Electronic Journal of Probability

Vol. 8 (2003) Paper no. 11, pages 1–15.

Journal URL
http://www.math.washington.edu/~ejpecp/

Clustering behavior of a continuous-sites stepping-stone model with Brownian migration

Xiaowen Zhou[1]
Department of Mathematics and Statistics, Concordia University
Montreal, Quebec, H4B 1R6, Canada
zhou@alcor.concordia.ca

Abstract: Clustering behavior is studied for a continuous-sites stepping-stone model with Brownian migration. It is shown that, if the model starts with the same mixture of different types of individuals over each site, then it will evolve in a way such that the site space is divided into disjoint intervals where only one type of individuals appear in each interval. Those intervals (clusters) are growing as time $t \to \infty$. The average size of the clusters at a fixed time $t > 0$ is of the order of $\sqrt{t}$. Clusters at different times or sites are asymptotically independent as the difference of either the times or the sites goes to infinity.

Key words: stepping-stone model, clustering, coalescing Brownian motion.

AMS Subject Classification (2000): Primary: 60G17. Secondary:60J25, 60K35.

Submitted to EJP on November 6, 2002. Final version accepted on July 1, 2003.

[1]Supported by NSERC grant 253124 and a startup grant at Concordia University.

1. Introduction

Stepping-stone models were first proposed by Kimura [9] as stochastic models in population genetics. Discrete-sites stepping-stone models describe the simultaneous evolutions of populations at different colonies, where it undergoes mutation, selection and resampling within each colony and migration among those colonies. They have been studied since by different authors (see Handa [8] and Sawyer [13]). Similar models (interacting Fleming-Viot models) were considered by Dawson, Greven and Vaillancourt [3]. Very loosely put, these models can be thought as collections of Fisher-Wright models or Fleming-Viot models with geographical structures. There is one model at each colony. Different populations interact with each other via migrations among colonies. Results on long-term behaviors of such models were obtained in [3, 8].

Continuous-sites stepping-stone models with two types of individuals were first introduced in Shiga [14]. Cluster formation of such models was considered by Evans and Fleischmann [6] for a particular class of sites, namely, the continuous hierarchical group. Another continuous-sites stepping-stone model with infinitely many types was defined and discussed by Evans [5]. Further properties of this model can be found in Donnelly et al. [4]. Duality plays an important role in these studies.

Clustering is a phenomenon observed among such models, namely, individuals over sites close to each other tend to have the same type. In models with hierarchically structured site space the cluster formation was discussed by Fleischmann and Greven [7], Evans and Fleischmann [6] and Klenke [10] through studying the time-site scaling of the original models. In this paper we will focus on the infinitely many types stepping-stone model over the real line. Using the scaling property for stable processes, Evans [5] (also see [4]) showed that if the migration process is a stable process with index $1 < \alpha \leq 2$, then there is only one type of individual appearing over each site as soon as time $t > 0$. In this paper we point out that the above mentioned phenomenon can actually occur across an interval. i.e. the system clusters. We call such an interval a cluster with a certain type.

When the migration is Brownian motion and the initial state of the model consists of the same mixture of different types of individuals over each site, the evolution of the clusters can be intuitively described as follows: If we start with the same mixture of different types of individuals over each site, then clustering happens across the site space simultaneously as soon as $t > 0$. The site space is divided into intervals where there is only one type of individuals over each interval. As time goes on, the clusters are getting bigger and bigger in size. The average size of those clusters is of the order of $\sqrt{t}$ at any *fixed time* t. The types of two clusters are asymptotically independent if they are separated by either a long distance or a long time. Those results are obtained by the moment duality and analysis of the dual process, the coalescing Brownian motion. If the initial mixing measure is diffuse, sharp results can be obtained. We remark that in this model the clustering phenomenon occurs in a clean-cut fashion in contrast to those in [7, 6, 10] where the clustering is described indirectly via scaled processes.

The clustering behavior described in Theorem 3.7 resembles the one in multi-type nearest neighbor voter models over the one-dimensional lattice $\mathbb{Z}$ (see Liggett [11] for an account on the two-type case). This suggests that the continuous-sites stepping-stone model should arise as an appropriate time-space scaling limit of voter models. We refer the reader to Mueller and Tribe [12] and Cox, Durrett and Perkins [1] for work along this line.

The rest of the paper is organized as follows. We first briefly introduce the setup and the moment duality of a continuous-sites stepping-stone model in Section 2. Then in Section 3 we apply the moment duality along with a result on coalescing Brownian motion flow to study the dynamics of cluster formation in this model. In Section 4 we prove a duality formula involving joint moments over different times, which will be used later to investigate the relationship between the types of clusters at different locations and different times.

2. Definitions and preliminary results

We first sketch the setup of an infinitely-many-types continuous-sites stepping-stone model X with Brownian migration.

Let real line $\mathbb{R}$ be the *site-space*. m denotes the Lebesgue measure on $\mathbb{R}$. Let $K := [0,1]$. We identify K with the coin-tossing space $\{0,1\}^{\mathbb{N}}$. K equipped with the product topology serves as the *type-space* of X. Evans later points out that the above-mentioned topology could also be replaced by the usual topology on $[0,1]$. Write $M(K)$ for the Banach space of finite signed measures on K equipped with the total variation norm $\|\cdot\|_{M(K)}$. Let $L^\infty(m, M(K))$ denote the Banach space of (equivalence classes of) maps $\mu : \mathbb{R} \to M(K)$ such that $\operatorname{ess\,sup}\{\|\mu(e)\|_{M(K)} : e \in \mathbb{R}\} < \infty$. Write $C(K)$ for the Banach space of continuous functions on K equipped with the usual supremum norm $\|\cdot\|_{C(K)}$. To simplify notations we always write $m(de)$ for de. Let $L^1(m, C(K))$ denote the Banach space of (equivalence classes of) maps $\mu : \mathbb{R} \to C(K)$ such that $\int \|\mu(e)\|_{C(K)} de < \infty$. Then $L^\infty(m, M(K))$ is isometric to a closed subspace of the dual of $L^1(m, C(K))$ under the pairing $(\mu, x) \mapsto \int \langle \mu(e), x(e)\rangle\, de$, $\mu \in L^\infty(m, M(K))$, $x \in L^1(m, C(K))$. 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 $M_1(K)$. Ξ equipped with the relative weak* topology is a compact, metrizable space. It serves as the state space of X.

The intuitive interpretation is that $\mu \in \Xi$ describes the relative frequencies of different populations at the various sites: $\mu(e)(L)$ is the "proportion of the population at site $e \in \mathbb{R}$ that has a type belonging to the set $L \subset K$".

More elaborate discussions on the set up of such processes can be found in [5].

The *nth moment* of $\mu \in \Xi$ corresponding to $\phi \in L^1(m^{\otimes n}, C(K^n))$ is defined as follows.

Definition 2.1. 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

$$\begin{aligned} I_n(\mu;\phi) &:= \int_{\mathbb{R}^n} \left\langle \bigotimes_{i=1}^n \mu(e_i), \phi(\mathbf{e}) \right\rangle d\mathbf{e} \\ &= \int_{\mathbb{R}^n} d\mathbf{e} \int_{K^n} \phi(\mathbf{e})(\mathbf{k}) \bigotimes_{i=1}^n \mu(e_i)(dk_i), \quad \mu \in \Xi. \end{aligned} \tag{2.1}$$

Write I for I_1.

