9.13}$$

A similar argument establishes that

$$\mathbb{P}^{[\infty]}\{T_{12} \le t,\, T_{34} \le t,\, T_{ij} \le t\} = O(t^{\frac{6}{5}}), \quad \text{as } t \downarrow 0, \tag{9.14}$$

for $i = 1, 2$ and $j = 3, 4$.

Substituting (9.13) and (9.14) into (9.11) gives

$$\operatorname{Var}\left(\sum_{i=1}^{N(t)} F_i(t)^2\right) = O(t^{\frac{6}{5}}), \quad \text{as } t \downarrow 0. \tag{9.15}$$

This establishes the desired result when combined with the expectation calculation (9.9), Chebyshev's inequality, a standard Borel–Cantelli argument, and the monotonicity of $\sum_{i=1}^{N(t)} F_i(t)^2$. □

We may suppose that on our probability space $(\Omega^{[\infty]}, \mathcal{F}^{[\infty]}, \mathbb{P}^{[\infty]})$ there is a sequence $B_1, B_2, \ldots$ of i.i.d. one–dimensional standard Brownian motions with initial distribution the uniform distribution on $[0, 2\pi]$ and that Z_i is defined by setting $Z_i(t)$ to be the image of $B_i(t)$ under the usual homomorphism from $\mathbb{R}$ onto $\mathbb{T}$. For $n \in \mathbb{N}$ and $0 \le j \le 2^n - 1$, let $I_1^{n,j} \le I_2^{n,j} \le \ldots$ be a list in increasing order of the set of indices $\{i \in \mathbb{N} : B_i(0) \in [2\pi j/2^n, 2\pi(j+1)/2^n[\}$. Put $B_i^{n,j} := B_{I_i^{n,j}}$ and $Z_i^{n,j} := Z_{I_i^{n,j}}$. Thus $(B_i^{n,j})_{i\in\mathbb{N}}$ is an i.i.d. sequence of standard $\mathbb{R}$–valued Brownian motions and $(Z_i^{n,j})_{i\in\mathbb{N}}$ is an i.i.d. sequence of standard $\mathbb{T}$–valued Brownian motions. In each case the corresponding initial distribution is uniform on $[2\pi j/2^n, 2\pi(j+1)/2^n[$. Moreover, for $n \in \mathbb{N}$ fixed the sequences $(B_i^{n,j})_{i\in\mathbb{N}}$ are independent as j varies and the same is true of the sequences $(Z_i^{n,j})_{i\in\mathbb{N}}$.

Let $\underline{W}$ (resp. $\underline{W}^{n,j}$, $W^{n,j}$) be the coalescing system defined in terms of $(B_i)_{i\in\mathbb{N}}$ (resp. $(B_i^{n,j})_{i\in\mathbb{N}}$, $(Z_i^{n,j})_{i\in\mathbb{N}}$) in the same manner that W is defined in terms of $(Z_i)_{i\in\mathbb{N}}$.

It is clear by construction that

$$N(t) = |W(t)| \le \sum_{i=0}^{2^n-1} |W^{n,i}(t)| \le \sum_{i=0}^{2^n-1} |\underline{W}^{n,i}(t)|, \quad t > 0,\, n \in \mathbb{N}. \tag{9.16}$$

Lemma 9.6. *The expectation $\mathbb{P}^{[\infty]}[\,|\underline{W}(1)|\,]$ is finite.*

Proof. There is an obvious analogue of the duality relation Proposition 9.1 for systems of coalescing and annihilating one–dimensional Brownian motions. Using this duality and arguing as in the proof of Corollary 9.3, it is easy to see that, letting $\bar{L}$ and $\bar{U}$ be two independent, standard, real-valued Brownian motions on some probability space $(\bar\Omega, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ with $\bar L(0) = \bar U(0) = 0$,

