Dispersal heterogeneity in the spatial Λ-Fleming-Viot process

Raphaël Forien

Dispersal heterogeneity in the spatial Λ-Fleming-Viot process

Raphaël Forien*

October 23, 2017

Abstract

We study the evolution of gene frequencies in a spatially distributed population when the dispersal of individuals is not uniform in space. We adapt the spatial Λ-Fleming-Viot process to this setting and consider that individuals spread their offspring farther from themselves at each generation in one halfspace than in the other. We study the large scale behaviour of this process and show that the motion of ancestral lineages is asymptotically close to a family of skew Brownian motions which coalesce upon meeting in one dimension, but never meet in higher dimension. This leads to a generalization of a result due to Nagylaki on the scaling limits of the gene frequencies: the non-uniform dispersal causes a discontinuity in the slope of the gene frequencies but the gene frequencies themselves are continuous across the interface.

Résumé

Cet article étudie l'évolution de la fréquence de certains gènes au sein d'une population structurée spatialement lorsque la dispersion des individus n'est pas uniforme dans l'espace. Nous adaptons le processus Λ-Fleming-Viot spatial à cette situation en considérant que la progéniture d'un individu donné se déplace en moyenne plus loin de son ascendant dans un demi espace que dans l'autre. Nous étudions le comportement à grande échelle (spatialle et temporelle) de ce processus et nous montrons que le processus des lignées ancestrales converge vers un système de skew mouvements Browniens qui coalescent dès

*CMAP, École polytechnique, CNRS, Université Paris-Saclay, 91128, Palaiseau Cedex, France; e-mail: raphael.forien@cmap.polytechnique.fr.
R.F. was supported in part by the chaire Modélisation Mathématique et Biodiversité of Veolia Environnement-École Polytechnique-Museum National d'Histoire Naturelle-Fondation X.

qu'ils se rencontrent en dimension 1 et qui ne se rencontrent jamais en dimension plus grande. Cela conduit à une généralisation d'un résultat sur le comportement à grande échelle des fréquences génétiques dû à Nagylaki : la dispersion inhomogène se traduit par une discontinuité de la pente des fréquences génétiques mais ces dernières sont continues à l'interface entre les deux domaines.

AMS 2010 subject classifications. Primary: 60J70, 60G57, 60F99, 92D10 ; Secondary: 60J25, 60J55.

Keywords: population genetics, generalised Fleming-Viot process, skew Brownian motion, duality.

Introduction

Landscape genetics studies the influence of geographical features of the environment on evolutionary processes and on the genetic composition of populations. Habitat fragmentation and ecological interfaces play a significant role in this field [MH13]. Scientists strive to detect, map and quantify the long term effects on genetic diversity of spatial heterogeneities by observing the genetic patterns that they have produced through evolution [Sla87]. For example, genetic differentiation between two subpopulations separated by a physical obstacle can be used to measure the reduction in gene flow caused by the obstacle [SQH+03, RPS+06, GCR+07].

Our focus in this work is the special case in which individuals spread their offspring farther from themselves in some parts of space than in others. By comparing the genomes of individuals and the frequencies of different genetic types (called *alleles*) at different locations, one tries to infer the strength of dispersal (or *gene flow*) in these regions and to measure the effect of the interface.

Simple models for the evolution of gene frequencies are then required which can be fitted to field data with reasonable computational power. That is why mathematicians in the field of population genetics establish large scale approximations of microscopic models which take into account the interaction between geographical features and evolutionary forces [Mal48, KW64, BDE02].

Nagylaki [Nag76] studied the effect of a discontinuity in the migration rate in the linear stepping stone model. He considered colonies located at the points $k/\sqrt{n}$, $k \in \mathbb{Z}$, which evolve in discrete generations spanning $1/n$ units of time. At each generation, adjacent colonies to the left of the origin exchange a proportion $m/2$ of migrants while adjacent colonies to the right exchange a proportion $v^2m/2$, as depicted in Figure 1.

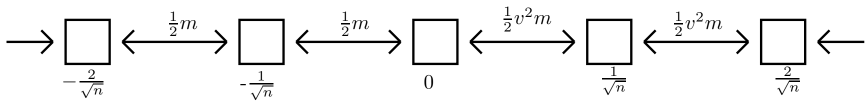

Figure 1: Discrete model with a discontinuity in the migration rate

Letting $n \to \infty$ and considering that the number of individuals in each colony is so large that genetic drift (*i.e.* fluctuations due to random sampling of individuals at each generation) can be ignored, Nagylaki showed that the proportion of individuals of a given type at location $x \in \mathbb{R}$ at time $t \geq 0$, denoted by $p(t,x)$, is well approximated by the solution to the following equation

$$\begin{cases} \dfrac{\partial p}{\partial t}(t,x) = \dfrac{m}{2}\dfrac{\partial^2 p}{\partial x^2}(t,x) & \text{if } x < 0 \\ \dfrac{\partial p}{\partial t}(t,x) = \dfrac{v^2m}{2}\dfrac{\partial^2 p}{\partial x^2}(t,x) & \text{if } x > 0 \end{cases}$$

and, for $t > 0$,

$$p(t,0^+) = p(t,0^-), \qquad\qquad \frac{\partial p}{\partial x}(t,0^-) = v^2\frac{\partial p}{\partial x}(t,0^+).$$

In words, allele frequencies must be continuous at zero but their first spatial derivative has a discontinuity which is given as a simple function of the ratio of the migration rates on each side of the habitat (see Figure 3). He extended this result [NB88] to the probability of identity by descent, *i.e.* the probability that two uniformly sampled individuals have inherited the same allele from a common ancestor without mutation as a function of the distance between the sampling locations. Nagylaki found similar conditions for the first derivative of the probability of identity as for the allele frequencies. Along with Ayati and Dupont [ADN99], he further investigated the qualitative properties of the probability of identity in this setting and provided numerical approximations.

In parallel to these developments, a diffusion process has been introduced [IM63, Wal78, HS81] and used to study diffusion in physical systems presenting an interface between different media [ABT$^+$11]. The so-called *skew Brownian motion* with parameter $\alpha \in [0,1]$ can be described as an $\mathbb{R}$-valued stochastic process which performs Brownian excursions from the origin, on the positive half line with probability α and on the negative half line with probability $1-\alpha$. See [Lej06] for a review of the definition and properties of skew Brownian motion.

In this paper, we study the genealogy of a sample of individuals in the presence of heterogeneous dispersal. This genealogy is described by a system of ancestral lineages which at time t correspond to the positions of the ancestors of the sample t generations in the past. We find that, in the diffusion limit, those ancestral lineages follow skew Brownian motions with different diffusion coefficients on each side of the interface (Proposition 3.3 below). The genealogy of a sample of individuals is then given by a system of skew Brownian motions which coalesce upon meeting in one dimension but never coalesce in higher dimensions (Theorem 2). As a consequence, allele frequencies follow a deterministic partial differential equation in dimensions two and higher while in one dimension, patches of different types form and evolve randomly (Theorem 1). Our method allows for more general assumptions on the microscopic model than [Nag76, Nag88] (*e.g.* continuous spatial structure and non-nearest neighbour migration).

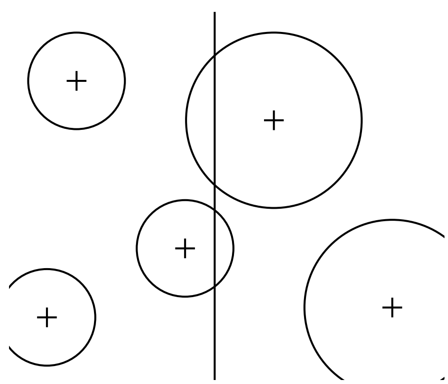

Figure 2: Size of reproduction events
The size of the region affected by a reproduction event depends on the halfspace in which its centre falls ($x_1 > 0$ or $x_1 < 0$).

We use the spatial Λ-Fleming-Viot process framework introduced in [BEV10] and [Eth08] to model the evolution of allele frequencies in a continuous space (see [BEV13a] for a review on this process). In this model, reproduction events occur according to a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^d$ which specifies their time and location. During these reproduction events, a proportion u - called the *impact parameter* - of individuals in a ball of radius r is replaced by the offspring of a uniformly sampled individual in this ball. To model heterogeneous dispersal, we assume that the radius of the reproduction event depends on the halfspace in which its centre falls, as illustrated in Figure 2. We study the large scale behaviour of the spatial Λ-Fleming-Viot process (SLFV) under a diffusive rescaling similar to the one considered in the homogeneous setting in [BEV13b]. In particular, the impact parameter is kept constant as we rescale space and time.

Our results and their proofs are similar in spirit to those in [BEV13b]. We use the fact that the SLFV has a dual in the form of a system of coalescing particles moving in $\mathbb{R}^d$ (interpreted as the locations in the past of the ancestors of a random sample of individuals). We show (Theorem 2) that the rescaled dual converges to a system of skew Brownian motions which evolve independently of each other until they meet, and then coalesce instantaneously upon meeting. In particular, when $d \geq 2$, the particles never meet and evolve independently of each other. Our approach improves on [BEV13b] as our proof covers any configuration where ancetral lineages converge to Markov processes with continuous paths.