Now we are going to define *coalescing Brownian motion* which is dual to the stepping-stone model we are interested in. Coalescing Brownian motion is a system of indexed one-dimensional interacting Brownian motions with the following intuitive description. All the processes evolve as independent Brownian motions until two of them first meet. After this moment, which we call a *coalescing time*, the process with higher index assumes the value of the process with lower index. We say the process with higher index is *attached* to the one with lower index which is still *free*. They move together according to a single

Brownian motion independent of the others until the next coalescing time. The system then evolves in the same fashion.

To keep track of the interactions within the coalescing system we have to introduce more notations. Given a positive integer n, let $\mathcal{P}_n$ denote the set of *partitions* of $\mathbb{N}_n := \{1, \dots, n\}$. That is, an element π of $\mathcal{P}_n$ is a collection $\pi = \{A_1(\pi), \dots, A_h(\pi)\}$ of disjoint subsets of $\mathbb{N}_n$ such that $\bigcup_i A_i(\pi) = \mathbb{N}_n$. The sets $A_1(\pi), \dots A_h(\pi)$ are the *blocks* of the partition π. The integer h is called the *length* of π and is denoted by $l(\pi)$. For convenience we always suppose that the blocks are indexed such that $\min A_i(\pi) < \min A_j(\pi)$ for $i < j$, i.e. they are indexed according to the order of their smallest elements. Equivalently, we can think of $\mathcal{P}_n$ 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 $\pi \in \mathcal{P}_n$.

Given $\pi \in \mathcal{P}_n$, Let

$$a_\pi(i) := \min\{j : j \sim_\pi i, 1 \leq j \leq n\},\ 1 \leq i \leq n.$$

Given $i \geq 1$, let

$$a_i(\pi) := \min A_i(\pi),\ \pi \in \mathcal{P}_n, l(\pi) \geq i.$$

What we really mean by $\mathbb{N}_n$ is that it is the collection of indices of all the processes in a coalescing system. A partition π describes the interaction in the system at a fixed time. Each block in π corresponds to a free process. The block consists of the index of that free process together with the indices of all the other processes attached to it. $a_\pi(i)$ is just the index of the free process to which the ith process is attached. $a_i(\pi), i = 1, \dots, l(\pi)$, are all the indices of the free processes left.

For $\pi' \in \mathcal{P}_n$, write $\pi \prec \pi'$ or $\pi' \succ \pi$ if π' is obtained by merging some of the blocks in π. Write $\pi \preceq \pi'(\pi' \succeq \pi)$ if $\pi \prec \pi'(\pi' \succ \pi)$ or $\pi' = \pi$.

Given $\pi \in \mathcal{P}_n$, we can define a $l(\pi)$-dimensional subspace $\mathbb{R}^n_\pi$ of $\mathbb{R}^n$ by identifying the coordinates with indices from the same block of π. More specifically,

$$\mathbb{R}^n_\pi := \{(x_{a_\pi(1)}, \dots, x_{a_\pi(n)}) : x_{a_\pi(i)} \in \mathbb{R}, 1 \leq i \leq n\}.$$

Put

$$\check{\mathbb{R}}^n_\pi := \mathbb{R}^n_\pi \setminus \bigcup_{\pi' \succ \pi, l(\pi') = l(\pi)-1} \mathbb{R}^n_{\pi'}.$$

$\check{\mathbb{R}}^n_\pi$ is just the effective state space of the coalescing system when the interaction is represented by π. Note that $\check{\mathbb{R}}^n_\pi$ and $\check{\mathbb{R}}^n_{\pi'}$ are disjoint for $\pi \neq \pi'$.

More precisely, let $\mathbf{W}^\mathbf{e} = (W_1, \dots, W_n)$ be a n-dimensional Brownian motion starting from $\mathbf{e} \in \mathbb{R}^n$. The n-dimensional coalescing Brownian motion $\check{\mathbf{W}}^\mathbf{e} = (\check{W}_1, \dots, \check{W}_n)$ can be constructed from $\mathbf{W}^\mathbf{e}$ inductively as follows. Suppose that times $0 =: \tau_0 \leq \dots \leq \tau_k \leq \infty$ and partitions $\{\{1\}, \dots, \{n\}\} =: \pi_0 \prec \dots \prec \pi_k \preceq \{\{1, \dots, n\}\}$ have already been defined and $\check{\mathbf{W}}^\mathbf{e}$ has been defined on $[0, \tau_k)$. If $\pi_k = \{\{1, \dots, n\}\}$, then $\check{\mathbf{W}}^\mathbf{e}_t = (W_1(t), \dots, W_1(t))$ for $t \geq \tau_k$. Otherwise, let $\pi_k = \{A_1(\pi_k), \dots, A_{l(\pi_k)}(\pi_k)\}$. Put

$$\tau_{k+1} := \inf\{t > \tau_k : \exists i < j, W_{a_i(\pi_k)}(t) = W_{a_j(\pi_k)}(t)\}. \tag{2.2}$$

Suppose that $W_{a_i(\pi_k)}(\tau_{k+1}) = W_{a_j(\pi_k)}(\tau_{k+1})$ for some $1 \leq i < j \leq l(\pi_k)$, then define

$$\pi_{k+1} := \{A_1(\pi_{k+1}), \dots, A_{l(\pi_k)-1}(\pi_{k+1})\}, \tag{2.3}$$

where

$$A_r(\pi_{k+1}) := \begin{cases} A_r(\pi_k), & \text{for } 1 \le r < i \text{ or } i < r < j, \\ A_i(\pi_k) \cup A_j(\pi_k), & \text{for } r = i, \\ A_{r+1}(\pi_k), & \text{for } j \le r \le l(\pi_k) - 1 \end{cases}$$

and $\check{\mathbf{W}}^{\mathbf{e}}(t) := (W_{a_{\pi_k}(1)}(t), \ldots, W_{a_{\pi_k}(n)}(t))$ for $\tau_k \le t < \tau_{k+1}$.

Theorem 2.2 was first obtained in [5]. It will be used repeatedly in the present paper.

Theorem 2.2. *(Moment duality) 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*

$$\begin{aligned} &\int Q_t(\mu, d\nu) I_n(\nu; \phi) \\ &= \sum_{\pi \in \mathcal{P}_n} \int_{\mathbb{R}^n} \mathbb{P}\left[1_{\{\check{\mathbf{W}}_t^{\mathbf{e}} \in \check{\mathbb{R}}_\pi^n\}} \int \bigotimes_{i=1}^{l(\pi)} \mu(\check{W}^{\mathbf{e}}_{a_i(\pi)}(t))(dk_{a_i(\pi)}) \phi(\mathbf{e})(k_{a_\pi(1)}, \ldots, k_{a_\pi(n)}) \right] d\mathbf{e}. \end{aligned} \tag{2.4}$$

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

Remark 2.3. The duality formula (2.4) doesn't have exactly the same expression as that in Theorem 4.1 of [5]. But one can easily check that they turn out to be the same.

Because coalescing Brownian motion is dual to the stepping-stone model, it plays a crucial role in analyzing the clustering behavior. We first introduce two results on a system of coalescing Brownian motions. Given $a < b$, let $\check{\mathbf{W}}^{a,b,n} := (\check{W}_1^{a,b,n}, \ldots, \check{W}_n^{a,b,n})$ be a collection of coalescing Brownian motions such that the initial values $\check{\mathbf{W}}^{a,b,n}(0) = (\check{W}_n^{a,b,n}(0), \ldots, \check{W}_n^{a,b,n}(0))$ are independent and uniformly distributed over interval $[a, b]$. We can define $\check{\mathbf{W}}^{a,b,n}, n = 1, 2, \ldots,$ on the same probability space in such a way that $\cup_{i=1}^n \{\check{\mathbf{W}}_i^{a,b,n}(t)\} \subset \cup_{i=1}^{n+1} \{\check{\mathbf{W}}_i^{a,b,n+1}(t)\}$ for all $n = 1, 2, \ldots$ and $t \ge 0$. Set $\check{\mathbf{W}}^{a,b}(t) := \cup_{n=1}^\infty \cup_{i=1}^n \{\check{\mathbf{W}}_i^{a,b,n}(t)\}$. Write $|\check{\mathbf{W}}^{a,b}(t)|$ for the cardinality of the collection of coordinates of $\check{\mathbf{W}}^{a,b}(t)$, i.e. the total number of "free" Brownian motions left in the coalescing system $\check{\mathbf{W}}^{a,b,n}$ by time t. Write $\check{\mathbf{W}}^a$ for $\check{\mathbf{W}}^{-a,a}$.