$$\begin{aligned}
&\mathbb{P}^{[\infty]}[|\underline{W}(1)|] \\
&= \lim_{M\to\infty} \sum_{i=-\infty}^{\infty} \mathbb{P}^{[\infty]}\{\underline{W}(1) \cap [2\pi i/M, 2\pi(i+1)/M] \neq \emptyset\} \\
&= \lim_{M\to\infty} \sum_{i=-\infty}^{\infty} \bar{\mathbb{P}}\Big\{ \min_{0\le t\le 1} \left((\bar U(t) + 2\pi(i+1)/M) - (\bar L(t) + 2\pi i/M)\right) > 0, \\
&\quad [\bar L(1) + 2\pi i/M, \bar U(1) + 2\pi(i+1)/M] \cap [0, 2\pi] \neq \emptyset \Big\} \\
&\le \limsup_{M\to\infty} c' M \bar{\mathbb{P}}\left[\mathbf{1}\left\{\min_{0\le t\le 1}\left(\bar U(t) - \bar L(t)\right) > -2\pi/M\right\}\left(\bar U(1) - \bar L(1) + c''\right)\right]
\end{aligned}$$

for suitable constants c' and c''. Noting that $(\bar U - \bar L)/\sqrt{2}$ is a standard Brownian motion, the result follows from a straightforward calculation with the joint distribution of the minimum up to time 1 and value at time 1 of such a process (see, for example, Corollary 30 in Section 1.3 of [Fre83]). □

Proof of Proposition 9.4. By the Cauchy–Schwarz inequality,

$$1 = \left(\sum_{i=1}^{N(t)} F_i(t)\right)^2 \le N(t) \sum_{i=1}^{N(t)} F_i(t)^2, \tag{9.17}$$

and hence, by Lemma 9.5,

$$\liminf_{t\downarrow} t^{\frac{1}{2}} N(t) \ge \frac{\pi^{\frac{3}{2}}}{2}, \quad \mathbb{P}^{[\infty]} - a.s. \tag{9.18}$$

On the other hand, for each $n \in \mathbb{N}$, $|\underline{W}^{n,i}(2^{-2n})|$, $i = 0, \ldots, 2^n - 1$, are i.i.d. random variables which, by Brownian scaling, have the same distribution as $|\underline{W}(1)|$. By (9.16),

$$t^{\frac{1}{2}} N(t) \le \frac{1}{2^{n-1}} \sum_{i=0}^{2^n-1} |\underline{W}^{n,i}(2^{-2n})| \tag{9.19}$$

for $2^{-2n} < t \le 2^{-2(n-1)}$. An application of Lemma 9.6 and the following strong law of large numbers for triangular arrays completes the proof. □

Lemma 9.7. *Consider a triangular array $\{X_{n,i} : 1 \le i \le 2^n, n \in \mathbb{N}\}$ of identically distributed, real–valued, mean zero, random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that the collection $\{X_{n,i} : 1 \le i \le 2^n\}$ is independent for each $n \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} 2^{-n}\left(X_{n,1} + \cdots + X_{n,2^n}\right) = 0, \ \mathbb{P} - a.s.$$

Proof. This sort of result appears to be known in the theory of complete convergence. For example, it follows from the much more general Theorem A in [AK80] by taking $N_n = 2^n$ and $\psi(t) = 2^t$ in the notation of that result (see also the Example following that result). For the sake of completeness, we give a short proof that was pointed out to us by Michael Klass.

Let $\{Y_n : n \in \mathbb{N}\}$ be an independent identically distributed sequence with the same common distribution as the $X_{n,i}$. By the strong law of large numbers, for any $\varepsilon > 0$ the probability that $|Y_1 + \cdots + Y_{2^n}| > \varepsilon 2^n$ infinitely often is 0. Therefore, by the triangle inequality, for any $\varepsilon > 0$ the probability that $|Y_{2^n+1} + \cdots + Y_{2^{n+1}}| > \varepsilon 2^n$ infinitely often is 0; and so, by the Borel–Cantelli lemma for sequences of independent events,

$$\sum_n \mathbb{P}\{|Y_{2^n+1} + \cdots + Y_{2^{n+1}}| > \varepsilon 2^n\} < \infty \tag{9.20}$$

for all $\varepsilon > 0$. The last sum is also

$$\sum_n \mathbb{P}\{|X_{n,1} + \cdots + X_{n,2^n}| > \varepsilon 2^n\}, \tag{9.21}$$

and an application of the "other half" of the Borel–Cantelli lemma for possibly dependent events establishes that for all $\varepsilon > 0$ the probability of $|X_{n,1} + \cdots + X_{n,2^n}| > \varepsilon 2^n$ infinitely often is 0, as required. □