As a consequence, we obtain a scaling limit of the process describing the evolution of allele frequencies across space (Theorem 1). The limit is deterministic as soon as $d \geq 2$ and solves a heat equation on each halfspace. The fact that ancestral lineages follow skew Brownian motions translates into a discontinuity of the first spatial derivative along the normal of the interface, in agreement with Nagylaki's result. When $d = 1$, each site is occupied by only one type of individuals at any positive time, and the boundaries between patches of different types evolve according to a system of annihilating skew Brownian motions.

The proof of the convergence of the motion of lineages to skew Brownian motion is adapted from the work of A. Iksanov and A. Pilipenko [IP16], where skew Brownian motion is obtained as a scaling limit of a Markov chain on $\mathbb{Z}$ which behaves like simple random walk outside a bounded region around the origin. The difficulty in proving convergence to skew Brownian motion comes from the fact that martingale problem characterizations of the limiting process are ill suited to this setting. (In particular, scale functions of the limiting process do not turn the random walk into a martingale.) Following [IP16], we circumvent this by studying the positive and negative parts of the process separately, and then linking the two by their respective local times at the origin. This method turns out to be readily applicable to more general migration patterns than originally studied in [Nag76], as we show here by dealing with a continuous spatial structure.

The paper is laid out as follows. We define the SLFV with heterogeneous dispersal in Section 1 and we state our main result (Theorem 1) in Section 2. Section 3 gives a description of the dual of the SLFV and states its convergence under the diffusive rescaling (Theorem 2). The latter is proved in Section 4 and implies Theorem 1. Finally, the convergence of the motion of an ancestral lineage to skew Brownian motion is proved in Section 5, following the arguments of [IP16].

Acknowledgements

The author would like to thank Amandine Véber and Alison Etheridge for many helpful discussions on this work and for their comments on this paper.

1 Definition of the model

Consider a model where individuals are scattered in a continuous space of dimension d and can be of two types, denoted by 0 or 1. We suppose that the density of individuals is constant in space. The population is represented by a random function $\{w(t,x), t \geq 0, x \in \mathbb{R}^d\}$, where $w(t,x) \in [0,1]$ is interpreted as the proportion of type 1 individuals at location x at time t. Define the two halfspaces $\mathbb{H}^+$, $\mathbb{H}^-$ by

$$\mathbb{H}^{\pm} = \left\{ x \in \mathbb{R}^d : \pm x_1 > 0 \right\}.$$

Take $u \in (0,1]$ and $0 < r_- \leq r_+ < +\infty$. We denote the volume of the ball of radius r in $\mathbb{R}^d$ by V_r. The SLFV with heterogeneous dispersal is defined as follows.

Definition 1.1 (SLFV with heterogeneous dispersal)**.** *Let $\Pi^{\pm}$ be a Poisson point process on $\mathbb{H}^{\pm} \times \mathbb{R}_+$ with intensity $\frac{1}{V_{r_\pm}} dxdt$. For each point (x,t) in $\Pi^{\pm}$, a reproduction event takes place in $B(x, r_\pm)$ at time t:*

1) Pick a location y uniformly at random in $B(x,r_\pm)$ and sample a type $k \in \{0,1\}$ from the types present at y (i.e. $k = 1$ with probability $\frac{1}{V_{r_\pm}} \int_{B(x,r_\pm)} w(t_-, y) dy$).

2) Update $w(t,z)$ for $z \in B(x,r_\pm)$ as follows:

$$w(t,z) = (1-u)w(t_-,z) + u\mathbb{1}_{\{k=1\}}.$$

In other words, a proportion u of individuals in the ball of centre x and radius $r_\pm$ dies and is replaced by the offspring of an individual sampled uniformly from this ball.

Remark. *The factor $\frac{1}{V_{r_\pm}}$ in the rate of the Poisson point process ensures that the mean lifetime of individuals is the same in both halfspaces (far enough from the interface).*

Theorem 4.2 in [BEV10] can be adapted without difficulty to show that there exists a unique càdlàg Markov process $(w(t,\cdot))_{t\geq 0}$ satisfying this definition and taking values in the quotient space Ξ of Lebesgue-measurable maps from $\mathbb{R}^d$ to $[0,1]$ that are identified when they coincide up to a Lebesgue-null set. This space can be identified with (a subset of) the space of measures on $\mathbb{R}^d$ that are absolutely continuous with respect to Lebesgue measure. It is endowed with the following metric d which induces the topology of vague convergence of measures on $\mathbb{R}^d$. Let $(f_n)_{n\geq 1}$ be a separating family of uniformly bounded and compactly supported real-valued functions on $\mathbb{R}^d$, then

$$d(w,w') = \sum_{n=1}^{\infty} \frac{1}{2^n} \left| \langle w, f_n \rangle - \langle w', f_n \rangle \right|, \qquad w, w' \in \Xi.$$

2 Large scale behaviour of the SLFV with heterogeneous dispersal

Fix $w_0 : \mathbb{R}^d \to [0,1]$. For $n \geq 1$, set $w^n(t,x) = w(nt, \sqrt{n}x)$ and assume that $w^n(0,x) = w_0(x)$ for all $n \geq 1$. For $\beta \in (-1,1)$, let $\mathcal{D}^\beta$ denote the set of all continuous functions $\phi : \mathbb{R}^d \to \mathbb{R}$, twice continuously differentiable on each halfspace $\mathbb{H}^\pm$, such that

$$(1+\beta) \left. \frac{\partial \phi}{\partial x_1} \right|_{x_1=0^+} = (1-\beta) \left. \frac{\partial \phi}{\partial x_1} \right|_{x_1=0^-}.$$

Theorem 1. *As $n \to \infty$, the sequence of Ξ-valued processes $\{w^n(t,\cdot), t \geq 0\}$ converges in the sense of finite dimensional distributions in the vague topology to a process $\{p(t,\cdot), t \geq 0\}$. In dimension one, $p(t,x)$ is a Bernoulli random variable with parameter $\rho(t,x)$ and the correlations between the values of $p(t,\cdot)$ at distinct sites are non trivial and are given in* (8) *(see also Figure 4). In dimensions two and higher, $p(t,x)$ is deterministic and equals $\rho(t,x)$. In both cases, there exists $\beta \in (0,1)$ such that $\rho(t,\cdot)$ is the solution in $\mathcal{D}^\beta$ to the following equation*

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(t,x) = \dfrac{u r_\pm^2}{d+2} \Delta \rho(t,x) & \text{if } x \in \mathbb{H}^\pm, \\ \rho(0,x) = w_0(x) & x \in \mathbb{R}^d. \end{cases} \tag{1}$$

Finding solutions to (1) in $\mathcal{D}^\beta$ can be reduced to finding classical solutions to the heat equation with discontinuous coefficients by a change of variables as shown in [Nag76]. Existence and uniqueness of the solution in $\mathcal{D}^\beta$ to (1) was also proved in [Por79a] and [Por79b], see also Proposition 1 in [Lej06]. We prove Theorem 1 by studying the dual of the SLFV with heterogeneous dispersal.

The fact that the solution to (1) has to be found in $\mathcal{D}^\beta$ with $\beta \geq 0$ agrees with the findings of Nagylaki [Nag76] (equations 8 and 9). This transmission condition reflects the fact that individuals living near the frontier between the two halfspaces are more likely to have ancestors coming from $\mathbb{H}^+$ than from $\mathbb{H}^-$ (recall that we take $r_- \leq r_+$), see Figure 3.

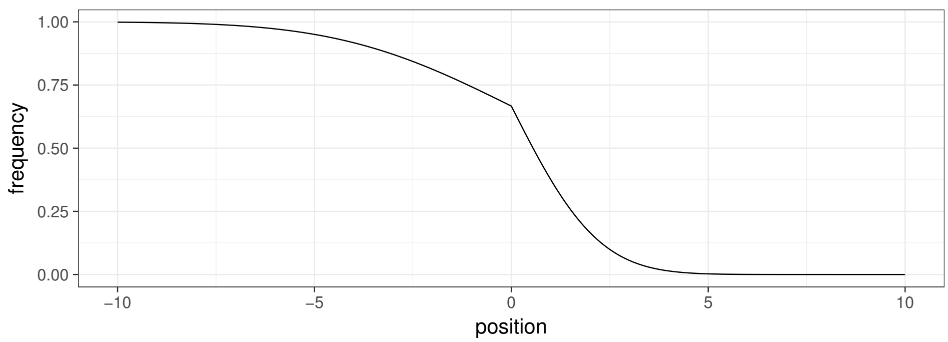

Figure 3: Diffusion of an allele with heterogeneous dispersal
Graphical representation of $x \mapsto \rho(t, x)$ started from a Heavyside initial condition $\mathbb{1}_{\{x<0\}}$ at time $t = 12$ with parameters: $\sigma_+ = 0.5$, $\sigma_- = 1$, $\beta = -0.6$. Note the discontinuity in the first spatial derivative at $x = 0$.

