BALLS–IN–BOXES DUALITY FOR COALESCING RANDOM WALKS AND COALESCING BROWNIAN MOTIONS

STEVEN N. EVANS AND XIAOWEN ZHOU

ABSTRACT. We present a duality relation between two systems of coalescing random walks and an analogous duality relation between two systems of coalescing Brownian motions. Our results extends previous work in the literature and we apply it to the study of a system of coalescing Brownian motions with Poisson immigration.

1. Introduction

Consider a system of m indexed particles with locations in $\mathbb{R}$ that evolve as follows. Each particle moves according to an independent standard Brownian motion on $\mathbb{R}$ until two particles are at the same location. At this moment a *coalescence* occurs and the particle of higher index starts to move together with the particle of lower index. We say the particle with higher index is *attached* to the particle with lower index, which is still *free*. The particle system then continues its evolution in the same fashion. Note that indices are not essential here, the collection of locations of the particles is Markovian in its own right, but it will be convenient to think of the process as taking values in $\mathbb{R}^m$ rather than subsets of $\mathbb{R}$ with at most m elements. For definiteness, we will further assume that the particles are indexed in increasing order of their initial positions: it it clear that the dynamics preserve this ordering. Call the resulting Markov process $\mathbf{X} = (X_1, \dots, X_m)$.

The analogous coalescing simple random walk has many applications. One successful example is in voter model, which is particularly well understood because of a duality relation with the coalescing random walk (see, for example, [Gri79, Lig99]). Similarly, coalescing Brownian motion plays a key role in analyzing certain complex interactive stochastic systems. For example, in [DEF$^+$00] the coalescing Brownian motion is dual to the Brownian stepping-stone model in the sense that it determines the joint "moments" of the latter. This interplay leads to further results on the Brownian stepping-stone model in [Zho03]. A "continuous family" of coalescing Brownian motions, usually referred to as the Arratia flow, serves as a fundamental example in the theory of stochastic flows. See [Arr79, Har84] for accounts of this topic. The Arratia flow is an example of an interesting noise that is not generated by Brownian motions or Poisson processes [Tsi98, Tsi04]. More general "sticky" flows have recently been considered in [LJR02, LJR03, War02].

Closed form analytic expressions for features of the joint distribution of coalescing Brownian motion are rarely known, but some intriguing relationships have been observed for stochastic systems involving coalescing Brownian motions. A self-duality relation for the Arratia flow is described in the Introduction of [Arr79], where the borders between

SNE supported in part by NSF grant DMS-0071468 and a Miller Institute for Basic Research in Science research professorship, XZ supported by an NSERC grant.

clusters (that is, pre-images of particles) are shown to have the same joint distribution as the locations of particles. A duality between a system of coalescing Brownain motions and a system of annihilating Brownian motions is established in [DEF+00]. A dual relationship is presented in [STW00] between two system of Brownian motions, in which one system runs forward in time, the other runs backward in time, Brownian motions from the same system coalesce and Brownian motions from different systems reflect on each other. Another result along this line is obtained in in [TW97], which involves a duality on two flows of Brownian motions moving at opposite directions of the time interval $(-\infty, \infty)$. Within each flow, the Brownian motions coalesce, and meanwhile each Brownian motion is either reflected or absorbed at 0 depending on when it reaches 0. There is no interaction between the two flows.

The distribution of $\mathbf{X}(t)$ is uniquely specified by knowing for each choice of $y_1 < y_2 < \ldots < y_n$ the joint probabilities of which "balls" $X_1(t), X_2(t), \ldots, X_m(t)$ lie which of the "boxes" $[y_1, y_2], [y_2, y_3], \ldots, [y_{n-1}, y_n]$. That is, the distribution of $\mathbf{X}(t)$ is determined by the joint distribution of the indicators

$$I_{ij}^{\rightarrow}(t, \mathbf{y}) := 1\{X_i(t) \in [y_j, y_{j+1}]\}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n-1$.

Suppose now that $\mathbf{Y} := (Y_1, \ldots, Y_n)$ is another coalescing Brownian motion. The distribution of $\mathbf{Y}(t)$ is uniquely specified by knowing for each choice of $x_1 < x_2 < \ldots < x_n$ the distribution of the indicators

$$I_{ij}^{\leftarrow}(t, \mathbf{x}) := 1\{x_i \in [Y_j(t), Y_{j+1}(t)]\}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n-1$.

Thus we can think of a coalescing Brownian motion as being a set of evolving balls with the distribution at time t determined by how the balls fall in a fixed set of boxes, or we can think of a coalescing Brownian motion as giving a set of evolving boxes with the distribution at time t determined by how these boxes contain a fixed set of balls. We show that these two points of view are *dual* to each other in the sense that if $\mathbf{X}(0) = \mathbf{x}$ and $\mathbf{Y}(0) = \mathbf{y}$, then for each $t \geq 0$ the arrays of indicators $(I_{ij}^{\rightarrow}(t, \mathbf{y}))$ and $(I_{ij}^{\leftarrow}(t, \mathbf{x}))$ have the same joint distribution. We derive this duality from an analogous, essentially combinatorial, fact about coalescing simple random walk.

Special cases of the above mentioned duality were proved earlier in [XZ]. Instead of using a discrete approximation approach, the results there were directly obtained from coalescing Brownian motions, and, as a result, the proofs were rather lengthy.

Moreover, we extend the Brownian motion result to a situation where the "balls" and the "boxes" are allowed to originate at different points in time (rather than all originating at time 0). This latter extended result is then used to analyse the asymptotic behaviour of a system of coalescing Brownian particles in which new particles arise according to a homogeneous space–time Poisson point process.