10. Finitely many pure types for circular Brownian migration

Recall that Z and $\hat{Z}$ are standard Brownian motions on the circle $\mathbb{T}$ and m is normalised Lebesgue measure. Recall also that $\mathcal{O}$ is the collection of open subsets of $\mathbb{T}$ that are either empty or the union of a finite number of disjoint intervals.

Definition 10.1. Let Ξ^o denote the subset of Ξ consisting of ν such that there exists a finite set $\{k_1^*, \ldots, k_N^*\} \subseteq K$ (depending on ν) with the property that for m-a.e. $e \in \mathbb{T}$ we can take $\nu(e) = \delta_{k_i^*}$ for some i, and, moreover, we can choose a version of ν such that the sets $\{e \in \mathbb{T} : \nu(e) = \delta_{k_j^*}\} \in \mathcal{O}$ for $1 \le j \le N$.

Theorem 10.2. *For all $\mu \in \Xi$, $\mathbb{Q}^\mu\{X_t \in \Xi^o$ for all $t > 0\} = 1$.*

Proof. Fix $\mu \in \Xi$ and $t > 0$. We will first show that

$$\mathbb{Q}^\mu\{X_t \in \Xi^o\} = 1. \tag{10.1}$$

By the same argument as in Proposition 5.1 of [Eva97], $\mathbb{Q}^\mu$-a.s. there is a random countable set of types K^* such that $X_t(e) \in \{\delta_k : k \in K^*\}$ for m-a.e. $e \in \mathbb{T}$. We can also require that K^* has been chosen "minimally" so that $m(\{e \in E : X_t(e) = \delta_k\}) > 0$ for all $k \in K^*$, $\mathbb{Q}^\mu$-a.s., and this requirement specifies K^* uniquely, $\mathbb{Q}^\mu$-a.s. For $n \in \mathbb{N}$ it is clear that on the event where K^* has cardinality at least n the dissimilarity $D_n(X_t)$ (recall Definition 6.1) is strictly positive $\mathbb{Q}^\mu$-a.s. It follows from Theorem 6.4 and Corollary 9.3 that K^* is finite $\mathbb{Q}^\mu$-a.s.

In order to show that a representative of the equivalence class of X_t in Ξ may be defined so that $\{e \in \mathbb{T} : X_t(e) = \delta_k\} \in \mathcal{O}$ for all $k \in K^*$, it suffices by the device used in the proof of Theorem 6.4 to consider the case where the probability measure $\mu(e)$ is diffuse for all $e \in \mathbb{T}$ and to show in this case that $\mathbb{Q}^\mu$-a.s. for all $k \in K^*$ the support of the measure $\mathbf{1}(X_t(e) = \delta_k)\, m(de)$ (which does not depend on the choice of equivalence class representative) is a connected set. For this, it in turn suffices to check that if $a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2$ are arranged in anti–clockwise order around

$\mathbb{T}$, then we have

$$\begin{aligned}(10.2)\qquad &\int m^{\otimes 4}(de)\mathbf{1}\left\{e_1 \in]a_1, b_1[,\ e_2 \in]a_2, b_2[,\ e_3 \in]c_1, d_1[,\ e_4 \in]c_2, d_2[\right\}\\ &\times \int \bigotimes_{i=1}^{4} X_t(e_i)(dk_i)\mathbf{1}\left\{k_3 \neq k_1 = k_2 \neq k_4\right\} = 0,\ \mathbb{Q}^{\mu}\text{-a.s.}\end{aligned}$$

or, equivalently by Remark 4.4,

$$\begin{aligned}(10.3)\qquad &\mathbf{1}\left\{Z_1^{[4]}(0) \in]a_1, b_1[,\ Z_2^{[4]}(0) \in]a_2, b_2[,\ Z_3^{[4]}(0) \in]c_1, d_1[,\ Z_4^{[4]}(0) \in]c_2, d_2[\right\}\\ &\times \mathbf{1}\left\{\gamma_3^{[4]}(t) \neq \gamma_1^{[4]}(t) = \gamma_2^{[4]}(t) \neq \gamma_4^{[4]}(t)\right\} = 0,\ \mathbb{P}^{[4]}\text{-a.s.}\end{aligned}$$