As already noted by Nagylaki [Nag76], β depends on the microscopic model in a rather intricate way. We give an expression for β in (26) when the microscopic model is the SLFV. This dependence on the choice of the model is a potential issue when trying to infer demographic parameters from genetic data. Inferring β as an independent parameter would reduce the power of such an inference scheme, so one would like to choose a particular model and make β a function of the other parameters in the model. However it isn't clear how one should choose among the great variety of possible microscopic models.

3 Duality

3.1 The dual of the SLFV with heterogeneous dispersal

We now define a system of coalescing particles whose displacements are driven by the same Poisson point process of reproduction events as the SLFV. The particles at time t describe the positions of the set of ancestors at time $-t$ of a sample of individuals alive at time 0. Since the Poisson point processes $\Pi^\pm$ are reversible with respect to time, the reproduction events which took place in the past have the same distribution as those which occur forwards in time.

Definition 3.1 (Dual of the SLFV with heterogeneous dispersal)**.** *Let $\Pi^\pm$ be Poisson point processes in $\mathbb{H}^\pm \times \mathbb{R}_+$ with intensity $\frac{1}{V_{r_\pm}} dxdt$. Let $(\mathcal{A}_t)_{t\geq 0}$ be a system of finitely many particles whose dynamics are as follows. For each point (x,t) in $\Pi^\pm$, a reproduction event takes place in $B(x,r_\pm)$ at time t:*

1) Pick a location y uniformly at random in $B(x,r_\pm)$.

2) Each particle sitting inside $B(x,r_\pm)$ at time t_- is marked with probability u, independently of each other.

3) All marked particles coalesce and move to y.

We denote the number of particles present at time t by N_t and their spatial locations by $\xi_t^1, \dots, \xi_t^{N_t}$, so that $\mathcal{A}_t = \{\xi_t^1, \dots, \xi_t^{N_t}\}$.

Let $B^\pm(x,r)$ denote the intersection of $B(x,r)$ and $\mathbb{H}^\pm$. The motion of one particle is a Markov process on $\mathbb{R}^d$ with infinitesimal generator

$$\mathcal{L}f(x) = u \int_{\mathbb{R}^d} \Phi(x,y)(f(y) - f(x))dy \tag{2}$$

with

$$\Phi(x,y) = \frac{|B^+(x,r_+) \cap B^+(y,r_+)|}{V_{r_+}^2} + \frac{|B^-(x,r_-) \cap B^-(y,r_-)|}{V_{r_-}^2}. \tag{3}$$

This is seen by noting that a particle located at x finds itself in the region of a reproduction event of $\Pi^\pm$ at rate

$$\frac{|B^\pm(x,r_\pm)|}{V_{r_\pm}}.$$

It is further affected by such an event with probability u and moves to a location y chosen uniformly in the ball of radius $r_\pm$ affected by the event. See [BEV13a] (paragraph 3.5) for a more detailed justification in the homogeneous case. The law of $(\mathcal{A}_t)_{t\geq 0}$ started from j lineages at locations $\underline{x} = (x_1, \dots, x_j)$ is denoted by $\mathbb{P}_{\underline{x}}(\cdot)$.

Let us now give the (weak) duality relation between $(w_t)_{t\geq 0}$ and $(\mathcal{A}_t)_{t\geq 0}$. Let $C_c(\mathbb{R}^d)$ be the space of compactly supported real valued functions on $\mathbb{R}^d$.

For $\psi : (\mathbb{R}^d)^j \to \mathbb{R}_+$ in $C_c((\mathbb{R}^d)^j)$ and $w \in \Xi$, set

$$I(w,\psi) = \int_{(\mathbb{R}^d)^j} \prod_{i=1}^{j} w(x_i)\psi(x_1,\ldots,x_j)dx_1\ldots dx_j.$$

Also set

$$\langle w, \mathcal{A}_t\rangle = \prod_{i=1}^{N_t} w(\xi_t^i).$$

Then, for any $j \in \mathbb{N}$, for $\psi \in C_c((\mathbb{R}^d)^j)$, [BEV10]

$$\mathbb{E}_{w_0}\left[I(w_t,\psi)\right] = \int_{(\mathbb{R}^d)^j} \mathbb{E}_{\underline{x}}\left[\langle w_0, \mathcal{A}_t\rangle\right]\psi(\underline{x})d\underline{x}. \tag{4}$$

Since the linear span of functions of the form $I(\cdot,\psi)$ and constant functions is dense in $C(\Xi)$ (Lemma 4.1 in [BEV10]), one can prove Theorem 1 by showing that, for any $0 \le t_1 < \ldots < t_k$ and $\psi_1,\ldots,\psi_k$ in $C_c((\mathbb{R}^d)^j)$,

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{i=1}^{k} I(w_{t_i}^n,\psi_i)\right] = \mathbb{E}\left[\prod_{i=1}^{k} I(p_{t_i},\psi_i)\right]. \tag{5}$$

We shall do this using the duality relation (4) above. For $n \ge 1$, define the rescaled dual process $(\mathcal{A}_t^n)_{t\ge 0}$ by

$$\mathbb{E}_{\underline{x}}\left[f(\mathcal{A}_t^n)\right] = \mathbb{E}_{\sqrt{n}\underline{x}}\left[f\left(\tfrac{1}{\sqrt{n}}\xi_{nt}^1,\ldots,\tfrac{1}{\sqrt{n}}\xi_{nt}^{N_{nt}}\right)\right].$$

Then $(\mathcal{A}_t^n)_{t\ge0}$ is dual to $(w_t^n)_{t\ge0}$ in the sense that

$$\mathbb{E}_{w_0}\left[I(w_t^n,\psi)\right] = \int_{(\mathbb{R}^d)^j} \mathbb{E}_{\underline{x}}\left[\langle w_0, \mathcal{A}_t^n\rangle\right]\psi(\underline{x})d\underline{x}.$$

In Section 4, we prove the convergence of $(\mathcal{A}_t^n)_{t\ge0}$ to a system of coalescing skew Brownian motions. Note that in dimensions two and higher, skew Brownian motions never meet and the dual of the SLFV with heterogeneous dispersal thus converges to a system of independent skew Brownian motions. This is the reason why the SLFV converges to a deterministic process when $d \ge 2$ in Theorem 1.

3.2 Skew Brownian motion

In [HS81] (see also [Wal78], [LG84] and [Lej06]) it is shown that for $\beta \in [-1, 1]$, there exists a unique solution to the equation

$$X_t = B_t + \beta L_t^0(X),$$

where B is standard Brownian motion and $L_t^0(X)$ is the local time at 0 of X. This process is called skew Brownian motion with parameter $\alpha = \frac{\beta+1}{2}$. (For $\beta = 1$, $(X_t)_{t\geq 0}$ is reflected Brownian motion.) This result can be extended to the d-dimensional case where the first coordinate of the process follows skew Brownian motion.

Proposition 3.2. *Let $B = \left(B_t^1, \ldots, B_t^d\right)_{t\geq 0}$ be standard (d dimensional) Brownian motion. Let $\sigma : \mathbb{R}^d \to (0, \infty)$ be defined by $\sigma^2(x) = \sigma_\pm^2 \mathbb{1}_{\{x\in\mathbb{H}^\pm\}}$ with $\sigma_\pm^2 > 0$ and take $x_0 = (x_0^1, \ldots, x_0^d) \in \mathbb{R}^d$. Then, for $\beta \in [-1, 1]$, there exists a unique $\mathbb{R}^d$-valued Markov process $(X_t)_{t\geq 0}$ satisfying*

$$\begin{aligned} X_t^1 &= x_0^1 + \int_0^t \sigma(X_s) dB_s^1 + \beta L_t^0(X^1) \\ X_t^i &= x_0^i + \int_0^t \sigma(X_s) dB_s^i \qquad\qquad \textit{for } 2 \leq i \leq d. \end{aligned} \tag{6}$$

Furthermore, the law of $(X_t)_{t\geq 0}$ is the unique solution to the (hence well posed) martingale problem associated with the generator L, defined on the domain $\mathcal{D}^\beta$ by

$$L\phi(x) = \frac{1}{2}\sigma^2(x)\Delta\phi(x), \qquad \forall \phi \in \mathcal{D}^\beta.$$

This result is proved in [Lej06] (Proposition 10) in the case $d = 1$ and $\sigma_+ = \sigma_-$. The extension to higher dimensions is straightforward and the case $\sigma_+ \neq \sigma_-$ can be treated with the help of [BP87]. In [Por79a], [Por79b], it is proved that L generates a Feller semigroup. Part of the work in showing Theorem 1 is the proof that the motion of particles in $\mathcal{A}^n$ converges to a solution to (6), as stated in the following Proposition. Its proof is given in Section 5.

Proposition 3.3 (Convergence to skew Brownian motion)**.** *Let $(\xi_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Markov process with infinitesimal generator $\mathcal{L}$ given in (2). For $n \geq 1$, set $\xi_t^n = \frac{1}{\sqrt{n}}\xi_{nt}$ and suppose ξ_0^n is deterministic and converges to $x_0 \in \mathbb{R}$ as $n \to \infty$. Fix $T > 0$. Then, as $n \to \infty$, $(X_t^n)_{t\in[0,T]}$ converges in distribution in the Skorokhod space $\mathrm{D}\left([0, T], \mathbb{R}^d\right)$ to $(X_t)_{t\in[0,T]}$, a solution to (6) with $\sigma_\pm^2 = u\frac{2r_\pm^2}{d+2}$, and $\beta \in (0, 1)$.*

The parameter β is given as a (complicated) function of the law of $(\xi_t)_{t\geq 0}$ in (26). Note however that $\beta \geq 0$ as soon as $r_+ \geq r_-$.