The rest of this paper is arranged as follows. Section 2 contains the preparation and the proof of our main result on the duality between two coalescing simple random walks. In section 3 we generalize this dual relationship to coalescing Brownian motions starting from possibly different times. Some known results are re-derived. In section 4 we further generalize the duality to a model with both coalescing Brownian motion and Poisson migration included.

2. Coalescing random walk

A p-simple random walk on $\mathbb{Z}$ is a continuous time simple random walk that makes jumps at unit rate, and when it makes a jump from some site it jumps to the right neighbour with probability p and to the left neighbour with probability $1-p$. An m-dimensional p-simple coalescing random walk is defined in the same way as the coalescing Brownian motion of the Introduction. When $p = \frac{1}{2}$ we just call this particle system a simple coalescing random walk.

Some notation is useful to keep track of the interactions among the particles in the coalescing system. Let $\mathcal{P}_m$ denote the set of *interval partitions* of the totality of indices $\mathbb{N}_m := \{1, \ldots, n\}$. That is, an element π of $\mathcal{P}_n$ is a collection $\pi = \{A_1(\pi), \ldots, A_h(\pi)\}$ of disjoint subsets of $\mathbb{N}_m$ such that $\bigcup_i A_i(\pi) = \mathbb{N}_m$ and $a < b$ for all $a \in A_i$, $b \in A_j$, $i < j$. The sets $A_1(\pi), \ldots A_h(\pi)$ consisting of consecutive indices are the *intervals* of the partition π. The integer h is the *length* of π and is denoted by $l(\pi)$. Equivalently, we can think of $\mathcal{P}_m$ as a set of equivalence relations on $\mathbb{N}_m$ and write $i \sim_\pi j$ if i and j belong to the same interval of $\pi \in \mathcal{P}_m$. Of course, if $i \sim_\pi j$, then $i \sim_\pi k \sim_\pi j$ for all $i \le k \le j$.

Given $\pi \in \mathcal{P}_m$, define

$$\alpha_i(\pi) := \min A_i(\pi)$$

to be the left-hand end-point of the i^{th} interval $A_i(\pi)$. Put

$$\mathbb{Z}^m_\pi := \{(x_1, \ldots, x_m) \in \mathbb{Z}^m : x_1 \le \ldots \le x_m \text{ and } x_i = x_j \text{ if } i \sim_\pi j\}$$

and

$$\hat{\mathbb{Z}}^m_\pi := \{(x_1, \ldots, x_m) \in \mathbb{Z}^m : x_1 \le \ldots \le x_m \text{ and } x_i = x_j \text{ if and only if } i \sim_\pi j\}.$$

Note that $\mathbb{Z}^m$ is the disjoint union of the sets $\hat{\mathbb{Z}}^m_\pi$, $\pi \in \mathcal{P}_m$.

Write $\mathbf{X} = (X_1, \ldots, X_m)$ for the coalescing random walk. If $\mathbf{X}(t) \in \hat{\mathbb{Z}}^m_\pi$, then the free particles at time t have indices $\alpha_1(\pi), \ldots, \alpha_{l(\pi)}(\pi)$ and the i^{th} particle at time t is attached to the free particle with index

$$\min\{j : 1 \le j \le m,\, j \sim_\pi i\} = \max\{\alpha_k(\pi) : \alpha_k(\pi) \le i\}.$$

In order to write down the generator of $\mathbf{X}$, we require a final piece of notation. Let $\{\mathbf{e}^k_i : 1 \le i \le k\}$ be the set of coordinate vectors in $\mathbb{Z}^k$; that is, $\mathbf{e}^k_i$ is the vector that has i^{th} coordinate 1 and all other coordinates 0. For $\pi \in \mathcal{P}_m$, define a map $K_\pi : \mathbb{Z}^m_\pi \to \mathbb{Z}^{l(\pi)}$ by

$$K_\pi(\mathbf{x}) = K_\pi(x_1, \ldots, x_m) := (x_{\alpha_1(\pi)}, \ldots, x_{\alpha_{l(\pi)}(\pi)})$$

Notice that K_π is a bijection between $\mathbb{Z}^m_\pi$ and $\{x \in \mathbb{Z}^{l(\pi)} : x_1 \le x_2 \le \ldots \le x_{l(\pi)}\}$. For brevity, we will sometimes write $\mathbf{x}_\pi$ for $K_\pi(\mathbf{x})$.

Write $B(\mathbb{Z}^m)$ for the collection of all bounded functions on $\mathbb{Z}^m$. The generator G of $\mathbf{X}$ is the operator $G : B(\mathbb{Z}^m) \to B(\mathbb{Z}^m)$ given by

$$\begin{aligned} Gf(\mathbf{x}) := p \sum_{i=1}^{l(\pi)} f \circ K_\pi^{-1}(\mathbf{x}_\pi + \mathbf{e}_i^{l(\pi)}) + (1-p) \sum_{i=1}^{l(\pi)} f \circ K_\pi^{-1}(\mathbf{x}_\pi - \mathbf{e}_i^{l(\pi)}) \\ - l(\pi) f \circ K_\pi^{-1}(\mathbf{x}_\pi), \quad f \in B(\mathbb{Z}^m),\ \mathbf{x} \in \hat{\mathbb{Z}}_\pi^m,\ \pi \in \mathcal{P}_m. \end{aligned}$$

This expression is well-defined, because if $\mathbf{x} \in \hat{\mathbb{Z}}_\pi^m$, then $\mathbf{x}_\pi$, $\mathbf{x}_\pi + \mathbf{e}_i^{l(\pi)}$ and $\mathbf{x}_\pi - \mathbf{e}_i^{l(\pi)}$ are all in $\{x \in \mathbb{Z}^{l(\pi)} : x_1 \leq x_2 \leq \ldots \leq x_{l(\pi)}\}$.