Write, for our fixed $t > 0$,

$$(10.4)\qquad T_{ij} = \inf\{0 \leq s \leq t : Z_i^{[4]}(s) = Z_j^{[4]}(s)\}, \quad 1 \leq i < j \leq 4,$$

for the first collision time of $Z_i^{[4]}$ and $Z_j^{[4]}$ before time t, with our standing convention that $\inf \emptyset = \infty$. We have $\mathbb{P}^{[4]}\{T_{ij} = T_{k\ell} \neq \infty\} = 0$ for $(i,j) \neq (k,\ell)$. Suppose that we have a realisation with the properties

$$(10.5)\qquad Z_1^{[4]}(0) \in]a_1, b_1[,\ Z_2^{[4]}(0) \in]a_2, b_2[,\ Z_3^{[4]}(0) \in]c_1, d_1[,\ Z_4^{[4]}(0) \in]c_2, d_2[,$$

$$(10.6)\qquad \gamma_3^{[4]}(t) \neq \gamma_1^{[4]}(t) = \gamma_2^{[4]}(t) \neq \gamma_4^{[4]}(t),$$

and

$$(10.7)\qquad T_{ij} \neq \infty \text{ implies } T_{ij} \neq T_{k\ell} \text{ for } (i,j) \neq (k,\ell).$$

In order that $\gamma_1^{[4]}(t) = \gamma_2^{[4]}(t)$ holds, we must have $T_{12} \neq \infty$. From the continuity of the paths of circular Brownian motion and (10.7), in order that (10.5) holds it must then be the case that

$$(10.8)\qquad T_{13} \wedge T_{14} \wedge T_{23} \wedge T_{24} < T_{12} \wedge T_{34}.$$

By construction, this would imply that $\gamma_3^{[4]}(t) = \gamma_1^{[4]}(t) = \gamma_2^{[4]}(t)$ or $\gamma_4^{[4]}(t) = \gamma_1^{[4]}(t) = \gamma_2^{[4]}(t)$, contradicting (10.6). Thus (10.3) holds and the proof of (10.1) is complete.

In order to establish the claim of the theorem, it suffices by (10.1) and the Markov property to consider the special case of $\mu \in \Xi^o$. Write $\{k_1^*, \ldots, k_N^*\} \subseteq K$ for the corresponding set of types k such that $m(\{e \in \mathbb{T} : \mu(e) = \delta_k\}) > 0$. Fix $1 \leq i \leq N$. Let $G \subseteq K$ be a closed and open set such that $k_i^* \in G$ and $k_j^* \notin G$ for $j \neq i$ (writing $k_i^* = (h_1, h_2, \ldots)$ one can take $G = \{(h_1', h_2', \ldots) \in K : h_1' = h_1, \ldots, h_n' = h_n\}$ for some sufficiently large n). It suffices to show for each such G that if we put $Y_t(e) := X_t(e)(G) \in [0,1]$, then $\mathbb{Q}^{\mu}$–a.s. for all $t \geq 0$ we can choose a representative of $Y_t \in L^\infty(\mathbb{T}, m)$ such that $Y_t(e) \in \{0,1\}$ for m-a.e. $e \in \mathbb{T}$ and $\{e \in \mathbb{T} : Y_t(e) = 1\} \in \mathcal{O}$.