3.3 Large scale behaviour of the dual process

Let $(\mathcal{A}_t^\infty)_{t\geq 0}$ be a system of particles moving in $\mathbb{R}^d$ according to independent skew Brownian motions (*i.e.* solutions to (6)) with $\sigma_\pm^2 = u\frac{2r_\pm^2}{d+2}$ and with the same parameter β which coalesce instantaneously upon meeting. In particular, in dimension 2 and higher, the particles never meet and $(\mathcal{A}_t^\infty)_{t\geq 0}$ is a system of independent skew Brownian motions. We denote the locations of the particles at time t by $\{X_t^1, \ldots, X_t^{N_t}\}$.

From [Eva97], we know that there exists a Ξ-valued process $\{p(t,x), t \geq 0, x \in \mathbb{R}^d\}$ which is dual to $\mathcal{A}^\infty$ in the sense that, for $\psi \in C_c((\mathbb{R}^d)^j)$,

$$\mathbb{E}_{w_0}\left[I(p_t, \psi)\right] = \int_{\left(\mathbb{R}^d\right)^j} \mathbb{E}_{\underline{x}}\left[\langle \mathcal{A}_t^\infty, w_0\rangle\right] \psi(\underline{x}) d\underline{x}. \tag{7}$$

Furthermore, by Lemma 3.2 in [BEV13b], in dimension one, $p(t,x)$ is a Bernoulli random variable with parameter $\rho(t,x) = \mathbb{E}_x\left[w_0(Z_t)\right]$ while in dimensions two and higher, $p(t,x)$ is deterministic and equals $\rho(t,x)$. The fact that ρ can be characterized as the solution to (1) is a direct consequence of operator semigroup theory (see [EK86] and recall that L generates a Feller semigroup). In [BEV13b], it is shown that the following theorem implies (5) and hence Theorem 1 (see their proof of Theorem 1.1).

Theorem 2. *As $n \to \infty$, $(\mathcal{A}_t^n)_{t\geq 0}$ converges in the sense of finite dimensional distributions to $(\mathcal{A}_t^\infty)_{t\geq 0}$.*

Moreover, for $k \in \mathbb{N}$ and $0 \leq t_1 < \ldots < t_k$, suppose that we start $\mathcal{A}^n$ with j_0 particles at locations $\underline{x}_0$, let the process evolve until time t_1, add j_1 lineages at locations $\underline{x}_1$, let the process evolve until time t_2 and so on. Call the resulting process $\hat{\mathcal{A}}^n$ and define $\hat{\mathcal{A}}^\infty$ analogously. Then for any $t \geq 0$, $\hat{\mathcal{A}}_t^n$ converges in distribution to $\hat{\mathcal{A}}_t^\infty$ as $n \to \infty$.

From (7), we obtain that for Lebesgue almost every $(x_1, \ldots, x_j) \in (\mathbb{R}^d)^j$,

$$\mathbb{E}_{w_0}\left[\prod_{i=1}^{j} p(t, x_i)\right] = \mathbb{E}_{x_1,\ldots,x_j}\left[\prod_{i=1}^{N_t} w_0(X_t^i)\right] \tag{8}$$

yielding the correlations between the values of $p(t,\cdot)$ at different sites.

In dimensions two and higher, lineages never coalesce and evolve independently of each other. As a result, one can show (see [BEV13b])

$$\mathbb{E}_{w_0}\left[p(t,x)^2\right] = \mathbb{E}_x\left[w_0(X_t)\right]^2 = \mathbb{E}_{w_0}\left[p(t,x)\right]^2,$$

which is only possible if p is deterministic.

In dimension one, since lineages coalesce when they meet, at any positive time each location is occupied by only one type of individuals. Small patches of type 1 and type 0 individuals then form, whose borders can be shown to follow anihilating skew Brownian motions. Neighbouring patches of the same type thus merge whenever their borders meet, as illustrated in Figure 4.

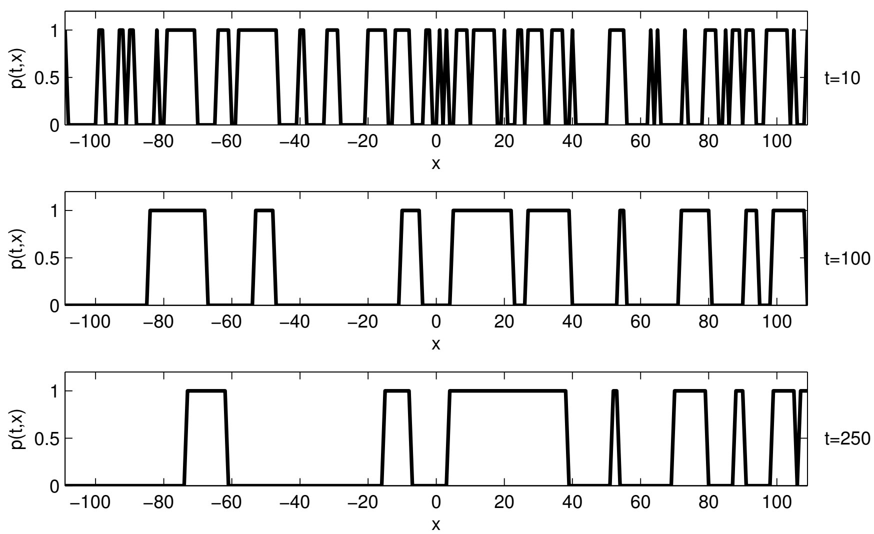

Figure 4: The limiting process in dimension one

Numerical simulation of $(p(t,\cdot))_{t\geq 0}$ in a one dimensional space of length 220 with $\sigma_-^2 = 0.2$, $\sigma_+^2 = 0.06$ and $\beta = 7/13$, started from the initial condition $w_0(x) \equiv 0.5$, shown at time $t = 10$, $t = 100$ and $t = 250$. Notice how the number of patches decreases with time as their interfaces meet and annihilate each other. Patches on the right are smaller and more numerous than patches on the left because diffusion is stronger on the left than on the right of the origin.

Remark. *Lineages coalesce instantaneously upon meeting because the impact parameter u (which should be interpreted as the inverse of the effective population size) is kept constant as we rescale time and space. Other scalings would result in different limiting behaviours. If u is of order $1/\sqrt{n}$, then we expect that, in the limit, lineages coalesce when they accumulate a local time together equal to an independent exponential random variable, as in [DR08]. The evolution of allele frequencies is then described by a stochastic partial* *differential equation in one spatial dimension (but remains deterministic in higher dimensions as skew Brownian motions never meet), as in [EVY14]. Moreover, if $u = o\left(1/\sqrt{n}\right)$, lineages never coalesce in the limit, even in one dimension, and the evolution of allele frequencies is deterministic (and equal to ρ).*

4 Proof of Theorem 2

Proposition 3.3 gives the convergence of the law of the motion of each particle in $\mathcal{A}^n$ to skew Brownian motion. To show Theorem 2, we thus need to control the coalescence of the particles. The following proposition helps fulfill this goal.

Proposition 4.1. *Let $O \subset \mathbb{R}^d$ be an open set and let $F \subset \mathbb{R}^d$ be a closed set. Suppose that a sequence of functions (or processes) $f_n : \mathbb{R}_+ \to \mathbb{R}^d$ converges uniformly on every compact interval to a continuous function $f : \mathbb{R}_+ \to \mathbb{R}^d$. Define $T_O^n = \inf\{t \geq 0 : f_n(t) \in O\}$ and T_F^n, T_O and T_F accordingly. Then*

$$T_F \leq \liminf_{n\to\infty} T_F^n, \qquad\qquad \limsup_{n\to\infty} T_O^n \leq T_O.$$

This Proposition is proved in Appendix A. An immediate consequence is that if a sequence of processes $\left\{\left(X_t^n\right)_{t\geq 0}, n \geq 1\right\}$ converges in distribution in $\mathrm{D}\left([0,T], \mathbb{R}^d\right)$ to a continuous process $(X_t)_{t\geq 0}$, and if $T_O = T_F$ a.s. when F is the closure of O (defining T_F, T_O, T_F^n and T_O^n as the hitting times of these sets by the processes $(X_t)_{t\geq 0}$ and $(X_t^n)_{t\geq 0}$ respectively), then, by the Skorokhod representation theorem, both T_O^n and T_F^n converge in distribution to $T_O = T_F$.