Lemma 2.4 was first obtained in [16]. It plays a key role in analyzing the cluster formation and the sizes of those clusters.

Lemma 2.4. $\mathbb{P}[|\check{\mathbf{W}}^a(t)|] = 1 + \frac{2a}{\sqrt{\pi t}}$.

Since $\mathbb{P}\{|\check{\mathbf{W}}^a(t)| \ge 2\}$ is equal to the probability that two independent Brownian motions, with initial values $-a$ and a respectively, have not met until time t. The next result is a consequence of reflection principle of Brownian motion.

Lemma 2.5.

$$\mathbb{P}\{|\check{\mathbf{W}}^a(t)| \ge 2\} = \frac{1}{\sqrt{2\pi t}} \int_0^\infty \exp\left(-\frac{(x - \sqrt{2}a)^2}{2t}\right) - \exp\left(-\frac{(x + \sqrt{2}a)^2}{2t}\right) dx.$$

3. Clustering of the continuous-sites Stepping-stone model

For $\theta \in M_1(K)$, write $\theta^{\mathbb{R}}$ for the element in Ξ such that $\theta^{\mathbb{R}}(e) = \theta$ for m a.e. $e \in \mathbb{R}$. $\delta_{\{k\}}$ denotes the point mass at $k \in K$. $\delta_{\delta_{\{k\}}^{\mathbb{R}}}$ is the point mass at $\delta_{\{k\}}^{\mathbb{R}} \in \Xi$. Then

$\int \delta_{\delta_{\{k\}}^{\mathbb{R}}}\theta(dk) \in M_1(\Xi)$ means that with probability $\theta(dk)$ only individuals of type k appear over the site space $\mathbb{R}$. Write P_t for the transition semigroup of one-dimensional Brownian motion.

Theorem 3.1. *Given $\mathcal{M} \in M_1(\Xi)$ and $\mathcal{M}^* = \int \delta_{\delta_{\{k\}}^{\mathbb{R}}}\theta(dk)$ for some $\theta \in M_1(K)$, then*

$$\lim_{t\to\infty} \mathcal{M}Q_t = \mathcal{M}^* \tag{3.1}$$

if and only if for any $\chi \in C(K)$,

$$\lim_{t\to\infty} \int \langle P_t\mu(e), \chi\rangle \mathcal{M}(d\mu) = \langle \theta, \chi\rangle \tag{3.2}$$

for m a.e. $e \in \mathbb{R}$.

Proof. The necessity of (3.2) follow readily from the moment duality formula (2.4). We only need to show that (3.2) is sufficient. Write $\mathbf{e}$ for $(e_1, ..., e_n)$. Write $\psi \otimes \chi$ for the tensor product of $\psi \in L^1(m^{\otimes n}) \cap C(\mathbb{R}^n)$ and $\chi \in C(K^n)$. Observe that in an n-dimensional coalescing Brownian motion, there will be only one free Brownian motion (with index 1) left eventually. It follows from moment duality (2.4) that

$$\begin{aligned}
&\lim_{t\to\infty} \int \mathcal{M}Q_t(d\mu) I_n(\mu; \psi \otimes \chi) \\
&= \lim_{t\to\infty} \int \mathcal{M}(d\mu) \int \psi(\mathbf{e}) \mathbb{P}\left[\int \chi(k, ..., k)\mu(\check{W}_1^{\mathbf{e}}(t))(dk)\right] d\mathbf{e} \\
&= \int \psi(\mathbf{e}) \langle \theta, \chi(k, ..., k)\rangle d\mathbf{e} \\
&= \int I_n(\delta_{\delta_{\{k\}}^{\mathbb{R}}}; \psi \otimes \chi)\theta(dk) \\
&= \left(\int \delta_{\delta_{\{k\}}^{\mathbb{R}}}\theta(dk)\right) I_n(\cdot; \psi \otimes \chi).
\end{aligned} \tag{3.3}$$

Hence, (3.1) follows from Lemma 3.1 in [5]. □

Remark 3.2. By Theorem 3.1, if the initial value of X is $\theta^{\mathbb{R}}, \theta \in M_1(K)$, i.e. X starts with the same mixture of individuals over each site, then $\theta^{\mathbb{R}}Q_t \to \int \delta_{\delta_{\{k\}}^{\mathbb{R}}}\theta(dk)$. As a result, we certainly expect that individuals of the same type clump together.

For $\mu \in \Xi$, define the *block average* $\mu_{[a,b]} \in M_1(K)$ of μ on $[a, b] \subset \mathbb{R}$ as

$$\mu_{[a,b]} := \frac{1}{b-a}\int_a^b \mu(e)de.$$

Notice that given $G \subset K$, $\mu_{[a,b]}(G) = 1$ if and only if $\mu_x(G) = 1$ for m a.e. $x \in [a, b]$ and if and only if $\mu_{[a',b']}(G) = 1$ for $a \le a' \le b' \le b$.

Let $G_1, \ldots, G_d$ be a partition of K, i.e. $\cup_{i=1}^d G_i = K$ and $G_i \cap G_j = \emptyset, i \ne j$.

Theorem 3.3. *Given $x \in \mathbb{R}$ and $t > 0$, $Q^{\theta^{\mathbb{R}}}$ almost surely, there exists a constant $A > 0$ and $1 \le i \le d$ such that $X_t(y)(G_i) = 1$ for m a.e. $y \in (x - A, x + A)$.*

Given $x \in \mathbb{R}$ and $a > 0$, $Q^{\theta^{\mathbb{R}}}$ almost surely, there exists a time $T > 0$ and $1 \le i \le d$ such that $X_T(y)(G_i) = 1$ for m a.e. $y \in (x - a, x + a)$.