Note: From now on we will suppress the dependence on dimension and write $\mathbf{e}_i^{l(\pi)}$ as $\mathbf{e}_i$.

Write $\mathbb{Z}' := \mathbb{Z} + \frac{1}{2} = \{i + \frac{1}{2} : i \in \mathbb{Z}\}$. An n-dimensional q-simple coalescing random walk on $\mathbb{Z}'^n$ and its generator H can be defined in the obvious way. Such a process, with $q = 1-p$, will serve as the process dual to the p-simple coalescing random walk on $\mathbb{Z}^m$ in the following way.

Fix $\mathbf{x} \in \mathbb{Z}^m$ with $x_1 \leq \ldots \leq x_m$ and $\mathbf{y} \in \mathbb{Z}'^n$ with $y_1 \leq \ldots \leq y_n$. By analogy with the notation introduced in the Introduction, put

$$I_{ij}^{\rightarrow}(t, \mathbf{y}) := 1\{X_i(t) \in [y_j, y_{j+1}]\}$$

and

$$I_{ij}^{\leftarrow}(t, \mathbf{x}) := 1\{x_i \in [Y_j(t), Y_{j+1}(t)]\}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n-1$.

Theorem 2.1. *Suppose in the notation above that $\mathbf{X} = (X_1, \ldots, X_m)$ is an m-dimensional p-simple coalescing random walk starting at $\mathbf{x} = (x_1, \ldots, x_m)$ with $x_1 \leq \ldots \leq x_m$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ is a n-dimensional $(1-p)$-simple coalescing random walk starting at $\mathbf{y} = (y_1, \ldots, y_n)$ with $y_1 \leq \ldots \leq y_m$. Then for each $t \geq 0$ the joint distribution of the $m \times (n-1)$-dimensional random array $(I_{ij}^{\rightarrow}(t, \mathbf{y}))$ coincides with that of the $m \times (n-1)$-dimensional random array $(I_{ij}^{\leftarrow}(t, \mathbf{x}))$.*

Proof. For a function $g : \{0,1\}^{m(n-1)} \to \mathbb{R}$, a vector $\tilde{\mathbf{x}} \in \mathbb{Z}^m$ with $\tilde{x}_1 \leq \ldots \leq \tilde{x}_m$, and a vector $\tilde{\mathbf{y}} \in \mathbb{Z}'^n$ with $\tilde{y}_1 \leq \ldots \leq \tilde{y}_n$, set

$$\bar{g}(\tilde{\mathbf{x}}; \tilde{\mathbf{y}}) := g(1_{[\tilde{y}_1, \tilde{y}_2]}(\tilde{x}_1), \ldots, 1_{[\tilde{y}_{n-1}, \tilde{y}_n]}(\tilde{x}_1), \ldots, 1_{[\tilde{y}_1, \tilde{y}_2]}(\tilde{x}_m), \ldots, 1_{[\tilde{y}_{n-1}, \tilde{y}_n]}(\tilde{x}_m)).$$

We may assume that $\mathbf{X}$ and $\mathbf{Y}$ are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We need to show that

$$\mathbb{P}[\bar{g}(\mathbf{X}_t; \mathbf{y})] = \mathbb{P}[\bar{g}(\mathbf{x}; \mathbf{Y}_t)]. \tag{2.1}$$

For $\tilde{\mathbf{x}} \in \mathbb{Z}^m$, put $\bar{g}_{\tilde{\mathbf{x}}}(\cdot) := \bar{g}(\tilde{\mathbf{x}}; \cdot)$, and for $\tilde{\mathbf{y}}' \in \mathbb{Z}'^n$, put $\bar{g}_{\tilde{\mathbf{y}}}(\cdot) := \bar{g}(\cdot; \tilde{\mathbf{y}})$. In order to establish (2.1), it suffices by a standard argument (cf. Section 4.4 in [EK86]) to show that

$$G(\bar{g}_{\mathbf{y}})(\mathbf{x}) = H(\bar{g}_{\mathbf{x}})(\mathbf{y}) \tag{2.2}$$

(recall that G and H are the generators of $\mathbf{X}$ and $\mathbf{Y}$, respectively).

Fix $\mathbf{x} \in \hat{\mathbb{Z}}_\pi^m$ and $\mathbf{y} \in \hat{\mathbb{Z}}_\varpi'^n$ for some $\pi \in \mathcal{P}_m$ and $\varpi \in \mathcal{P}_n$. Set

$$I^+ := \{i : 1 \leq i \leq l(\pi),\ x_{\alpha_i(\pi)} + \frac{1}{2} = y_{\alpha_j(\varpi)} \text{ for some } 1 \leq j \leq l(\varpi)\}$$

and

$$I^- := \{i : 1 \le i \le l(\pi),\ x_{\alpha_i(\pi)} - \frac{1}{2} = y_{\alpha_j(\varpi)} \text{ for some } 1 \le j \le l(\varpi)\}.$$

Similarly, put

$$J^- := \{j : 1 \le j \le l(\varpi),\ y_{\alpha_j(\varpi)} - \frac{1}{2} = x_{\alpha_i(\pi)} \text{ for some } 1 \le i \le l(\pi)\}$$

and

$$J^+ := \{j : 1 \le j \le l(\varpi),\ y_{\alpha_j(\varpi)} + \frac{1}{2} = x_{\alpha_i(\pi)} \text{ for some } 1 \le i \le l(\pi)\}.$$