Proof of Theorem 2. We prove the first part of the result when starting from two particles; the proof is easily extended to a larger sample (see [BEV13b]). The two particles in $\mathcal{A}^n$ evolve independently of each other until they come within a distance $2r_+/\sqrt{n}$ of each other (since $r_- \leq r_+$). Let us then define T_n as the first time at which the two particles come close to each other in the rescaled setting

$$T_n = \inf\left\{t \geq 0 : \left|\xi_t^{n,1} - \xi_t^{n,2}\right| \leq \frac{2r_+}{\sqrt{n}}\right\}. \tag{9}$$

When $d \geq 2$, we show that $\mathbb{P}_{x_1,x_2}\left(T_n \leq t\right) \to 0$ as $n \to \infty$ for all $t > 0$. For $\varepsilon > 0$, define

$$T_n^\varepsilon = \inf\left\{t \geq 0 : \left|\xi_t^{n,1} - \xi_t^{n,2}\right| \leq 2r_+\varepsilon\right\}.$$

This is the hitting time of the closed set $\{(x,y) : |x-y| \le 2r_+\varepsilon\}$ by the process $(\xi_t^{n,1}, \xi_t^{n,2})_{t\ge 0}$. Since $\xi^{n,1}$ and $\xi^{n,2}$ are independent up to time T_n and, for n large enough, $T_n \ge T_n^\varepsilon$, by Proposition 3.3 and Proposition 4.1, T_n^ε converges in distribution to T^ε, defined as the hitting time of $\{(x,y) : |x-y| \le 2r_+\varepsilon\}$ by two independent solutions to (6) started from x_1 and x_2. As a result, since $T_n \ge T_n^\varepsilon$ a.s. for n large enough,

$$\limsup_{n\to\infty} \mathbb{P}_{x_1,x_2}\left(T_n \le t\right) \le \mathbb{P}_{x_1,x_2}\left(T^\varepsilon \le t\right).$$

The right-hand-side vanishes as $\varepsilon \downarrow 0$ when $d \ge 2$, yielding the result in this case.

We treat the case $d = 1$ in two steps. First we prove that the trajectory of the two particles up to time T_n converges in distribution to the motion of two independent skew Brownian motions up to their meeting time. Then we argue that the coalescence happens soon enough once the two particles are close to each other that the delay between T_n and the coalescence time (denoted by T_n^c) vanishes in the limit.

By the Skorokhod representation theorem and by Proposition 3.3, there exist sequences of processes $(\tilde{\xi}_t^{n,1}, \tilde{\xi}_t^{n,2})_{t\ge 0}$ and $(\tilde{X}_t^1, \tilde{X}_t^2)_{t\ge 0}$ defined on some probability space such that

i) $(\tilde{\xi}_t^{n,1})_{t\ge 0}$ and $(\tilde{\xi}_t^{n,2})_{t\ge 0}$ are independent Markov processes with infinitesimal generator $\mathcal{L}$,

ii) $(\tilde{X}_t^1)_{t\ge 0}$ and $(\tilde{X}_t^2)_{t\ge 0}$ are independent solutions to (6),

iii) $(\tilde{\xi}_t^{n,i})_{t\ge 0}$ converges uniformly on compact time intervals to $(\tilde{X}_t^i)_{t\ge 0}$ almost surely for $i \in \{1,2\}$.

Defining $\tilde{T}_n$ analogously to (9), $(\tilde{\xi}_t^{n,1}, \tilde{\xi}_t^{n,2})_{t\le \tilde{T}_n}$ has the same distribution as $(\xi_t^{n,1}, \xi_t^{n,2})_{t\le T_n}$. Suppose that $\tilde{X}_0^1 > \tilde{X}_0^2$ and define the hitting time of the diagonal by $(\tilde{X}_t^1, \tilde{X}_t^2)_{t\ge 0}$ as

$$\tilde{T}^\Delta = \inf\{t \ge 0 : \tilde{X}_t^1 \le \tilde{X}_t^2\}.$$

Let us show that $\tilde{T}_n \underset{n\to\infty}{\longrightarrow} \tilde{T}^\Delta$ almost surely. Set

$$\tilde{T}_n^\Delta = \inf\{t \ge 0 : \tilde{\xi}_t^{n,1} \le \tilde{\xi}_t^{n,2}\}$$

and note that since the jumps of $\tilde{\xi}^{n,i}$ are of size at most $2r_+/\sqrt{n}$, the two lineages cannot jump over one another without coming within a distance $2r_+/\sqrt{n}$ of each other, *i.e.* $\tilde{T}_n \le \tilde{T}_n^\Delta$ almost surely. Moreover, define $\tilde{T}_n^\varepsilon$

and $\tilde{T}^\varepsilon$ as the hitting times of $\{(x,y) : |x-y| \le 2r_+\varepsilon\}$ by $(\tilde{\xi}^{n,1}_t, \tilde{\xi}^{n,2}_t)_{t\ge 0}$ and $(\tilde{X}^1_t, \tilde{X}^2_t)_{t\ge 0}$ respectively. By Proposition 4.1, $\tilde{T}^\Delta_n \underset{n\to\infty}{\longrightarrow} \tilde{T}^\Delta$ a.s. and $\tilde{T}^\varepsilon_n \underset{n\to\infty}{\longrightarrow} \tilde{T}^\varepsilon$ a.s. As a result, for all $\varepsilon > 0$,

$$\tilde{T}^\varepsilon \le \liminf_{n\to\infty} \tilde{T}_n \le \limsup_{n\to\infty} \tilde{T}_n \le \tilde{T}^\Delta \qquad a.s.$$

By the continuity of $t \mapsto (\tilde{X}^1_t, \tilde{X}^2_t)$, $\tilde{T}^\varepsilon \to \tilde{T}^\Delta$ almost surely as $\varepsilon \downarrow 0$, yielding the almost sure convergence of $\tilde{T}_n$ to $\tilde{T}^\Delta$. As a result, $(\tilde{\xi}^{n,1}_t, \tilde{\xi}^{n,2}_t)_{t\le \tilde{T}_n}$ converges almost surely to $(\tilde{X}^1_t, \tilde{X}^2_t)_{t\le \tilde{T}^\Delta}$. In other words, $(\xi^{n,1}_t, \xi^{n,2}_t)_{t\le T_n}$ converges in distribution to $(X^1_t, X^2_t)_{t\le T^\Delta}$, the trajectory of two independent skew Brownian motions stopped at the time when they hit each other.

We now show that the two particles coalesce quickly once they come within a distance $2r_+/\sqrt{n}$ of each other. This is a consequence of the following result, which is proved as in [BEV10], Proposition 6.4.

Lemma 4.2. *Let T^c denote the coalescence time of the two particles ξ^1_t, ξ^2_t in $(\mathcal{A}_t)_{t\ge 0}$ (i.e. in the original time scale). Then*

$$\lim_{t\to\infty} \sup_{|y_1-y_2|\le 2r_+} \mathbb{P}_{y_1,y_2}(T^c > t) = 0.$$

By the strong Markov property,

$$\mathbb{P}_{x_1,x_2}(T^c_n - T_n > t) = \mathbb{E}_{x_1,x_2}\left[\mathbb{P}_{\sqrt{n}\xi^{n,1}_{T_n}, \sqrt{n}\xi^{n,2}_{T_n}}(T^c > nt)\right]. \tag{10}$$

The term inside the expectation on the right-hand-side is bounded by $\sup_{|y_1-y_2|\le 2r_+} \mathbb{P}_{y_1,y_2}(T^c > nt)$, which converges to zero as $n\to\infty$ by Lemma 4.2. In addition, the distance covered by $\xi^{n,i}$ between T_n and T^c_n is of the order of $\frac{1}{\sqrt{n}}$. Indeed, in Section 5, we prove the following.

Lemma 4.3. *For any $\varepsilon > 0$ and $T > 0$,*

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \mathbb{P}\left(\sup_{\substack{s,t\in[0,nT] \\ |s-t|\le \delta n}} |\xi_s - \xi_t| > \varepsilon\sqrt{n} \right) = 0.$$

Write

$$\begin{aligned}\mathbb{P}\left(\left|\xi^{n,i}_{T^c_n} - \xi^{n,i}_{T_n}\right| > \varepsilon\right) \le \mathbb{P}&\left(\sup_{\substack{s,t\in[0,nT] \\ |s-t|\le \delta n}} |\xi_s - \xi_t| > \varepsilon\sqrt{n} \right) \\ &+ \mathbb{P}(|T^c_n - T_n| > \delta) + \mathbb{P}(T_n > nT) + \mathbb{P}(T^c_n > nT).\end{aligned}$$

Letting $n \to \infty$, the second term on the right-hand-side converges to zero by (10). So do the last two terms since both T_n and T_n^c converge in distribution as $n \to \infty$. Then letting $\delta \downarrow 0$, the first term vanishes by Lemma 4.3. As a consequence, $(\xi^{n,1}_{T_n^c}, T_n^c)$ converges in distribution (and even in probability) to $(X^1_{T^\Delta}, T^\Delta)$. Since the remaining particle after the coalescence event follows a Markov process with infinitesimal generator $\mathcal{L}$, we know by Proposition 3.3 that $(\xi^{n,1}_{T_n^c+t})_{t\geq 0}$ converges in distribution to skew Brownian motion started at $X^1_{T^\Delta}$.