Proof. Suppose that $x = 0$. For any $1 \le i \le d, a > 0$ and positive integer n, apply moment duality (2.4), we have

$$Q^{\theta^{\mathbb{R}}}\left[X_{t[-a,a]}(G_i)^n\right] = \int I_n(\nu; \frac{1}{2a} 1_{[-a,a]}^{\otimes n} \otimes 1_{G_i}^{\otimes n}) Q_t(\theta^{\mathbb{R}}, d\nu) = \mathbb{P}\left[\theta_i^{|W^{a,n}(t)|}\right],$$

where $X_{t[-a,a]}$ denotes the block average of X_t and $\theta_i := \mathbb{P}\{G_i\}$. It follows from $X_{t[-a,a]}(G_i)^n \to 1_{\{X_{t[-a,a]}(G_k)=1\}}$, $n \to \infty$, that

$$\begin{aligned} & Q^{\theta^{\mathbb{R}}}\{X_t(y)(G_i) = 1 \text{ for } m \text{ a.e. } y \in (x-a, x+a)\} \\ & = Q^{\theta^{\mathbb{R}}}\{X_{t[-a,-a]}(G_i) = 1\} \\ & = \lim_{n\to\infty} Q^{\theta^{\mathbb{R}}}\left[X_{t[-a,a]}(G_i)^n\right] \\ & = \mathbb{P}\left[\theta_i^{|W^a(t)|}\right]. \end{aligned} \tag{3.4}$$

By Lemma 2.4, for fixed $t > 0$, $|W^a(t)| \to 1$ in probability as $a \to 0+$. In addition, for fixed $a > 0$, $|W^a(t)| \to 1$ in probability as $t \to \infty$. Then

$$\begin{aligned} & \lim_{a\to 0+} Q^{\theta^{\mathbb{R}}}\{X_t(y)(G_i) = 1 \text{ for } m \text{ a.e. } y \in (x-a, x+a)\} \\ & = \lim_{a\to 0+} \mathbb{P}\left[\theta_i^{|W^a(t)|}\right] \\ & = \theta_i \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} & \lim_{t\to\infty} Q^{\mu^{\theta}}\{X_t(y)(G_i) = 1 \text{ for } m \text{ a.e. } y \in (x-a, x+a)\} \\ & = \lim_{t\to\infty} \mathbb{P}\left[\theta_i^{|W^a(t)|}\right] \\ & = \theta_i. \end{aligned}$$

Notice that $\sum_{i=1}^d \theta_i = 1$, the assertion in this theorem is verified. □

Remark 3.4. It seems the initial value $\theta^{\mathbb{R}}$ is necessary to obtain the desired result in Theorem 3.3. This is similar to the study on voter models where a typical initial distribution is a renewal measure. Also notice the similar requirements on the initial values of related models in [7, 6, 10].

Given a partition $\{G_1, \ldots, G_d\}$ of K, we say $\mu \in \Xi$ has a *cluster* of type G_i over the interval (u, v) if $\mu(.)(G_i) = 1$ m a.e. on (u, v). The length of this cluster is just the length of the largest interval containing (c, d) such that μ has a cluster of type G_i on it. For any $M > 0$ and $a > 0$, let $N_{M,a}(\mu)$ be the total number of different clusters of μ over the interval $[-M, M]$ with length greater than a. Let $L_{M,a}(\mu)$ be the summation of lengths of those clusters of μ on $[-M, M]$ with lengths greater than a.

The following lemma says that clustering occurs not only locally, as described in Theorem 3.3, but also across the interval $[-M, M]$ at time $t > 0$.

Lemma 3.5. *For any $M > 0$ and $t > 0$, $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely,*

$$\lim_{a\to 0+} \frac{L_{M,a}(X_t)}{2M} = 1. \tag{3.6}$$

Proof. Set $g_{x,a}(\mu) := 1_V(\mu), \mu \in M_1(\Xi), V := \bigcup_{i=1}^d \{\nu \in \Xi : \nu_{[x,x+a]}(G_i) = 1\}$. Then

$$\begin{aligned}
\mathbb{Q}^{\theta^{\mathbb{R}}}\left[\int_{-M}^{M} g_{x,a}(X_t)dx\right] &= \int_{-M}^{M} \mathbb{Q}^{\theta^{\mathbb{R}}}\left\{\bigcup_{i=1}^{\infty}\{X_{t[x,x+a]}(G_i) = 1\right\} dx \\
&= \int_{-M}^{M} \sum_{i=1}^{\infty} \mathbb{Q}^{\theta^{\mathbb{R}}}\{X_{t[x,x+a]}(G_i) = 1\} dx \\
&= 2M \sum_{i=1}^{d} \mathbb{P}[\theta_i^{|W^{\frac{a}{2}}(t)|}].
\end{aligned} \tag{3.7}$$

Since

$$\int_{-M}^{M} g_{x,a}(X_t)dx \le L_{M,a}(X_t) \le 2M, a > 0,$$

and

$$\lim_{a\to 0+} \mathbb{Q}^{\theta^{\mathbb{R}}}\left[\int_{-M}^{M} g_{x,a}(X_t)dx\right] = 2M,$$

then (3.6) holds. □

Let $N_M(\mu) := \lim_{a\to 0+} N_{M,a}(\mu), \mu \in \Xi$, be the total number of clusters in μ over $[-M, M]$. The next result shows that clustering happens simultaneously over $\mathbb{R}$. It also gives an estimate on the average size of the clusters at a fixed time $t > 0$. Notice that Lemma 3.5 does not exclude the possibility that $N_M(X_t)$ could be infinite.

Denote by $\mathbb{Q}_t^{\theta^{\mathbb{R}}}$ the distribution of X_t under $\mathbb{Q}^{\theta^{\mathbb{R}}}$. Given $x \in \mathbb{R}$, define a shifting operator τ_x on Ξ by $\tau_x(\mu) := \mu(x + .)$, $\mu \in \Xi$. τ_x can induce a shift operator (also denoted by τ_x) on $C(\Xi)$ and $M_1(\Xi)$ by

$$\tau_x\Phi(\mu) := \Phi(\tau_x\mu), \Phi \in C(\Xi), \mu \in \Xi,$$

and

$$(\tau_x\mathbb{Q})\Phi := \mathbb{Q}(\tau_x\Phi), \mathbb{Q} \in M_1(\Xi), \Phi \in C(\Xi).$$

We refer to [15] for the definition of *strong mixing* and other results concerning ergodic theory.

Lemma 3.6. *For any $t > 0, \mathbb{Q}_t^{\theta^{\mathbb{R}}}$ is strong mixing. Therefore, it is ergodic with respect to τ_x.*

Proof. The initial value $\theta^{\mathbb{R}}$ is shifting-invariant. Then $\tau_x\mathbb{Q}_t^{\theta^{\mathbb{R}}} = \mathbb{Q}_t^{\theta^{\mathbb{R}}}, x \in \mathbb{R}$, follows from the moment duality (2.4). τ_x is thus measure preserving. By Lemma 3.1 in [5] we need to show that for any $n_1, n_2 \in \mathbb{N}_+$, any $\phi_1 = \psi_1 \otimes \chi_1, \psi_1 \in L^1(m^{\otimes n_1}) \cap C(\mathbb{R}^{n_1}), \chi_1 \in C(K^{n_1})$, and any $\phi_2 = \psi_2 \otimes \chi_2, \psi_2 \in L^1(m^{\otimes n_2}) \cap C(\mathbb{R}^{n_2}), \chi_2 \in C(K^{n_2})$, it holds that

$$\lim_{x\to\infty} \int I_{n_1}(\tau_x\mu; \phi_1) I_{n_2}(\mu; \phi_2) \mathbb{Q}_t^{\theta^{\mathbb{R}}}(d\mu) = \int I_{n_1}(\mu; \phi_1)\mathbb{Q}_t^{\theta^{\mathbb{R}}}(d\mu) \int I_{n_2}(\mu; \phi_2)\mathbb{Q}_t^{\theta^{\mathbb{R}}}(d\mu). \tag{3.8}$$

Notice that

$$\begin{aligned}
I_{n_1}(\tau_x\mu; \phi_1) &= \int_{\mathbb{R}^{n_1}} \left\langle \bigotimes_{i=1}^{n_1} \mu(x + e_i), \phi_1(\mathbf{e}) \right\rangle d\mathbf{e} \\
&= \int_{\mathbb{R}^{n_1}} \psi_1(\mathbf{e} - \mathbf{x}) \left\langle \bigotimes_{i=1}^{n_1} \mu(e_i), \chi_1 \right\rangle d\mathbf{e},
\end{aligned} \tag{3.9}$$

where $\mathbf{x} := (x, \dots, x)$, (3.8) can then be easily verified by moment duality (2.4) and the following facts. $\mathbb{P}\{T^{e_1, e_2 - x} < t\} \to 0$ as $|x| \to +\infty$, where $T^{e_1, e_2 - x} := \inf\{s > 0 : W^{e_1}(s) = W^{e_2 - x}(s)\}$. W^{e_1} and $W^{e_2 - x}$ are two independent Brownian motions starting from e_1 and $e_2 - x$ respectively. □