Recall that $x_{\alpha_1(\pi)} < \ldots < x_{\alpha_{l(\pi)}(\pi)}$ and $y_{\alpha_1(\varpi)} < \ldots < y_{\alpha_{l(\varpi)}(\varpi)}$. Therefore, for each $i \in I^+$ there is a unique $j \in J^-$ such that $x_{\alpha_i(\pi)} + \frac{1}{2} = y_{\alpha_j(\varpi)}$ and *vice versa*. Fix such a pair (i, j). Observe that

$$\mathbf{x}' := \mathbf{x} + \sum_{k \in A_i(\pi)} \mathbf{e}_k^m = K^{-1}(\mathbf{x}_\pi + \mathbf{e}_i)$$

and

$$\mathbf{y}' := \mathbf{y} - \sum_{k \in A_j(\varpi)} \mathbf{e}_k^n = K^{-1}(\mathbf{y}_\varpi - \mathbf{e}_j).$$

Writing $1(B)(\cdot)$ for the indicator function of a set B, we are going to verify that

$$(1([y_{j'}, y_{j'+1}])(x'_{i'})) = (1([y'_{j'}, y'_{j'+1}])(x_{i'})) \tag{2.3}$$

by considering all the possible scenarios.

Given any $i' \in A_i(\pi)$ we have:

- for $j' = \alpha_j(\varpi) - 1$,
$$\begin{aligned} 1([y_{j'}, y_{j'+1}])(x'_{i'}) &= 1([y_{j'}, y_{j'+1}])(x_{i'} + 1) \\ &= 0 \\ &= 1([y_{j'}, y_{j'+1} - 1])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}), \end{aligned}$$
- for $\alpha_j(\varpi) \le j' < \max A_j(\varpi)$,
$$\begin{aligned} 1([y_{j'}, y_{j'+1}])(x'_{i'}) &= 1([y_{j'}, y_{j'+1}])(x_{i'} + 1) \\ &= 0 \\ &= 1([y_{j'} - 1, y_{j'+1} - 1])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}), \end{aligned}$$
- for $j' = \max A_j(\varpi)$,
$$\begin{aligned} 1([y_{j'}, y_{j'+1}])(x'_{i'}) &= 1([y_{j'}, y_{j'+1}])(x_{i'} + 1) \\ &= 1 \\ &= 1([y_{j'} - 1, y_{j'+1}])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}), \end{aligned}$$
- and for $j' < \alpha_j(\varpi) - 1$ or $j' > \max A_j(\varpi)$,
$$\begin{aligned} 1([y_{j'}, y_{j'+1}])(x'_{i'}) &= 1([y_{j'}, y_{j'+1}])(x'_{i'}) = 1([y_{j'}, y_{j'+1}])(x_{i'} + 1) \\ &= 0 \\ &= 1([y_{j'}, y_{j'+1}])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}). \end{aligned}$$

Moreover, given any $i' \notin A_i(\pi)$, we have $x_{i'} \neq x_{\alpha_i(\pi)}$. Hence

- for $j' = \alpha_j(\varpi) - 1$,

$$1([y_{j'}, y_{j'+1}])(x'_{i'}) = 1([y_{j'}, y_{j'+1}])(x_{i'}) = 1([y_{j'}, y_{j'+1} - 1])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}),$$

- for $j' = \max A_j(\varpi)$,

$$1([y_{j'}, y_{j'+1}])(x'_{i'}) = 1([y_{j'}, y_{j'+1}])(x_{i'}) = 1([y_{j'} - 1, y_{j'+1}])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}),$$

- for $\alpha_j(\varpi) \le j' < \max A_j(\varpi)$,

$$1([y_{j'}, y_{j'+1}])(x'_{i'}) = 1([y_{j'}, y_{j'+1}])(x_{i'}) = 1([y_{j'} - 1, y_{j'+1} - 1])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}),$$

- and for $j' < \alpha_j(\varpi) - 1$ or $j' > \max A_j(\varpi)$,

$$1([y_{j'}, y_{j'+1}])(x'_{i'}) = 1([y_{j'}, y_{j'+1}])(x_{i'}) = 1([y'_{j'}, y'_{j'+1}])(x_{i'}).$$

Combining the above observations yield (2.3).

Therefore,

$$\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi + \mathbf{e}_i) = \bar{g}_{\mathbf{x}} \circ K_\pi^{-1}(\mathbf{y}_\varpi - \mathbf{e}_j).$$

Furthermore, it is easy to see for $i' \notin I^+$ that

$$\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi + \mathbf{e}_{i'}) = \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi)$$

and for $j' \notin J^-$ that

$$\bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi - \mathbf{e}_{j'}) = \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi).$$

Similarly, for any $i \in I^-$ there exists a unique $j \in J^+$ such that $x_{\alpha_i(\pi)} - \frac{1}{2} = y_{\alpha_j(\varpi)}$ and *vice versa*. For such a pair (i, j) we have

$$\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi - \mathbf{e}_i) = \bar{g}_{\mathbf{x}} \circ K_\pi^{-1}(\mathbf{y}_\varpi + \mathbf{e}_j).$$

Furthermore, we see for $i' \notin I^-$ that

$$\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi - \mathbf{e}_{i'}) = \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi)$$

and for $j' \notin J^+$ that

$$\bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi + \mathbf{e}_{j'}) = \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi).$$