This proves the convergence in distribution of $\mathcal{A}_t^n$ to $\mathcal{A}_t^\infty$ when started from two particles. For larger samples, it is enough to note that three particles (or more) almost never simultaneously come within a distance $2r_+/\sqrt{n}$ of each other. The proof of the convergence of the finite dimensional distributions and that of the second part of the statement follow the same lines, using the Markov property at suitable times. Details can be found in [BEV13b]. □

5 Convergence to skew Brownian motion

We now give the proof of Proposition 3.3. The arguments are adapted from the work of Iksanov and Pilipenko [IP16]. We limit ourselves to the one dimensional case for the proof, but the generalisation to higher dimensions is straightforward. Iksanov and Pilipenko treat the case of a discrete time Markov chain on $\mathbb{Z}$ which behaves like a simple random walk outside a bounded region centered at the origin. We extend their proof to continuous time jump Markov processes with continuous state space.

5.1 Proof of Proposition 3.3

Recall that $(\xi_t)_{t\geq 0}$ is a Markov process with generator $\mathcal{L}$ given by (2) and $\xi^n(t) = \frac{1}{\sqrt{n}}\xi_{nt}$.

As announced above, we restrict ourselves to $d = 1$ and we follow the lines of [IP16]. Set

$$\tilde{X}^{\pm}(t) = \pm\xi_t \, \mathbb{1}_{\{\pm\xi_t > r_+\}}$$

and

$$\begin{aligned}
\tau_0^{\pm} &= \inf\{t > 0 : |\xi_t| \leq r_+\}, \\
\sigma_k^{\pm} &= \inf\left\{t > \tau_k^{\pm} : \pm\xi_t > r_+\right\}, && k \geq 0, \\
\tau_{k+1}^{\pm} &= \inf\left\{t > \sigma_k^{\pm} : \pm\xi_t \leq r_+\right\}, && k \geq 0.
\end{aligned}$$

One can then write the decomposition (see formula 2.1 in [IP16])

$$\tilde{X}^{\pm}(t) = \tilde{X}^{\pm}(0) + M^{\pm}(t) + L^{\pm}(t) \mp \sum_{i\geq 0} \xi(\tau_i^{\pm}) \mathbb{1}_{\left\{\tau_i^{\pm} \leq t < \sigma_i^{\pm}\right\}} \tag{11}$$

with

$$\begin{aligned} M^{\pm}(t) &= \pm \int_0^t \mathbb{1}_{\{\pm\xi(s^-) > r_+\}} d\xi_s, \\ L^{\pm}(t) &= \pm \sum_{i\geq 0} \left(\xi(\sigma_i^{\pm}) - \xi(\tau_i^{\pm})\right) \mathbb{1}_{\left\{\sigma_i^{\pm} \leq t\right\}}. \end{aligned}$$

Also set

$$M_n^{\pm}(t) = \frac{1}{\sqrt{n}} M^{\pm}(nt), \qquad\qquad L_n^{\pm}(t) = \frac{1}{\sqrt{n}} L^{\pm}(nt).$$

Let $\xi_t^+ = \xi_t \vee 0$ and $\xi_t^- = (-\xi_t) \vee 0$. The following now holds.

Lemma 5.1. *For any fixed $T > 0$, the sequence of random variables $(\xi_n^{\pm}$, $M_n^{\pm}$, $L_n^{\pm})_{n\geq 1}$ is tight in $\mathrm{D}\left([0,T], \mathbb{R}^6\right)$. Furthermore, any limit point $(X_\infty^{\pm}$, $M_\infty^{\pm}$, $L_\infty^{\pm})$ of the sequence is a continuous process satisfying*

$$\int_0^T \mathbb{1}_{\left\{X_\infty^{\pm}(t) = 0\right\}} dt = 0, \quad a.s. \tag{12}$$

Lemma 5.2. *Let $(X_\infty^{\pm}, M_\infty^{\pm}, L_\infty^{\pm})$ be the limit point of a converging subsequence of $(\xi_n^{\pm}, M_n^{\pm}, L_n^{\pm})$ in $\mathrm{D}\left([0,T], \mathbb{R}^6\right)$. Then*

i) the processes $L_\infty^{\pm}$ are non-decreasing almost surely and satisfy

$$\int_0^T \mathbb{1}_{\left\{X_\infty^{\pm}(t) > 0\right\}} dL_\infty^{\pm}(t) = 0 \quad a.s.$$

ii) the processes $M_\infty^{\pm}$ are continuous $\mathcal{F}_t$-martingales with $\mathcal{F}_t = \sigma(X_\infty^{\pm}(s)$, $L_\infty^{\pm}(s)$, $M_\infty^{\pm}(s)$, $s \in [0,t])$ with predictable quadratic variation

$$\left\langle M_\infty^{\pm} \right\rangle_t = \sigma_{\pm}^2 \int_0^t \mathbb{1}_{\left\{X_\infty^{\pm}(s) > 0\right\}} ds$$

where $\sigma_{\pm}^2 = u \frac{2r_{\pm}^2}{d+2}$.

Lemma 5.3. *There exists $\beta \in [-1,1]$ such that, for $t \geq 0$,*

$$L^+_\infty(t) = \frac{1+\beta}{1-\beta} L^-_\infty(t)$$

almost surely.

Proposition 3.3 follows from the above lemmas and Proposition 2.1 in [IP16]. Lemma 5.1 is proved in Subsection 5.3. The proof of Lemma 5.2 does not differ from the one given for Lemma 2.2 in [IP16] and we omit the details. The proof of Lemma 5.3 is given in Subsection 5.4.

5.2 Occupation time of the boundary

We begin with the following result controlling the time spent by $(\xi_t)_{t\geq 0}$ in the region $[-r_+, r_+]$.

Lemma 5.4. *For $t \geq 0$, define $\nu(t) = \int_0^t \mathbb{1}_{\{|\xi_s| \leq r_+\}} ds$. Then*

i) $\lim_{t\to\infty} \nu(t) = +\infty$ *almost surely,*

ii) $\sup_{x\in\mathbb{R}} \mathbb{E}_x [\nu(t)] = \mathcal{O}\left(\sqrt{t}\right)$ *a.s. as $t \to \infty$.*

Proof. The fact that $\nu(t) \to \infty$ as $t \to \infty$ is well known. Set $\zeta_0 = 0$ and

$$\begin{aligned} \varsigma_i &= \inf\{t > \zeta_{i-1} : |\xi_t| \leq r_+\}, && i \geq 1, \\ \zeta_i &= \inf\{t > \varsigma_i : |\xi_t| > r_+\}, && i \geq 1. \end{aligned}$$

Then $\nu(t)$ can be written as the sum of the lengths of the excursions inside $[-r_+, r_+]$ up to time t,

$$\nu(t) = \sum_{i\geq 1} (\zeta_i \wedge t - \varsigma_i \wedge t).$$

Hence

$$\mathbb{E}_x [\nu(t)] \leq \mathbb{E}_x \left[\sum_{i\geq 1} \mathbb{E}\left[\zeta_i - \varsigma_i \mid \mathcal{F}_{\varsigma_i}\right] \mathbb{1}_{\{\varsigma_i \leq t\}} \right].$$

Noting that there exists $\varepsilon > 0$ such that $\mathbb{P}\left(|\xi(t+dt)| > r_+ \mid \xi_t = x\right) \geq \varepsilon dt$ for all $|x| \leq r_+$, we see that $\zeta_i - \varsigma_i$ is stochastically dominated by an exponential random variable with parameter ε. Hence

$$\mathbb{E}_x [\nu(t)] \leq \frac{1}{\varepsilon} \mathbb{E}_x \left[\sum_{i\geq 1} \mathbb{1}_{\{\varsigma_i \leq t\}} \right].$$

In addition, the number of visits to $[-r_+, r_+]$ before time t is less than the number of visits to this set before the first excursion longer than t, *i.e.*

$$\sum_{i\geq 1} \mathbb{1}_{\{\varsigma_i \leq t\}} \leq m(t) := \inf\{i \geq 1 : \varsigma_{i+1} - \zeta_i > t\}.$$

Let $(W_t)_{t\geq 0}$ be a continuous time random walk on $\mathbb{R}$ with jump rate u and independent increments distributed according to

$$\frac{|B(0,1) \cap B(y,1)|}{V_1^2} dy.$$

Then for any $x > r_+$,

$$\mathbb{P}_{\pm x}\left(\varsigma_1 - \zeta_0 > t\right) \geq \mathbb{P}_0\left(\inf_{0\leq s\leq t} W_s \geq 0\right).$$

(Notice that the right-hand-side isn't changed if W is replaced by $r_\pm W$.) As a result $m(t)$ is stochastically dominated by a geometric random variable with parameter

$$p(t) = \mathbb{P}_0\left(\inf_{0\leq s\leq t} W_s \geq 0\right).$$

Furthermore, there exists $\eta > 0$ such that, for all $t \geq 0$, $p(t) \geq \frac{\eta}{\sqrt{t}}$, (see pp. 381-382 in [BGT89] or equations (3.4) and (3.5) in [IP16]). As a result,

$$\mathbb{E}_x\left[\nu(t)\right] \leq \frac{1}{\varepsilon p(t)} \leq \frac{\sqrt{t}}{\varepsilon\eta}.$$

$\square$

5.3 Tightness of $(\xi_n^\pm,\ M_n^\pm,\ L_n^\pm)_{n\geq 1}$

Let us now give the proof of Lemma 5.1. To prove that the sequence $(\xi_n^\pm,$ $M_n^\pm$, $L_n^\pm)_{n\geq 1}$ is tight in $\mathrm{D}\left([0,T], \mathbb{R}^6\right)$, we use the following criterion proved by Aldous [Ald78].