Theorem 3.7. *Given a partition $\{G_1, \dots, G_d\}$ of K and $t > 0$, $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely, there exists a sequence $\dots < b_{-2} < b_{-1} < b_0 < b_1 < b_2 < \dots, \lim_{n \to -\infty} b_n = -\infty, \lim_{n \to \infty} b_n = \infty$, such that X_t has a cluster of type G_i for some $1 \le i \le d$ on each (b_{j-1}, b_j) and the clusters over neighboring intervals are of different types. Moreover, there exists a constant c_t such that $\lim_{M \to \infty} \frac{2M}{N_M(X_t)} = c_t$ and $\sqrt{\pi t} \le c_t \le \frac{\sqrt{\pi t}}{1 - \sum_{i=1}^d \theta_i^2}$.*

Proof. For $a > 0$ and $\mu \in \Xi$, write $f_{x,a}(\mu) := 1_V(\mu)$, where $V := \bigcap_{i=1}^d \{\nu \in \Xi : \nu_{[x,x+a]}(G_i) \neq 1\}$. Then

$$
\begin{aligned}
\mathbb{Q}^{\theta^{\mathbb{R}}}\left[\int_{-M}^{M} f_{x,a}(X_t)dx\right] &= \int_{-M}^{M} \mathbb{Q}^{\theta^{\mathbb{R}}}\left\{\bigcap_{i=1}^{d}\{X_{t[x,x+a]}(G_i) \neq 1\right\} dx \\
&= \int_{-M}^{M} \left(1 - \sum_{i=1}^{d} \mathbb{Q}^{\theta^{\mathbb{R}}}\{X_{t[x,x+a]}(G_i) = 1\}\right) dx \qquad (3.10)\\
&= 2M\left(1 - \sum_{i=1}^{d} \mathbb{P}[\theta_i^{|W^{\frac{a}{2}}(t)|}]\right).
\end{aligned}
$$

It is easy to see that

$$
1 - \sum_{i=1}^{d} \mathbb{P}[\theta_i^{|W^{\frac{a}{2}}(t)|}] \le 1 - \sum_{i=1}^{d} \theta_i \mathbb{P}\{|W^{\frac{a}{2}}(t)| = 1\} = \mathbb{P}\{|W^{\frac{a}{2}}(t)| \ge 2\} \qquad (3.11)
$$

and

$$
\begin{aligned}
1 - \sum_{i=1}^{d} \mathbb{P}[\theta_i^{|W^{\frac{a}{2}}(t)|}] &\ge 1 - \sum_{i=1}^{d} \theta_i \mathbb{P}\{|W^{\frac{a}{2}}(t)| = 1\} - \sum_{i=1}^{d} \theta_i^2 \mathbb{P}\{|W^{\frac{a}{2}}(t)| \ge 2\} \qquad (3.12)\\
&\ge \left(1 - \sum_{i=1}^{d} \theta_i^2\right) \mathbb{P}\{|W^{\frac{a}{2}}(t)| \ge 2\}.
\end{aligned}
$$

Since $a(N_{M,a}(\mu) - 1) \le \int_{-M}^{M} f_{x,a}(\mu)dx$, then

$$
\begin{aligned}
\mathbb{Q}^{\theta^{\mathbb{R}}}[N_{M,a}(X_t)] &\le \frac{1}{a}\mathbb{Q}^{\theta^{\mathbb{R}}}\left[\int_{-M}^{M} f_{x,a}(X_t)dx\right] + 1 \qquad (3.13)\\
&\le \frac{2M}{a}\mathbb{P}\{|W^{\frac{a}{2}}(t)| \ge 2\} + 1.
\end{aligned}
$$

Let $a \to 0+$, by Lemma (2.5) we have

$$
\mathbb{Q}^{\theta^{\mathbb{R}}}[N_M(X_t)] \le \frac{2M}{\sqrt{\pi t}} + 1. \qquad (3.14)
$$

Since $N_M(X_t) < \infty$ $Q^{\theta^{\mathbb{R}}}$ a.s., this together with Lemma 3.5 imply that $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely, except on a m-null set, the interval $[-M, M]$ is divided into finite subintervals where X_t has one type over each interval. M is arbitrary, the first assertion of this theorem is thus proved.

On the other hand,

$$\mathbb{Q}^{\theta^{\mathbb{R}}}[N_M(X_t)] \geq \frac{1}{a}\mathbb{Q}^{\theta^{\mathbb{R}}}\left[\int_{-M}^{M} f_{x,a}(X_t)dx\right] \geq \frac{2M}{a}\left(1-\sum_{i=1}^{d}\theta_i^2\right)\mathbb{P}\{|W^{\frac{a}{2}}(t)| \geq 2\}.$$

Let $a \to 0+$, it follows form Lemma (2.5) again that

$$\mathbb{Q}^{\theta^{\mathbb{R}}}[N_M(X_t)] \geq \frac{2M}{\sqrt{\pi t}}\left(1-\sum_{i=1}^{d}\theta_i^2\right). \tag{3.15}$$

Write $N_m(\mu), m > 0$, for the total number of clusters of $\mu \in \Xi$ over the interval $[0, m]$. It follows from (3.13) that $\mathbb{Q}_t^{\theta^{\mathbb{R}}}[N_m] < \infty$. τ_m is measure-preserving under $\mathbb{Q}_t^{\theta^{\mathbb{R}}}$. By definition one can also verify that $N_{m+n} \leq N_m + \tau_m N_n$. Then the subadditive ergodic theorem (see Theorem 10.1 in [15]) implies that

$$\lim_{m\to\infty}\frac{N_m}{m} \text{ exists } \mathbb{Q}_t^{\theta^{\mathbb{R}}} \text{ a.s..}$$

Then

$$\lim_{m\to\infty}\frac{N_m(X_t)}{m} \text{ exists } \mathbb{Q}^{\theta^{\mathbb{R}}} \text{ a.s..}$$

Therefore,

$$\lim_{M\to\infty}\frac{N_M(X_t)}{2M} \text{ exists } \mathbb{Q}^{\theta^{\mathbb{R}}} \text{ a.s..}$$

It follows from Lemma 3.6 that, $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely, $\lim_{M\to\infty}\frac{N_M(X_t)}{2M}$ is a constant. Using the subadditive ergodic theorem again, we have

$$\frac{1}{\sqrt{\pi t}} \leq \lim_{M\to\infty}\frac{N_M(X_t)}{2M} \leq \frac{1-\sum_{i=1}^{d}\theta_i^2}{\sqrt{\pi t}} \quad \mathbb{Q}^{\theta^{\mathbb{R}}} \text{ a.s..}$$

□

Assume that the initial mixture $\theta \in M_1(K)$ is a diffuse measure, i.e. $\theta(\{k\}) = 0, k \in K$, then Theorem 3.7 can be improved.