Lastly, note that

$$\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi) = \bar{g}(\mathbf{x}; \mathbf{y}) = \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi)$$

and so

$$
\begin{aligned}
&G(\bar{g}_{\mathbf{y}})(\mathbf{x}) - H(\bar{g}_{\mathbf{x}})(\mathbf{y}) \\
&\quad = p \sum_{i=1}^{l(\pi)} \left(\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi + \mathbf{e}_i) - \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi) \right) \\
&\qquad + (1-p) \sum_{i=1}^{l(\pi)} \left(\bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi - \mathbf{e}_i) - \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi) \right) \\
&\qquad - p \sum_{j=1}^{l(\varpi)} \left(\bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi - \mathbf{e}_i) - \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi) \right) \\
&\qquad - (1-p) \sum_{j=1}^{l(\varpi)} \left(\bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi + \mathbf{e}_i) - \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi) \right) \\
&\quad = p \sum_{i \in I^+} \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi + \mathbf{e}_i) - p \sum_{j \in J^-} \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi - \mathbf{e}_j) \\
&\qquad + (1-p) \sum_{i \in I^-} \bar{g}_{\mathbf{y}} \circ K_\pi^{-1}(\mathbf{x}_\pi - \mathbf{e}_i) - (1-p) \sum_{j \in J^+} \bar{g}_{\mathbf{x}} \circ K_\varpi^{-1}(\mathbf{y}_\varpi + \mathbf{e}_j) \\
&\quad = 0,
\end{aligned}
$$

as required. □

Remark 2.2. One can see from the proof that it is crucial that the random walks make only nearest neighbor jumps.

3. Coalescing Brownian motion

In this section we will show that the duality in Theorem 2.1 also holds when the coalescing random walks are replaced by coalescing Brownian motions. Coalescing Brownian motion can be defined similarly to coalescing random walk. This duality between two coalescing Brownian motions follows if one can show the unsurprising fact that a coalescing random walk scaled in time and space in the usual way converges weakly to a coalescing Brownian motion.

Proposition 3.1. *The conclusion of Theorem 2.1 holds when the coalescing random walks are replaced by coalescing Brownian motions in the definition of* $(I_{ij}^{\rightarrow}(t, \mathbf{y}))$ *and* $(I_{ij}^{\leftarrow}(t, \mathbf{x}))$.

We omit the the proof, but remark that a particularly straightforward martingale argument proof of the convergence of coalescing random walk to coalescing Brownian motion can be given using the following result that parallels Lévy's celebrated martingale characterization of Brownian motion (and is a fairly simple consequence of that result). We also omit the proof of this theorem.

Theorem 3.2. *Let* $\mathbf{X}$ *be an* m*-dimensional continuous process with* $\mathbf{X}(0) = \mathbf{x}$*, where* $x_1 \leq \ldots \leq x_m$*, and let* $\mathcal{F}^{\mathbf{X}}$ *denote the filtration generated by* $\mathbf{X}$*. Then the following are equivalent.*

(i) The process $\mathbf{X}$ *is a coalescing Brownian motion.*
(ii) For each $1 \le i \le m$*, the process* X_i *is a Brownian motion with respect to* $\mathcal{F}^{\mathbf{X}}$*, and for each pair* $1 \le i < j \le n$*, the process* $\frac{1}{\sqrt{2}}(X_j - X_i)$ *is a Brownian motion stopped at* 0 *with respect to* $\mathcal{F}^{\mathbf{X}}$*.*
(iii) The process $\mathbf{X}$ *is a continuous martingale with quadratic variation* $\langle X_i, X_j\rangle_t = t - T_{ij} \wedge t$*, where* $T_{ij} := \inf\{s \ge 0 : X_i(s) = X_j(s)\}$*,* $1 \le i \le j \le n$*.*

For a function $g : \{0,1\}^{m(n-1)} \to \mathbb{R}$, a vector $\tilde{\mathbf{x}} \in \mathbb{Z}^m$ with $\tilde{x}_1 \le \ldots \le \tilde{x}_m$, and a vector $\tilde{\mathbf{y}} \in \mathbb{Z}'^m$ with $\tilde{y}_1 \le \ldots \le \tilde{y}_n$, set

$$\bar{g}(\tilde{\mathbf{x}};\tilde{\mathbf{y}}) := g(1_{[\tilde{y}_1,\tilde{y}_2]}(\tilde{x}_1), \ldots, 1_{[\tilde{y}_{n-1},\tilde{y}_n]}(\tilde{x}_1), \ldots, 1_{[\tilde{y}_1,\tilde{y}_2]}(\tilde{x}_m), \ldots, 1_{[\tilde{y}_{n-1},\tilde{y}_n]}(\tilde{x}_m)).$$

Proposition 3.1 says that for any $\mathbf{x}$ and $\mathbf{y}$

$$\mathbb{P}[\bar{g}(\mathbf{X}_t;\mathbf{y})] = \mathbb{P}[\bar{g}(\mathbf{x};\mathbf{Y}_t)].$$

By choosing the right function g we can recover some known dualities. For example, given $\pi = (A_1, \ldots, A_h) \in \mathcal{P}_n$ and $y_1 < \ldots < y_{2h}$, put

$$\bar{g}(\mathbf{x};\mathbf{y}) = \prod_{j=1}^{h} \prod_{i \in A_j} 1_{[y_{2j-1},y_{2j}]}(x_i), \ \mathbf{x} \in \mathbb{R}^n, \ \mathbf{y} \in \mathbb{R}^m.$$