By the remarks at the end of Section 4 of [Eva97], we have that Y is a Feller process with state-space the subset $L^\infty(\mathbb{T}, m; [0,1])$ of $L^\infty(\mathbb{T}, m)$ consisting of $[0,1]$-valued functions (where $L^\infty(\mathbb{T}, m; [0,1])$ is equipped with the relative weak* topology). Put $B := \{e \in \mathbb{T} : \mu(e) = \delta_{k_i^*}\} \in \mathcal{O}$. By the definition of X in Theorem 4.1

and Proposition 9.1, for $\psi \in L^1(m)$,

$$
\begin{aligned}
&\mathbb{Q}^\mu\left[\int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e})\prod_{i=1}^{n} Y_t(e_i)\right] \\
&= \int m^{\otimes n}(d\mathbf{e})\mathbb{P}\left[\psi^{\otimes n}(\mathbf{e})\prod_{j\in\Gamma^{\mathbf{e}}(t)} \mathbf{1}_B(Z_j^{\mathbf{e}}(t))\right] \\
&= \int m^{\otimes n}(d\mathbf{e})\mathbb{P}\left[\psi^{\otimes n}(\mathbf{e})\mathbf{1}\left\{W_t^{\{e_1,\dots,e_n\}} \subseteq B\right\}\right] \\
&= \mathbb{Q}\left[\int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e})\mathbf{1}\left\{\{e_1,\dots,e_n\} \subseteq V^B(t)\right\}\right] \\
&= \mathbb{Q}\left[\int m^{\otimes n}(d\mathbf{e})\psi^{\otimes n}(\mathbf{e})\prod_{i=1}^{n} \mathbf{1}_{V^B(t)}(e_i)\right]
\end{aligned}
\tag{10.9}
$$

(recall that V^B is defined on the probability space $(\Sigma, \mathcal{G}, \mathbb{Q})$). Consequently, the $L^\infty(\mathbb{T}, m; [0,1])$–valued processes Y and $\mathbf{1}_{V^B}$ have the same finite–dimensional distributions. Clearly, $t \mapsto \mathbf{1}_{V^B(t)}$ is continuous (in the weak* topology). Therefore, choosing our representative of Y_t to be $\mathbf{1}_{V^B(t)}$ for all $t \geq 0$ establishes the desired conclusion. □

11. The tree associated with coalescing circular Brownian motions

Recall that Z and $\hat{Z}$ are standard Brownian motions on $\mathbb{T}$ and m is normalised Lebesgue measure.

Definition 11.1. Given $i, j \in \mathbb{N}$, let $\tau_{ij} := \inf\{t \geq 0 : i \sim_{\xi(t)} j\}$ denote the first time that i and j belong to the same block. By Remark 6.3, the τ_{ij} are identically distributed. Metrise $\mathbb{N}$ with the (random) metric ρ given by $\rho(i,j) := \tau_{ij}$. Observe that ρ is an *ultrametric*; that is, the *strong triangle inequality* $\rho(i,j) \leq \rho(i,k) \vee \rho(k,j)$ holds for all i, j, k. Let $(\mathbb{F}, \rho)$ denote the completion of $(\mathbb{N}, \rho)$. The space $(\mathbb{F}, \rho)$ is also ultrametric. We refer the reader to Sections 18 and 19 of [Sch84] for basic facts about ultrametric spaces.

Some discussion of the space $(\mathbb{N}, \rho)$ can be found in Section 4 of [Ald93]. The analogue of $(\mathbb{F}, \rho)$ for another process of coalescing exchangeable partitions of $\mathbb{N}$, namely Kingman's coalescent, is considered in [Eva98] and the counterpart of Theorem 11.2 below is obtained.

For background on Hausdorff and packing dimension see [Mat95]. In order to establish some notation, we quickly recall the definitions of energy and capacity. Let (T, ρ) be a metric space. Write $M_1(T)$ for the collection of (Borel) probability measures on T. A *gauge* is a continuous, non–increasing function $f : [0, \infty[\to [0, \infty]$, such that $f(r) < \infty$ for $r > 0$, $f(0) = \infty$, and $\lim_{r\to\infty} f(r) = 0$. Given $\mu \in M_1(T)$ and a gauge f, the *energy of* μ *in the gauge* f is the quantity

$$\mathcal{E}_f(\mu) := \int \mu(dx) \int \mu(dy)\, f(\rho(x,y)).$$

The *capacity of* T *in the gauge* f is the quantity

$$\mathrm{Cap}_f(T) := (\inf\{\mathcal{E}_f(\mu) : \mu \in M_1(T)\})^{-1}$$

(note by our assumptions on f that we need only consider diffuse $\mu \in M_1(T)$ in the infimum).