Theorem 3.8. *Suppose that $\theta \in M_1(K)$ is a diffuse measure, then given $t > 0$, $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely, there exists a sequence $\ldots < b_{-2} < b_{-1} < b_0 < b_1 < b_2 < \ldots, \lim_{n\to-\infty} b_n = -\infty, \lim_{n\to\infty} b_n = \infty$, such that X_t has a cluster of type k for some $k \in K$ on each (b_{j-1}, b_j). Moreover, $\lim_{M\to\infty}\frac{2M}{N_M(X_t)} = \sqrt{\pi t}$.*

Proof. For each positive integer n, let $\Pi_n := \{[\frac{i}{2^n}, \frac{i+1}{2^n}) : 1 \leq i \leq 2^n - 1\}$ be a partition of K. Since Π_n is getting finer and finer as n increase, a cluster with respect to Π_n can break into new clusters with respect to Π_{n+1}. Apply Theorem 3.7 to Π_n and let $(b_i^{(n)})$ be the correspondent partition of the real line, we see that the set $\{b_i^{(n)} : -\infty < i < \infty\}$ is increasing with respect to n. By (3.14), $\cup_{n=1}^{\infty}\{b_i^{(n)} : -\infty < i < \infty\}$ has no subsequence converging to a finite limit and we choose it as the collection of those b_is in the present theorem. Since any $k_1, k_2 \in K, k_1 \neq k_2$ are separated by Π_n for n big enough, then on each interval (b_j, b_{j+1}) X_t can only have a cluster of a single type $k \in K$. The last assertion in this theorem follows from the subadditive ergodic theorem, (3.14), (3.15) and the fact that θ is diffuse. □

For $\mu \in \Xi$, define $L(\mu)$ and $U(\mu), U(\mu) > 0$ as the essential lower and upper bounds of the cluster of μ at 0. Given that θ is diffuse, we can obtain the joint distribution of $L(X_t)$ and $U(X_t)$.

Theorem 3.9. *Suppose that $\theta \in M_1(K)$ is diffuse, then for any $a > 0, b > 0$ and $t > 0$,*

$$\begin{aligned} &Q^{\theta^{\mathbb{R}}}\{L(X_t) < -a, U(X_t) > b\} \\ &= 1 - \frac{1}{\sqrt{2\pi t}} \int_0^\infty \exp\left(-\frac{(x - \frac{\sqrt{2}}{2}(a+b))^2}{2t}\right) - \exp\left(-\frac{(x + \frac{\sqrt{2}}{2}(a+b))^2}{2t}\right) dx. \end{aligned} \tag{3.16}$$

Proof. Using the same partition Π_n defined in the proof of Theorem 3.7, by (3.4) we have

$$\begin{aligned} &\sum_{G \in \Pi_n} Q^{\theta^{\mathbb{R}}}\{X_t(y) \in \delta_G \text{ for } m \text{ a.e. } y \in (-a, b)\} \\ &= \sum_{G \in \Pi_n} \mathbb{P}\left[\theta(G)^{|W^{\frac{a+b}{2}}(t)|}\right]. \end{aligned} \tag{3.17}$$

Let $n \to \infty$, since θ is diffuse, then

$$\begin{aligned} &Q^{\theta^{\mathbb{R}}}\{L(X_t) < -a, U(X_t) > b\} \\ &= \lim_{n \to \infty} \sum_{G \in \Pi_n} Q^{\theta^{\mathbb{R}}}\{X_t(y) \in \delta_G \text{ for } m \text{ a.e. } y \in (-a, b)\} \\ &= \mathbb{P}\{|W^{-a,b}(t)| = 1\} \\ &= \mathbb{P}\{|W^{\frac{a+b}{2}}(t)| = 1\}. \end{aligned} \tag{3.18}$$

So we can conclude (3.16) immediately from Lemma 2.5. □

Remark 3.10. As an consequence of Theorem 3.9, the probability that two sites a and b belong to the same cluster of the stepping-stone model at time t is also the probability that two independent Brownian motions with initial values a and b respectively meet each other before time t.

4. Correlation between types over different sites

In the rest of the paper, we focus on understanding the relationship between the types of clusters over different sites and at different times. To accomplish that we need to generalize the moment duality formula to one involving joint moments over different times.

We first define a system of coalescing Brownian motions in which the Brownian motions are allowed to have different starting times. Given $s > 0$, $\mathbf{e}_1 := (e_{11}, \ldots, e_{1n_1}) \in \mathbb{R}^{n_1}$ and $\mathbf{e}_2 := (e_{21}, \ldots, e_{2n_2}) \in \mathbb{R}^{n_2}$, an $(n_1 + n_2)$-dimensional coalescing process $\check{\mathbf{W}}^{(\mathbf{e}_1;0;\mathbf{e}_2;s)}$ is defined intuitively as follows: An n_1-dimensional process starts at time 0 from value $\mathbf{e}_1$ and evolves according to a coalescing Brownian motion, while another n_2-dimensional process is "frozen" until time s. Starting at time s from $\mathbf{e}_2$, the second process joins the first process and together they evolve according to an $(n_1 + n_2)$-dimensional process.

Lemma 4.1. *Let X be the continuous-sites stepping-stone model with Brownian motion migration, then for any $0 < t_2 < t_1, \mu \in \Xi$ and $\phi_i = L^1(m^{\otimes n_i}, C(K^{n_i})), i = 1, 2,$*

$$\tag{4.1}
\begin{aligned}
&\mathbb{Q}^{\mu}[I_{n_1}(X_{t_1};\phi_1)I_{n_2}(X_{t_2};\phi_2)] \\
&\quad = \sum_{\pi \in \mathcal{P}_{n_1+n_2}} \int_{\mathbb{R}^{n_1}} d\mathbf{e}_1 \int_{\mathbb{R}^{n_2}} d\mathbf{e}_2 \\
&\qquad \mathbb{P}\left[1_{\{\check{\mathbf{W}}^A_{t_1} \in \check{\mathbb{R}}^{n_1+n_2}_{\pi}\}} \int \phi(\mathbf{e}_1) \otimes \phi(\mathbf{e}_2)(k_{a_\pi(1)}, \ldots, k_{a_\pi(n_1+n_2)}) \bigotimes_{i=1}^{l(\pi)} \mu(\check{W}^A_{a_i(\pi)}(t_1))(dk_{a_i(\pi)})\right],
\end{aligned}$$

where $A = (\mathbf{e}_1; 0; \mathbf{e}_2; t_1 - t_2)$.

Proof. For any $\phi_i = \psi_i \otimes \chi_i, \psi_i \in L^1(m^{\otimes n_i}) \cap C(R^{n_i}), \chi_i \in C(K^{n_i}), i = 1, 2,$ the moment duality (2.4) yields that

$$\tag{4.2}
\begin{aligned}
&\mathbb{Q}^{\nu}[I_{n_1}(X_{t_1-t_2};\phi_1)] \\
&= \int I_{n_1}(\nu_1;\phi_1)Q_{t_1-t_2}(\nu, d\nu_1) \\
&= \sum_{\pi \in \mathcal{P}_{n_1}} \int_{\mathbb{R}^{n_1}} \psi_1(\mathbf{e}_1)\mathbb{P}\left[1_{\{\check{\mathbf{W}}^{\mathbf{e}_1}_{t_1-t_2} \in \check{\mathbb{R}}^{n_1}_{\pi}\}} \int \chi_1(k_{a_\pi(1)}, \ldots, k_{a_\pi(n_1)}) \bigotimes_{i=1}^{l(\pi)} \nu(\check{W}^{\mathbf{e}_1}_{a_i(\pi)}(t_1 - t_2))(dk_{a_i(\pi)})\right] d\mathbf{e}_1 \\
&= \sum_{\pi \in \mathcal{P}_{n_1}} \int_{\mathbb{R}^{n_1}} \psi_1(\mathbf{e}_1) w_{\mathbf{e}_1\pi} d\mathbf{e}_1 \int_{\mathbb{R}^{l(\pi)}} f_{\mathbf{e}_1\pi}(e_{a_1(\pi)}, \ldots, e_{a_{l(\pi)}(\pi)}) d\mathbf{e}_\pi \\
&\quad \int \bar{\chi}_\pi(k_{a_1(\pi)}, \ldots, k_{a_{l(\pi)}(\pi)}) \bigotimes_{i=1}^{l(\pi)} \nu(e_{a_i(\pi)})(dk_{a_i(\pi)}) \\
&= \sum_{\pi \in \mathcal{P}_{n_1}} I_{l(\pi)}\left(\nu; \left(\int_{\mathbb{R}^{n_1}} \psi_1(\mathbf{e}_1) w_{\mathbf{e}_1\pi} f_{\mathbf{e}_1\pi} d\mathbf{e}_1\right) \otimes \bar{\chi}_\pi\right),
\end{aligned}$$