Then Proposition 3.1 implies that

$$\mathbb{P}\left\{\bigcap_{j=1}^{h} \bigcap_{i \in A_j} \{X_i(t) \in [y_{2j-1}, y_{2j}]\}\right\} = \mathbb{P}\left\{\bigcap_{j=1}^{h} \bigcap_{i \in A_j} \{x_i \in [Y_{2j-1}(t), Y_{2j}(t)]\}\right\},$$

which gives the duality in Theorem 1.1 of [XZ]. If we choose

$$\bar{g}(\mathbf{x};\mathbf{y}) = \prod_{i=1}^{n}\left(1 - \prod_{j=1}^{m}(1 - 1_{[y_{2j-1},y_{2j}]}(x_i))\right), \ \mathbf{x} \in \mathbb{R}^n, y_1 < \ldots < y_{2m},$$

then

$$\mathbb{P}\left\{\bigcap_{i=1}^{n}\{X_i(t) \in \bigcup_{j=1}^{m}[y_{2j-1}, y_{2j}]\}\right\} = \mathbb{P}\left\{\bigcap_{i=1}^{n}\{X_i(0) \in \bigcup_{j=1}^{m}[Y_{2j-1}(t), Y_{2j}(t)]\}\right\}.$$

Therefore, Proposition 3.7 in [XZ] follows readily.

The duality Proposition 3.1 can be generalized to one involving coalescing Brownian motion starting from different times. In order to state this result, it will be convenient to think of coalescing Brownian motion a little differently from what we have done so far. As we have defined it, the coalescing Brownian motion $\mathbf{X}$ takes values in the space $\{\mathbf{x} \in \mathbb{R}^m : x_1 \le x_2 \le \ldots \le x_m\}$. It will be more convenient to work with a related process for which we don't impose this condition. Given an arbitrary $\mathbf{x} \in \mathbb{R}^m$, let σ be any permutation of the indices $\{1, 2, \ldots, m\}$ such that $x_{\sigma(1)} \le x_{\sigma(2)} \le \ldots x_{\sigma(m)}$. Let $\tilde{\mathbf{X}}$ be an $\mathbb{R}^m$-valued process that has the same distribution as the process $(X_{\sigma^{-1}(1)}, X_{\sigma^{-1}(2)}, \ldots, X_{\sigma^{-1}(m)})$, where $\mathbf{X}(0) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(m)})$. It is not difficult to see that $\tilde{\mathbf{X}}$ is a time-homogeneous strong Markov process. The following result is obvious.

Corollary 3.3. *The duality in Proposition 3.1 holds when the ordered coalescing Brownian motion* $\mathbf{X}$ *is replaced by the unordered coalescing Brownian motion* $\tilde{\mathbf{X}}$.

Given $((s_1, x_1), \ldots, (s_m, x_m)) \in (\mathbb{R}_+ \times \mathbb{R})^m$ with $0 \leq s_1 \leq s_2 \leq \ldots \leq s_m$, define a process $\bar{\mathbf{X}}$ taking values in $\{\epsilon\} \cup \bigcup_{k=1}^m (\mathbb{R}_+ \times \mathbb{R})^k$, where ϵ is the null vector, as follows. Let $0 \leq \sigma_1 < \ldots < \sigma_\ell$ denote the distinct elements of $(s_1, \ldots, s_m)$ written in order. For $t \in [0, \sigma_1[$, $\bar{\mathbf{X}}(t) = \epsilon$. For $t \in [\sigma_1, \sigma_2[$, $\bar{\mathbf{X}}$ evolves as $\tilde{\mathbf{X}}(\cdot - \sigma_1)$ under the initial condition $\tilde{\mathbf{X}}(0) = (x_i : s_i = \sigma_1)$. Inductively, if $\bar{\mathbf{X}}(t)$ has been defined on $[0, \sigma_h[$, then for $t \in [\sigma_h, \sigma_{h+1}[$ (with the convention $\sigma_{\ell+1} = \infty$), $\bar{\mathbf{X}}$ evolves conditionally independently of $\{\bar{\mathbf{X}}(u) : u \in [0, \sigma_h[\}$ given $\bar{\mathbf{X}}(u-)$ as $\tilde{\mathbf{X}}(\cdot - \sigma_h)$ under the initial condition $\tilde{\mathbf{X}}(0) = \bar{\mathbf{X}}(u-) \cup (x_i : s_i = \sigma_h)$ (where $\cup$ denotes the operation of appending one vector to the end of the other).

The following result is a straightforward consequence of Proposition 3.1 and repeated applications of the Markov property.

Proposition 3.4. *Let* $((s_1, x_1), \ldots, (s_m, x_m))$ *and* $\bar{\mathbf{X}}$ *be as above, and let* $\mathbf{Y} = (Y_1, \ldots, Y_n)$ *be a coalescing Brownian motion starting at* $\mathbf{y} = (y_1, \ldots, y_n)$*, with* $y_1 \leq \ldots \leq y_m$*. Then, for* $t \geq \max_i s_i$*, the* $m \times (n-1)$*-dimensional random array*

$$\left(1_{[y_j, y_{j+1}]}(\bar{X}_i(t))\right)$$

has the same distribution as

$$\left(1_{[Y_j(t-s_i), Y_{j+1}(t-s_i)]}(x_i)\right).$$

Given two functions $f, g : \mathbb{R}_+ \to \mathbb{R}$ with $f(t) \leq g(t)$ for all t, let $D_t^{\rightarrow}(f, g) \subset [0, t] \times \mathbb{R}$ denote the region sandwiched between the graphs of f and g up to time t. That is,

$$D_t^{\rightarrow}(f, g) := \{(s, y) : 0 \leq s \leq t, f(s) < y < g(s)\}.$$

Let $D_t^{\leftarrow}(f, g) := \{(t - s, y) : (s, y) \in D_t(f, g)\}$ be the region $D_t^{\rightarrow}(f, g)$ time-reversed at time t. The conclusion of in Proposition 3.4 is that the random array

$$\left(1_{[y_j, y_{j+1}]}(\tilde{X}_i(t))\right)$$

has the same distribution as the random array

$$\left(1\{(s_i, x_i) \in D_t^{\leftarrow}(Y_j, Y_{j+1})\}\right).$$

4. Coalescing Brownian motion with Poisson migration

In this section we are going to study a particle system which can be described intuitively as follows. Given a time-space Poisson random measure Π^+ on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $\lambda \times$ Lebesgue. Particles appear at the atoms of Π^+. Once a particle appears, it starts to move. The existing particles execute coalescing Brownian motion with possibly different initial times. Define a set-valued process S by taking S_t to be the set of locations of those particles at time $t > 0$.