Theorem 3 (Aldous [Ald78])**.** *Suppose $(X_n, n \geq 0)$ is a sequence of random variables taking values in $\mathrm{D}\left([0,T], \mathbb{R}\right)$ such that*

i) $(X_n(0), n \geq 0)$ and $\left(\sup_{t\geq 0} |X_n(t) - X_n(t^-)|\,, n \geq 0\right)$ are tight in $\mathbb{R}$, *ii) for any sequence $\{\tau_n, \delta_n\}$ such that τ_n is a stopping time with respect to the natural filtration of X_n and $\delta_n \in [0,1]$ is a constant such that $\delta_n \to 0$ as $n \to \infty$,*

$$X_n(\tau_n + \delta_n) - X_n(\tau_n) \xrightarrow[n\to\infty]{\mathcal{P}} 0.$$

Then $(X_n, n \geq 0)$ is tight in $\mathrm{D}([0,T], \mathbb{R})$.

Proof of Lemma 5.1. From (11), and the fact that

$$\left| \sum_{i\geq 0} \xi(\tau_i^{\pm}) \mathbb{1}_{\{\tau_i^{\pm} \leq t < \sigma_i^{\pm}\}} \right| \leq r_+,$$

it is enough to prove the tightness of $\xi_n^{\pm}$ and $M_n^{\pm}$. We use Aldous' criterion to prove that $M_n^{\pm}$ is tight and then we use the fact that the increments of ξ are bounded by those of $M := M^+ - M^-$ (equation (13) below) to show that ξ_n is tight.

From the definition of ξ, we have $M_n^{\pm}(0) = 0$ and

$$\sup_{t\geq 0} \left| M_n^{\pm}(t) - M_n^{\pm}(t_-) \right| \leq \frac{2r_+}{\sqrt{n}}.$$

Moreover, for any stopping time T and $\delta > 0$, since outside $[-r_+, r_+]$, ξ behaves as a simple random walk,

$$\mathbb{E}\left[\left(M_n^{\pm}(T+\delta) - M_n^{\pm}(T) \right)^2 \right] \leq \sigma_{\pm}^2 \delta.$$

The assumptions of Theorem 3 are thus satisfied, proving the tightness of $(M_n^{\pm})_n$.

Now take $0 \leq s \leq t$. If ξ does not visit $[-r_+, r_+]$ between time s and time t, then $\xi_t - \xi_s = M(t) - M(s)$. If it does visit this set, then let α be the first time ξ enters $[-r_+, r_+]$ after time s and θ the last time ξ leaves this set before time t. Then

$$\begin{aligned} |\xi_t - \xi_s| &\leq |\xi_t - \xi_\theta| + |\xi_\theta - \xi_\alpha| + |\xi_\alpha - \xi_s| \\ &\leq 4r_+ + |M(t) - M(\theta)| + |M(\alpha) - M(s)|. \end{aligned}$$

As a result, for $\delta > 0$,

$$\sup_{|s-t|\leq \delta n} |\xi_s - \xi_t| \leq 4r_+ + 2 \sup_{|s-t|\leq \delta n} |M(s) - M(t)|. \tag{13}$$

This bound is proved in [IP16] (equation (3.10)).

The tightness of $(\xi_n)_n$ then follows from that of $(M_n^{\pm})_n$ by writing

$$\begin{aligned}\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\Big(&\sup_{\substack{|s-t|\le\delta n\\ s,t\in[0,nT]}}|\xi_s-\xi_t|>\varepsilon\sqrt{n}\Big)\\ &\le\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\Big(4r_+ + 2\sup_{\substack{|s-t|\le\delta n\\ s,t\in[0,nT]}}|M(s)-M(t)|>\varepsilon\sqrt{n}\Big)=0. \quad (14)\end{aligned}$$

It remains to prove (12). Note that any limit point $(X_\infty^{\pm}, M_\infty^{\pm}, L_\infty^{\pm})$ satisfies

$$X_\infty(t)=X_\infty^+(t)-X_\infty^-(t)=M_\infty^+(t)-M_\infty^-(t)+L_\infty^+(t)-L_\infty^-(t)=M_\infty(t)+L_\infty(t).$$

From the definition of $M_n^{\pm}$ and Lemma 5.4, one shows, as in [IP16], that M_∞ is a stochastic integral with respect to standard Brownian motion $(B_t)_{t\ge 0}$

$$M_\infty(t)=\int_0^t \sigma(X_\infty(s))dB_s.$$

In addition, $L_\infty^{\pm}$ is a continuous process with locally bounded variation. As a result $\langle X_\infty\rangle_t=\langle M_\infty\rangle_t$ and (12) follows from the occupation density formula. $\square$

Note that (14) proves Lemma 4.3.

5.4 The left and right local time at zero of $(\xi_t)_{t\ge 0}$

The proof of Lemma 5.3 is adapted from that of Lemma 2.3 in [IP16]. Recall the expression for the left and right local time of $(\xi_t)_{t\ge 0}$,

$$L^{\pm}(t)=\pm\sum_{i\ge 0}(\xi(\sigma_i^{\pm})-\xi(\tau_i^{\pm}))\mathbb{1}_{\{\sigma_i^{\pm}\le t\}}.$$

For any particular visit of ξ to $[-r_+,r_+]$, the value of $\xi(\sigma_i^{\pm})-\xi(\tau_i^{\pm})$ depends on the value of ξ when it enters this set. However, over many visits to $[-r_+,r_+]$, $L^{\pm}(t)$ only records an average of these values. The "typical" value of $\xi(\sigma_i^{\pm})-\xi(\tau_i^{\pm})$ can thus be expressed with the help of the stationary distribution of the process describing the visits of ξ to $[-r_+,r_+]$ (Y below). The left and right local time of ξ then become asymptotically proportional to the occupation time of the boundary $\nu(t)$, with different coefficients whose expressions can be found below.

Recall from Lemma 5.4 that $\nu(t)=\int_0^t\mathbb{1}_{\{|\xi_s|\le r_+\}}ds$ and set, for $t\ge 0$,

$$\alpha(t)=\inf\{\alpha>0:\nu(\alpha)>t\}.$$

Define $Y(t) = \xi(\alpha(t))$ for $t \geq 0$. The process $(Y(t))_{t\geq 0}$ is a jump Markov process taking values in $[-r_+, r_+]$, describing the values taken by ξ inside this region. Let $\bar{\alpha}$ denote the left-continuous version of α, *i.e.* for $t \geq 0$,

$$\bar{\alpha}(t) = \sup\{\alpha \geq 0 : \nu(\alpha) < t\}.$$

If $t \geq 0$ is such that $\bar{\alpha}(t) \neq \alpha(t)$, then ξ makes an excursion outside $[-r_+, r_+]$ between time $\bar{\alpha}(t)$ and time $\alpha(t)$.

Let $V^{\pm}$ be defined by

$$V^{\pm}(t) = \pm \sum_{0<s\leq t} (Y(s) - Y(s^-)) \pm \sum_{0<s\leq t} (\xi(\bar{\alpha}(s)) - \xi(\alpha(s)))\mathbb{1}_{\{\pm\xi(\bar{\alpha}(s))>r_+\}}.$$

Lemma 5.5. *There exist $C > 0$ and $\beta \in [-1, 1]$ such that, as $t \to \infty$,*

$$\frac{1}{t} V^{\pm}(t) \to C(1 \pm \beta)$$

almost surely.

To prove Lemma 5.5, we use the Nummelin splitting technique [Num78] to turn Y into a renewal process. We can then build its stationary probability distribution (see Subsection 5.5), following Chapter 6.8 in [Dur10]. Lemma 5.5 is then reduced to the strong law of large numbers for renewal processes. The detailed argument is given in Subsection 5.6.