where $w_{\mathbf{e}_1\pi} := \mathbb{P}\{\check{\mathbf{W}}^{\mathbf{e}_1}(t_1 - t_2) \in \check{\mathbb{R}}^{n_1}_{\pi}\}$, $\mathbf{e}_\pi := (e_{a_1(\pi)}, \ldots, e_{a_{l(\pi)}(\pi)})$, $\bar{\chi}_\pi(k_{a_1(\pi)}, \ldots, k_{a_{l(\pi)}(\pi)}) := \chi_1(k_{a_\pi(1)}, \ldots, k_{a_\pi(n_1)})$ and $f_{\mathbf{e}_1\pi}(e'_{a_{l(\pi)}}, \ldots, e'_{a_{l(\pi)}(\pi)})$ is the conditional density of $(\check{W}^{\mathbf{e}_1}_{a_1(\pi)}(t_1 - t_2), \ldots, \check{W}^{\mathbf{e}_1}_{a_{l(\pi)}(\pi)}(t_1 - t_2))$ given $\{\check{\mathbf{W}}^{\mathbf{e}_1}(t_1 - t_2) \in \check{\mathbb{R}}^{n_1}_{\pi}\}$. Write $\check{\mathbf{W}}^{\mathbf{e}'\mathbf{e}_2} := \check{\mathbf{W}}^{(e'_1, \ldots, e'_l, e_{21}, \ldots, e_{2n_2})}, \mathbf{e}' = (e'_1, \ldots, e'_l), \mathbf{e}_2 = (e_{21}, \ldots, e_{2n_2})$. Then by Markov property for X, (4.2) and moment duality, we have

(4.3)

$$
\begin{aligned}
&\mathbb{P}^{\mu}[I_{n_1}(X_{t_1};\phi_1)I_{n_2}(X_{t_2};\phi_2)]\\
&= \int I_{n_2}(\nu_2;\phi_2)Q_{t_2}(\mu,d\nu_2)\int I_{n_1}(\nu_1;\phi_1)Q_{t_1-t_2}(\nu_2,d\nu_1)\\
&= \sum_{\pi\in\mathcal{P}_{n_1}}\int I_{n_2}(\nu_2;\phi_2)I_{l(\pi)}\left(\nu_2;\left(\int_{\mathbb{R}^{n_1}}\psi_1(\mathbf{e}_1)w_{\mathbf{e}_1\pi}f_{\mathbf{e}_1\pi}d\mathbf{e}_1\right)\otimes\bar{\chi}_\pi\right)Q_{t_2}(\mu,d\nu_2)\\
&= \sum_{\pi\in\mathcal{P}_{n_1}}\int I_{l(\pi)+n_2}\left(\nu_2;\left(\int_{\mathbb{R}^{n_1}}\psi_1(\mathbf{e}_1)w_{\mathbf{e}_1\pi}f_{\mathbf{e}_1\pi}d\mathbf{e}_1\right)\otimes\psi_2\otimes\bar{\chi}_\pi\otimes\chi_2\right)Q_{t_2}(\mu,d\nu_2)\\
&= \sum_{\pi\in\mathcal{P}_{n_1}}\sum_{\pi^*\in\mathcal{P}_{l(\pi)+n_2}}\int_{\mathbb{R}^{n_1}}\psi_1(\mathbf{e}_1)w_{\mathbf{e}_1\pi}d\mathbf{e}_1\int_{\mathbb{R}^{l(\pi)}}f_{\mathbf{e}_1\pi}(\mathbf{e}_\pi)d\mathbf{e}_\pi\int_{\mathbb{R}^{n_2}}\psi_2(\mathbf{e}_2)\\
&\quad\mathbb{P}\left[1_{\{\check{\mathbf{W}}^{\mathbf{e}_\pi\mathbf{e}_2}_{t_2}\in\check{\mathbb{R}}^{l(\pi)+n_2}_{\pi^*}\}}\int\bar{\chi}_\pi\otimes\chi_2(k_{a_{\pi^*}(1)},\ldots,k_{a_{\pi^*}(l(\pi)+n_2)})\bigotimes_{i=1}^{l(\pi^*)}\mu(\check{W}^{\mathbf{e}_\pi\mathbf{e}_2}_{a_i(\pi^*)}(t_2))(dk_{a_i(\pi^*)})\right]d\mathbf{e}_2\\
&= \sum_{\pi\in\mathcal{P}_{n_1+n_2}}\int_{\mathbb{R}^{n_1}}\psi_1(\mathbf{e}_1)d\mathbf{e}_1\int_{\mathbb{R}^{n_2}}\psi_2(\mathbf{e}_2)\\
&\quad\mathbb{P}\left[1_{\{\check{\mathbf{W}}^{A}_{t_1}\in\check{\mathbb{R}}^{n_1+n_2}_{\pi}\}}\int\chi_1\otimes\chi_2(k_{a_\pi(1)},\ldots,k_{a_\pi(n_1+n_2)})\bigotimes_{i=1}^{l(\pi)}\mu(\check{W}^{A}_{a_i(\pi)}(t_1))(dk_{a_i(\pi)})\right]d\mathbf{e}_2.
\end{aligned}
$$

Hence, (4.1) is a consequence of Lemma 3.1 in [5]. □

Remark 4.2. Lemma 4.1 can be generalized to a duality involving a n-fold joint (over different times) moment of X. We leave the details to the readers. Also notice that there is a similar duality in voter model. See [2] for related accounts.

Given $z\in\mathbb{R}$ and $t>0$, write $X_{[z]}(t) := \lim_{a\to 0+}X_{t[z-a,z+a]}$ for the block average of X_t at site z. Notice that $X_{[z]}(t)$ always exists by Theorem 3.3. Since $X_t(z)$ is defined for m-almost all $z\in\mathbb{R}$, $X_{[z]}(t)$ seems to be more appropriate to describe the distribution of types at time t and at site z. The joint(over different times) moment duality in Lemma 4.1 can be used to study the correlation of $X_{[z_1]}(t_1)(A)$ and $X_{[z_2]}(t_2)(B)$, $A,B\subset K$.