Let $C_{\frac{1}{2}} \subseteq [0,1]$ denote the middle–$\frac{1}{2}$ Cantor set equipped with the usual Euclidean metric inherited from $[0,1]$. One of the assertions of the following result is, in the terminology of [PP95] (see, also, [BP92, PPS96, Per96]), that $\mathbb{F}$ is a.s. *capacity–equivalent* to $C_{\frac{1}{2}}$. Hence, by the results of [PPS96], $\mathbb{F}$ is also a.s. capacity–equivalent to the zero set of (one–dimensional) Brownian motion.

Theorem 11.2. *With $\mathbb{P}^{[\infty]}$–probability one, the ultrametric space $(\mathbb{F}, \rho)$ is compact with Hausdorff and packing dimensions both equal to $\frac{1}{2}$. There exist random variables K^*, K^{**} such that $\mathbb{P}^{[\infty]}$–almost surely $0 < K^* \le K^{**} < \infty$ and for every gauge f*

$$K^* \mathrm{Cap}_f(C_{\frac{1}{2}}) \le \mathrm{Cap}_f(\mathbb{F}) \le K^{**} \mathrm{Cap}_f(C_{\frac{1}{2}}). \tag{11.1}$$

Proof. The proof is essentially a reprise of the proof of Theorem 1.1 in [Eva98], with our Proposition 9.4 and Lemma 9.5 playing the role of the statements (2.1) and (2.2) in [Eva98]. □

Acknowledgements: We thank Jürgen Gärtner, Michael Klass, Tom Liggett, Jim Pitman and Tom Salisbury for helpful discussions. Some of the research was conducted while Evans, Fleischmann and Kurtz were visiting the Pacific Institute for the Mathematical Sciences in Vancouver and the Mathematical Sciences Research Institute in Berkeley in Fall 1997.

References

[AK80] S. Asmussen and T. G. Kurtz. Necessary and sufficient conditions for complete convergence in the law of large numbers. *Ann. Probab.*, 8(1):176–182, 1980.

[Ald85] D. J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII – 1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin – New York, 1985.

[Ald93] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.

[Arr79] R. A. Arratia. *Coalescing Brownian motions on the line.* PhD thesis, University of Wisconsin–Madison, 1979.

[BCG86] M. Bramson, J.T. Cox, and D. Griffeath. Consolidation rates for two interacting systems in the plane. *Probab. Theory Related Fields*, 73:613–625, 1986.

[Ber96] J. Bertoin. *Lévy Processes*, volume 121 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge – New York, 1996.

[BG68] R.M. Blumenthal and R.K. Getoor. *Markov Processes and Potential Theory.* Academic Press, New York, 1968.

[BG80] M. Bramson and D. Griffeath. Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}^1$. *Ann. Probab.*, 8:183–213, 1980.

[BP92] I. Benjamini and Y. Peres. Random walks on a tree and capacity in the interval. *Ann. Inst. H. Poincaré Probab. Statist.*, 28:557–592, 1992.

[BPY98] P. Biane, J. Pitman, and M. Yor. Representations of the Jacobi theta and Riemann zeta functions in terms of Brownian excursions. Preprint, Department of Statistics, University of California at Berkeley, 1998.

[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.*, 347:2277–2360, 1995.

[DK96] P. Donnelly and T.G. Kurtz. A countable representation of the Fleming–Viot measure–valued diffusion. *Ann. Probab.*, 24:698–742, 1996.

[DM78] C. Dellacherie and P.-A. Meyer. *Probabilities and Potential*, volume 29 of *North–Holland Mathematical Studies.* North–Holland, Amsterdam, 1978.

[DS58] N. Dunford and J.T. Schwartz. *Linear Operators, Part I: General Theory*. Interscience Publishers, 1958.