The easiest way to define S formally is via the coalescing Brownian flow ϕ of Arratia [Arr79]. Here $\phi(s, t, x)$ for $s, t, x \in \mathbb{R}$ with $s \leq t$ is a collection of random variables with the properties

- the random map $(s,t,x) \mapsto \phi(s,t,x)$ is jointly measurable,
- for each s and x, the map $t \mapsto \phi(s,t,x)$, $t \geq s$, is continuous,
- for each s and t with $s \leq t$, the map $x \mapsto \phi(s,t,x)$ is non-decreasing and right-continuous,
- for $s \leq t \leq u$, $\phi(t,u,\cdot) \circ \phi(s,t,\cdot) = \phi(s,u,\cdot)$,
- for $u \in \mathbb{R}$, $(s,t,x) \mapsto \phi(s+u,t+u,x)$ has the same distribution as ϕ,
- for $x_1,\ldots,x_m \in \mathbb{R}$ the process $(\phi(0,t,x_1),\ldots,\phi(0,t,x_m))_{t\geq 0}$ has the same distribution as $\tilde{\mathbf{X}}$ started at $(x_1,\ldots,x_m)$.

We then set

$$S_t = \{\phi(s,t,x) : (s,x) \in \Pi^+,\, 0 \leq s \leq t\},$$

where we use the short-hand notation $(s,x) \in \Pi^+$ to mean that (s,x) is an atom of Π^+.

For any $b > 0$ we have that almost surely

$$S_t \cap [-b,b] = \{\phi(s,t,x) \in [-b,b] : (s,x) \in \Pi^+,\, 0 \leq s \leq t, x \in [-a,a]\}$$

for all $a > 0$ sufficiently large: in particular, S_t is almost surely a discrete set and we identify S_t interchangeably with the simple point process obtained by placing a unit mass on each point. Using this observation, conditioning on Π^+, by Proposition 3.4, and taking limits, we get the following result which characterizes the *avoidance function* and hence the distribution of S_t (see Theorem 3.3 of [Kal76]).

Proposition 4.1. *Given $y_1 < \ldots < y_{2n}$, let $\mathbf{Y}$ be a coalescing Brownian motion starting from $(y_1,\ldots,y_{2n})$. Then*

$$\mathbb{P}\left\{S_t \cap \bigcup_{j=1}^{n}]y_{2j-1}, y_{2j}] = \emptyset\right\} = \mathbb{P}\left[\exp\left(-\lambda \sum_{j=1}^{n} \int_0^t Y_{2j}(s) - Y_{2j-1}(s) ds\right)\right].$$

We can re-phrase Proposition 4.1 as follows. For fixed $a_0 \in \mathbb{R}$, the function

$$b \mapsto |D_t^{\rightarrow}(\phi(0,\cdot,b),\phi(0,\cdot,a_0))| = \int_0^t \phi(0,s,b) - \phi(0,s,a_0)\, ds, \quad b \geq a_0,$$

where $|\cdot|$ denotes Lebesgue measure in the plane, is non-negative, non-decreasing, and right-continuous. It follows that there is a unique random Radon measure M_t on $\mathbb{R}$ such that

$$M_t(]a,b]) = \int_0^t \phi(0,s,b) - \phi(0,s,a)\, ds, \quad b \geq a.$$

Proposition 4.1 then says that S_t is the simple point process obtained by placing a unit mass at each atom of the Cox process with the random intensity measure λM_t; that is, conditional on $M_t = m$, S_t is distributed as the random measure which places a unit mass at each atom of a Poisson process with intensity measure λm. Note that M_t has atoms, and so the resulting Cox process will not be a simple point process; that is, it can have atoms with mass greater than one. Consequently, S_t is not a Cox process.

There is a unique random Radon measure M_∞ on $\mathbb{R}$ such that

$$\begin{aligned} M_\infty(]a,b]) &= \int_0^\infty \phi(0,s,b) - \phi(0,s,a)\, ds \\ &= \lim_{t\to\infty} M_t(]a,b]) \\ &= \sup_{t\geq 0} M_t(]a,b]), \quad b \geq a, \end{aligned}$$

(the finiteness of $M_\infty(]a,b])$ is assured by the continuity of $s \mapsto \phi(0,s,a)$ and $s \mapsto \phi(0,s,b)$ and the fact that $\phi(0,s,a) = \phi(0,s,b)$ for all s sufficiently large). Hence S_t converges in distribution as $t \to \infty$ to the simple point process obtained by placing a unit mass at each atom of the Cox process with the random intensity measure λM_∞. As with S_t, the point process S_∞ is not a Cox process.