Theorem 4.3. *Let θ be a diffuse measure in $M_1(K)$. Then for any $z\in\mathbb{R}$ and $t>0$, $\mathbb{Q}^{\theta^{\mathbb{R}}}$ almost surely, $X_{[z]}(t)$ is a point mass. It satisfies*

$$
\mathbb{Q}^{\theta^{\mathbb{R}}}\{X_{[z]}(t)(A)=1\}=\theta(A), A\subset K. \tag{4.4}
$$

Moreover, for any $0<t_2<t_1, z_1,z_2\in\mathbb{R}$ and sets $A\subset K, B\subset K$,

$$
\begin{aligned}
&\mathbb{Q}^{\theta^{\mathbb{R}}}[X_{[z_1]}(t)(A)X_{[z_2]}(t)(B)]\\
&= \theta(A\cap B)\int_{\frac{|z_1-z_2|}{\sqrt{2t}}}^{\infty}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy+\theta(A)\theta(B)\int_0^{\frac{|z_1-z_2|}{\sqrt{2t}}}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy
\end{aligned} \tag{4.5}
$$

and

$$\begin{aligned}
&\mathbb{Q}^{\mathbb{R}^\theta}[X_{[z_1]}(t_1)(A)X_{[z_2]}(t_2)(B)] \\
&= \theta(A\cap B)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(t_1-t_2)}}e^{-\frac{x^2}{2(t_1-t_2)}}dx\int_{\frac{|x+z_1-z_2|}{\sqrt{2t_2}}}^{\infty}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy \\
&\quad + \theta(A)\theta(B)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(t_1-t_2)}}e^{-\frac{x^2}{2(t_1-t_2)}}dx\int_0^{\frac{|x+z_1-z_2|}{\sqrt{2t_2}}}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy.
\end{aligned} \tag{4.6}$$

Proof. (4.4) just follows from Theorem 3.8 and (3.4).

To prove (4.5) we apply moment duality,

$$\begin{aligned}
&\mathbb{Q}^{\mathbb{R}^\theta}[X_{[z_1-\frac{1}{n},z_1+\frac{1}{n}]}(t)(A)X_{[z_2-\frac{1}{n},z_2+\frac{1}{n}]}(t)(B)] \\
&= Q_tI_2(\mathbb{R}^\theta; 2n1_{[z_1-\frac{1}{n},z_1+\frac{1}{n}]}\otimes 2n1_{[z_1-\frac{1}{n},z_1+\frac{1}{n}]}\otimes 1_A\otimes 1_B) \\
&= \theta(A\cap B)\int_{z_1-\frac{1}{n}}^{z_1+\frac{1}{n}}2ndx\int_{z_2-\frac{1}{n}}^{z_2+\frac{1}{n}}2n\mathbb{P}\{T^{x,y}\le t\}dy \\
&\quad + \theta(A)\theta(B)\int_{z_1-\frac{1}{n}}^{z_1+\frac{1}{n}}2ndx\int_{z_2-\frac{1}{n}}^{z_2+\frac{1}{n}}2n\mathbb{P}\{T^{x,y}> t\}dy,
\end{aligned}$$

where $T^{x,y} := \inf\{t\ge 0 : B_t^x = B_t^y\}$ is the first meeting time of two independent Brownian motions B^x and B^y with initial values x and y respectively. Letting $n\to\infty$, the reflection principle for Brownian motion yields

$$\begin{aligned}
&\mathbb{Q}^{\mathbb{R}^\theta}[X_{[z_1]}(t)(A)X_{[z_2]}(t)(B)] \\
&= \theta(A\cap B)\mathbb{P}\{T^{z_1,z_2}\le t\} + \theta(A)\theta(B)\mathbb{P}\{T^{z_1,z_2}>t\} \\
&= \theta(A\cap B)\int_{\frac{|z_1-z_2|}{\sqrt{2t}}}^{\infty}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy + \theta(A)\theta(B)\int_0^{\frac{|z_1-z_2|}{\sqrt{2t}}}\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy.
\end{aligned}$$

So (4.5) follows.

By Lemma 4.1,

$$\begin{aligned}
&\mathbb{Q}^{\mathbb{R}^\theta}[X_{[z_1-\frac{1}{n},z_1+\frac{1}{n}]}(t_1)(A)X_{[z_2-\frac{1}{n},z_2+\frac{1}{n}]}(t_2)(B)] \\
&= \theta(A\cap B)\int_{-\infty}^{\infty}dx\int_{z_1-\frac{1}{n}}^{z_1+\frac{1}{n}}\frac{2n}{\sqrt{2\pi(t_1-t_2)}}e^{-\frac{(x-x')^2}{2(t_1-t_2)}}dx' \\
&\qquad \int_{z_2-\frac{1}{n}}^{z_2+\frac{1}{n}}2n\mathbb{P}\{T^{x,y}\le t_2\} \\
&\quad + \theta(A)\theta(B)\int_{-\infty}^{\infty}dx\int_{z_1-\frac{1}{n}}^{z_1+\frac{1}{n}}\frac{2n}{\sqrt{2\pi(t_1-t_2)}}e^{-\frac{(x-x')^2}{2(t_1-t_2)}}dx' \\
&\qquad \int_{z_2-\frac{1}{n}}^{z_2+\frac{1}{n}}2n\mathbb{P}\{T^{x,y}> t_2\}.
\end{aligned} \tag{4.7}$$

Hence, (4.6) can be obtained by letting $n\to\infty$.

$\square$

Remark 4.4. It follows from Theorem 4.3 that the types of the two clusters at site z_1, z_2 and at time t_1, t_2 respectively are asymptotically independent as $|z_2 - z_1| + |t_2 - t_1| \to \infty$. This together with Theorem 3.1 shows that $X(t)$ converges to $\theta^{\mathbb{R}}$ in distribution but not in probability.

Acknowledgments The author is grateful to Steven Evans for helpful discussions. His suggestion of considering the case when θ is diffuse results in Theorem 3.8 and Theorem 4.3. The author also appreciates the numerous comments form anonymous referees.

References

[1] Cox, J.T., Durrett, R. and Perkins, E.: Rescaled voter models converge to super-Brownian motion. *Ann. Probab.*, **28**, 185–234 (2000).

[2] Cox, J.T. and Griffeath, D.: Occupation time limit theorems for the voter model. *Ann. Probab.*, **11**, 876–893 (1983).

[3] Dawson, D.A., Greven, A. and Vaillancourt, J.: Equilibria and quasi–equilibria for infinite collections of interacting Fleming–Viot processes. *Trans. Amer. Math. Soc.*, **347**, 2277–2360 (1995).

[4] Donnelly, P., Evans, S.N., Fleischmann, K., Kurtz T.G. and X. Zhou, X.: Continuum-sites stepping-stone models, coalescing exchangeable partitions, and random trees. *Ann. Probab.*, **28**, 1063–1110 (2000).

[5] Evans, S.N.: Coalescing Markov labeled partitions and continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincaré Probab.*, **33**, 339–358 (1997).

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

[7] Fleischmann, K. and Greven, A.: Time-space analysis of the cluster-formation in interacting diffusions. *Electron. J. Probab.*, **1** (1994).

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

[9] Kimura, M.: "Stepping-stone" models of population. Technical report 3, Institute of Genetics, Japan, 1953.

[10] Klenke, A.: Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.*, **24**, 660–697 (1996).

[11] Liggett, T.M.: *Interacting Particle Systems.* New York: Springer-Verlag, 1985.

[12] Müller, C. and Tribe, R.: Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Th. Rel. Fields*, **102**, 519–546 (1995).

[13] Sawyer, S.: Results for the stepping stone models for migration in population genetics. *Ann. Probab.*, **4**, 699–728 (1976).

[14] Shiga, T.: Stepping stone models in population genetics and population dynamics. In S. Albeverio et al., ed, *Stochastic Processes in Physics and Engineering*, Mathematics and Its Applications, 345–355. D. Reidel Publishing Company, 1988.

[15] Walters, P.: An introduction to ergodic theory. New York: Springer-Verlag, 1982.

[16] Zhou X.: Duality between coalescing and annihilating Brownian motions. Preprint.