[DU77] J. Diestel and J.J. Uhl Jr. *Vector Measures*. American Mathematical Society, Providence, 1977.

[EF96] S.N. Evans and K. Fleischmann. Cluster formation in a stepping stone model with continuous, hierarchically structured sites. *Ann. Probab.*, 24:1926–1952, 1996.

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

[EP98] S.N. Evans and Y. Peres. Eventual intersection of sequences of Lévy processes. *Elect. Comm. in Probab.*, 3:21–27, 1998.

[Eva97] S.N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincaré Probab. Statist.*, 33:339–358, 1997.

[Eva98] S.N. Evans. Kingman's coalescent as a random metric space. Preprint available via http://www.stat.berkeley.edu/users/evans, 1998.

[FG94] K. Fleischmann and A. Greven. Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Related Fields*, 98:517–566, 1994.

[FG96] K. Fleischmann and A. Greven. Time-space analysis of the cluster-formation in interacting diffusions. *Electron. J. Probab.*, 1, 1996.

[Fre83] D. Freedman. *Brownian Motion and Diffusion*. Springer, New York, 1983.

[Get84] R.K. Getoor. Capacity theory and weak duality. In *Seminar on Stochastic Processes, 1983*, pages 97–130. Birkhäuser, Boston, 1984.

[GS84] R.K. Getoor and M.J. Sharpe. Naturality, standardness, and weak duality for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 67:1–62, 1984.

[Han90] K. Handa. A measure-valued diffusion process describing the stepping stone model with infinitely many alleles. *Stochastic Process. Appl.*, 36:269–296, 1990.

[Kle95] A. Klenke. Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab*, 24:660–697, 1995.

[Kni81] F.B. Knight. *Essentials of Brownian Motion and Diffusion*, volume 18 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, 1981.

[Mat95] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge – New York, 1995.

[MT95] C. Mueller and R. Tribe. Large time behavior of interface solutions to the heat equation with Fisher-Wright white noise. *Probab. Theory Related Fields*, 102:289–311, 1995.

[Per96] Y. Peres. Remarks on intersection–equivalence and capacity–equivalence. *Ann. Inst. H. Poincaré Phys. Théor.*, 64:339–347, 1996.

[PP95] R. Pemantle and Y. Peres. Galton–Watson trees with the same mean have the same polar sets. *Ann. Probab.*, 23:1102–1124, 1995.

[PPS96] R. Pemantle, Y. Peres, and J.W. Shapiro. The trace of spatial Brownian motion is capacity–equivalent to the unit square. *Probab. Theory Related Fields*, 106:379–399, 1996.

[Roy68] H. L. Royden. *Real Analysis*. Collier MacMillan – New York, 2nd edition, 1968.

[Sch76] D. Schwartz. On hitting probabilities for an annihilating particle model. *Ann. Probab.*, 6:398–403, 1976.

[Sch84] W. H. Schikhof. *Ultrametric Calculus: an Introduction to p-adic Analysis*, volume 4 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge – New York, 1984.

[Sha88] M. Sharpe. *General Theory of Markov Processes*. Academic Press, Boston, 1988.

[Shi80] T. Shiga. An interacting system in population genetics. *J. Math. Kyoto Univ.*, 20:213–242, 1980.

[Tsi98] B. Tsirelson. Brownian coalescence as a black noise I. Preprint, School of Mathematics, Tel Aviv University, 1998.

[TW97] B. Tóth and W. Werner. The true self–repelling motion. To appear in *Probab. Theory Related Fields*, 1997.

E-mail address: donnelly@stats.ox.ac.uk

Department of Statistics, Oxford University, 1 South Parks Road, Oxford OX1 3TG, U.K.

E-mail address: evans@stat.berkeley.edu

Department of Statistics #3860, University of California at Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A.

E-mail address: fleischmann@wias-berlin.de

Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, D-10117 Berlin, GERMANY

E-mail address: kurtz@math.wisc.edu

Departments of Mathematics and Statistics, University of Wisconsin–Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706-1388, U.S.A.

E-mail address: xzhou@stat.berkeley.edu

Department of Statistics #3860, University of California at Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A.