We can give an almost sure construction of S_∞ as follows. Consider a Poisson random measure Π^- on $\mathbb{R}_- \times \mathbb{R}$ with intensity measure $\lambda \times$ Lebesgue. Then S_t has the same distribution as

$$\{\phi(s,0,x) : (s,x) \in \Pi^-,\ -t \leq s \leq 0\},$$

and so S_∞ has the same distribution as

$$\{\phi(s,0,x) : (s,x) \in \Pi^-,\ -\infty < s \leq 0\}.$$

We can do some explicit computations for S_∞. In what follows, let Ai denote the Airy function – see [AS72] for its definition and related properties.

Proposition 4.2. *For $a < b$,*

$$\mathbb{P}\{S_\infty \cap]a,b] = \emptyset\} = \frac{\mathrm{Ai}\left(\lambda^{\frac{1}{3}}(b-a)\right)}{\mathrm{Ai}(0)}$$

and

$$\mathbb{P}[\#\, S_\infty \cap]a,b]] = (3\lambda)^{\frac{1}{3}} \frac{\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})}(b-a).$$

Proof. Let (Y_1, Y_2) be a two-dimensional coalescing Brownian motion starting at (a,b). Then $\frac{1}{\sqrt{2}}(Y_2 - Y_1)$ is a Brownian motion stopped at 0. By Theorem 1 in [Lef89], Theorem 1 in [Lac93], or Proposition 5.14 in [GS79], we have

$$\begin{aligned} \mathbb{P}\{S_\infty \cap]a,b] = \emptyset\} &= \mathbb{P}\left[\exp(-\lambda M_\infty(]a,b]))\right] \\ &= \mathbb{P}\left[\exp\left(-\lambda \int_0^\infty Y_2(s) - Y_1(s)\, ds\right)\right] \\ &= \frac{\mathrm{Ai}\left(\lambda^{\frac{1}{3}}(b-a)\right)}{\mathrm{Ai}(0)}. \end{aligned}$$

Note that

$$\begin{aligned}\lim_{d-c\downarrow 0}\frac{\mathbb{P}\{S_\infty\cap]c,d]\neq\emptyset\}}{d-c} &= -\frac{d}{dx}\frac{\mathrm{Ai}\left(\lambda^{\frac{1}{3}}x\right)}{\mathrm{Ai}(0)}\Bigg|_{x=0}\\ &= -\lambda^{\frac{1}{3}}\frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}\\ &= (3\lambda)^{\frac{1}{3}}\frac{\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{P}[\#\, S_\infty\cap]a,b]] &= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{P}\left\{S_\infty\cap\left]a+\frac{(i-1)(b-a)}{n},a+\frac{i(b-a)}{n}\right]\neq\emptyset\right\}\\ &= (3\lambda)^{\frac{1}{3}}\frac{\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})}(b-a).\end{aligned}$$

□

References

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

[AS72] M. Abramowitz and L. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Wiley, New York, 1972.

[DEF+00] P. Donnelly, S. N. Evans, F. Fleischmann, T. G. Kurtz, and Z. Zhou. Continuum-sites stepping-stone models, coalescing exchangeable partitions, and random trees. *Ann. Probab.*, 28:1063–1110, 2000.

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

[Gri79] D. Griffeath. *Additive and cancellative interacting particle systems*, volume 724 of *Lecture Notes in Mathematics.* Springer, Berlin, 1979.

[GS79] R. K. Getoor and M. J. Sharpe. Excursions of Brownian motion and Bessel processes. *Z. Wahrsch. Verw. Gebiete*, 47(1):83–106, 1979.

[Har84] T. E. Harris. Coalescing and noncoalescing stochastic flows in $\mathbb{R}_1$. *Stoch. Proc. Appl.*, 17:187–210, 1984.

[Kal76] O. Kallenberg. *Random Measures.* Academic Press, New York, 1976.

[Lac93] A. Lachal. L'intégrale du mouvement brownien. *J. Appl. Prob.*, 30:17–27, 1993.

[Lef89] M. Lefebvre. First-passage densities of a two-dimensional process. *SIAM J. Appl. Math.*, 49(5):1514–1523, 1989.

[Lig99] T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1999.

[LJR02] Y. Le Jan and O. Raimond. Integration of Brownian vector fields. *Ann. Probab.*, 30(2):826–873, 2002.

[LJR03] Y. Le Jan and O. Raimond. Flots de noyaux et flots coalescents. *C. R. Math. Acad. Sci. Paris*, 336(2):181–184, 2003.

[STW00] F. Soucaliuc, B. Tóth, and W. Werner. Reflection and coalescence between independent one-dimensional brownian paths. *Ann. Inst. H. Poincaré Probab.*, 36:509–545, 2000.

[Tsi98] B. Tsirelson. Within and beyond the reach of Brownian innovation. In *Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998)*, number Extra Vol. III, pages 311–320 (electronic), 1998.

[Tsi04] B. Tsirelson. Scaling limit, noise, stability. In J. Picard, editor, *Lectures on probability theory and statistics : Ecole d'été de probabilites de Saint-Flour XXXII-2002*, volume 1840, pages 1–106. Springer, Berlin, 2004.

[TW97] B. Tóth and W. Werner. The true self–repelling motion. *Probab. Theory Related Fields*, 111:375–452, 1997.

[War02] J. Warren. The noise made by a Poisson snake. *Electron. J. Probab.*, 7:no. 21, 21 pp. (electronic), 2002.

[XZ] J. Xiong and X. Zhou. On the duality between coalescing brownian motions. To appear in *Can. J. Math.*

[Zho03] X. Zhou. Clustering behavior of continuous-site stepping-stone model with brownian migration. *Elec. J. Probab.*, 8(11):1–15, 2003.

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: xzhou@mathstat.concordia.ca

Department of Mathematics and Statistics, Concordia University, 7141 Sherbrooke Street West, Montreal, Quebec, H4B 1R6, CANADA