Proof of Lemma 5.3. We first show that $V^{\pm}(\nu(t))$ provides a good approximation of $L^{\pm}(t)$ and then conclude with the help of Lemma 5.5. Note that $\pm\xi(\bar{\alpha}(s)) > r_+$ with $s > 0$ if and only if $\bar{\alpha}(s) = \sigma_i^{\pm}$ for some $i \geq 1$, and in this case, $\alpha(s) = \tau_{i+1}^{\pm}$. In addition, $s \leq \nu(t)$ if and only if $\bar{\alpha}(s) \leq t$, as a result

$$V^{\pm}(\nu(t)) = \pm(Y(\nu(t)) - Y(0)) \pm \sum_{i\geq 1} (\xi(\sigma_i^{\pm}) - \xi(\tau_{i+1}^{\pm}))\mathbb{1}_{\{\sigma_i^{\pm}\leq t\}}.$$

Hence

$$\begin{aligned} \left|V^{\pm}(\nu(t)) - L^{\pm}(t)\right| \leq |Y(\nu(t))| + |Y(0)| + \left|\xi(\sigma_0^{\pm})\right| + \left|\xi(\tau_0^{\pm})\right| + \left|\xi(\tau_1^{\pm})\right| \\ + \sum_{i\geq 2} \left|\xi(\tau_i^{\pm})\right| \mathbb{1}_{\{\sigma_{i-1}^{\pm}\leq t<\sigma_i^{\pm}\}}. \end{aligned}$$

Since $|Y(t)| \leq r_+$, $\left|\xi(\tau_i^{\pm})\right| \leq r_+$ and $\left|\xi(\sigma_i^{\pm})\right| \leq 3r_+$,

$$\left|V^{\pm}(\nu(t)) - L^{\pm}(t)\right| \leq 8r_+.$$

From this, Lemma 5.5, and using Lemma 5.4.i, we obtain

$$\lim_{t\to\infty} \frac{L^+(t)}{L^-(t)} = \frac{1+\beta}{1-\beta}. \tag{15}$$

$\square$

5.5 The Stationary distribution of Y

Let $\Phi^Y : [-r_+, r_+]^2 \to \mathbb{R}_+$ be such that

$$\mathcal{L}^Y f(x) = u \int_{[-r_+,r_+]} \Phi^Y(x,y)(f(y) - f(x))dy$$

is the infinitesimal generator of $(Y(t))_{t\geq 0}$. Clearly, from (3), for $x, y \in [-r_+, r_+]$,

$$\Phi^Y(x,y) \geq \Phi(x,y).$$

Note that Φ is continuous on the compact set $[-r_+, r_+]^2$ and that it stays strictly positive on sets of the form $U_{a,b} = [-r_+, r_+] \times [a, b]$ with $-r_+ < a < b < -r_+ + 2r_-$. Fix one such set $U_{a,b}$ and set $\Phi_{min} = \inf_{U_{a,b}} \Phi > 0$. As a result

$$\Phi^\varepsilon(x,y) := \Phi^Y(x,y) - \frac{\varepsilon}{b-a} \mathbb{1}_{\{[a,b]\}}(y) \geq 0$$

for $\varepsilon = (b-a)\Phi_{min} > 0$.

We now follow Chapter 6.8 of [Dur10] to build the (unique) stationary probability measure of Y. Define an operator $\mathcal{L}^Z$ on real-valued functions f on $[-r_+, r_+] \cup \{\partial\}$ by

$$\mathcal{L}^Z f(x) = \begin{cases} u \int_{[-r_+,r_+]} \Phi^\varepsilon(x,y)(f(y) - f(x))dy + u\varepsilon(f(\partial) - f(x)) & \\ & \text{if } x \in [-r_+, r_+], \\ \dfrac{1}{b-a} \int_a^b (f(y) - f(\partial))dy & \text{if } x = \partial, \end{cases} \tag{16}$$

and let $(Z(t))_{t\geq 0}$ be a Markov process on $[-r_+, r_+] \cup \{\partial\}$ with generator $\mathcal{L}^Z$. Let

$$\lambda(t) = \inf \left\{ \lambda > 0 : \int_0^\lambda \mathbb{1}_{\{Z(s)\neq\partial\}} ds > t \right\}, \tag{17}$$

then

$$(Z(\lambda(t)), t \geq 0) \stackrel{d}{=} (Y(t), t \geq 0).$$

Set $E_0 = 0$ and, for $k \geq 0$,

$$R_k = \inf\{t \geq E_k : Z(t) = \partial\},$$
$$E_{k+1} = \inf\{t \geq R_k : Z(t) \neq \partial\}.$$

Then $R_k - E_k$ is an exponential random variable with parameter $u\varepsilon$ for all $k \geq 1$ and ∂ is a positive recurrent state for Z. We can then use this fact to build a stationary probability measure for Y. Let $\mathbb{E}_\partial$ denote the expectation with respect to $\mathbb{P}(\cdot \mid Z(0) = \partial)$.

Lemma 5.6. *The measure π defined by*

$$\int_{[-r_+,r_+]} f(x)\pi(dx) = u\varepsilon\mathbb{E}_\partial\left[\int_{E_1}^{R_1} f(Z(s))ds\right] \tag{18}$$

is an invariant probability measure for $(Y(t))_{t\geq 0}$.

Since Y is irreducible with respect to the Lebesgue measure on $[-r_+, r_+]$, *i.e.* any two sets of positive Lebesgue measure communicate with each other (see [Dob40] or [Num04]), π is unique.

Proof. Let $f : [-r_+, r_+] \cup \{\partial\} \to \mathbb{R}$ be bounded and measurable. Since $\mathcal{L}^Z$ is the generator of Z, by the optional stopping time theorem,

$$\mathbb{E}_\partial\left[f(Z(R_1)) - f(Z(E_1)) - \int_{E_1}^{R_1} \mathcal{L}^Z f(Z(s))ds\right] = 0.$$

By the definition of R_1 and E_1,

$$f(Z(R_1)) = f(\partial), \qquad \mathbb{E}_\partial\left[f(Z(E_1))\right] = \frac{1}{b-a}\int_a^b f(y)dy.$$

And by the definition of $\mathcal{L}^Z$ in (16),

$$\mathcal{L}^Z f(x) = \mathcal{L}^Y f(x) + u\varepsilon\left(f(\partial) - \frac{1}{b-a}\int_a^b f(y)dy\right).$$

Combining these equalities with the fact that $\mathbb{E}_\partial\left[R_1 - E_1\right] = \dfrac{1}{u\varepsilon}$, we obtain

$$\int_{[-r_+,r_+]} \mathcal{L}^Y f\, d\pi = u\varepsilon\mathbb{E}_\partial\left[\int_{E_1}^{R_1} \mathcal{L}^Y f(Z(s))ds\right] = 0.$$

□

Furthermore, using the fact that $\Phi(x, y) = \Phi(y, x)$, we are able to identify π.

Lemma 5.7. *The measure π is the uniform probability distribution on $[-r_+, r_+]$.*

Proof. For f and g two bounded and measurable functions on $[-r_+, r_+]$, let

$$\langle f, g\rangle_\pi = \int_{-r_+}^{r_+} f(x)g(x)dx.$$

We want to show
$$\left\langle \mathcal{L}^Y f, g \right\rangle_\pi = \left\langle f, \mathcal{L}^Y g \right\rangle_\pi . \tag{19}$$
For $f : [-r_+, r_+] \to \mathbb{R}$ and $x \in \mathbb{R}$, let
$$Ef(x) := \mathbb{E}_x \left[f(Y(0)) \right] \tag{20}$$
and note that $\mathcal{L}^Y f(x) = \mathcal{L} E f(x)$. In addition, since $\Phi(x, y) = \Phi(y, x)$, for any $f, g \in L^2(\mathbb{R})$,
$$\langle \mathcal{L} f, g \rangle_{\mathbb{R}} = \langle f, \mathcal{L} g \rangle_{\mathbb{R}} .$$
However, $Ef \notin L^2(\mathbb{R})$. For $n \geq 1$, define
$$\mathcal{T}^n = \inf\{t > 0 : |\xi(t)| \leq r_+ \text{ or } |\xi(t)| \geq n\}$$
and, for $f : [-r_+, r_+] \to \mathbb{R}$ bounded,
$$E^n f(x) = \mathbb{E}_x \left[f(\xi(\mathcal{T}^n)) \mathbb{1}_{\{|\xi(\mathcal{T}^n)| \leq r_+\}} \right] .$$
Then
$$E^n f(x) = \begin{cases} f(x) & \text{if } |x| \leq r_+ \\ 0 & \text{if } |x| \geq n, \end{cases} \tag{21}$$
$$\mathcal{L} E^n f(x) = 0 \qquad \text{if } r_+ < |x| < n. \tag{22}$$
In particular, $E^n f \in L^2(\mathbb{R})$. As a result,
$$\langle \mathcal{L} E^n f, E^n g \rangle_{\mathbb{R}} = \langle E^n f, \mathcal{L} E^n g \rangle_{\mathbb{R}} . \tag{23}$$
In addition, from (21) and (22),
$$\begin{aligned} \langle \mathcal{L} E^n f, E^n g \rangle_{\mathbb{R}} &= \langle \mathcal{L} E^n f, E^n g \rangle_\pi \\ &= \langle \mathcal{L} E^n f, g \rangle_\pi . \end{aligned}$$
Finally, for any $x \in \mathbb{R}$, $\mathcal{T}^n \underset{n \to \infty}{\longrightarrow} Y(0) = \inf\{t > 0 : |\xi(t)| \leq r_+\}$ almost surely. By dominated convergence, for $x \in \mathbb{R}$, $E^n f(x) \underset{n \to \infty}{\longrightarrow} Ef(x)$ and, using dominated convergence once more, we obtain
$$\langle \mathcal{L} E^n f, g \rangle_\pi \underset{n \to \infty}{\longrightarrow} \langle \mathcal{L} E f, g \rangle_\pi .$$
Applying the same argument to the right-hand-side of (23), we obtain (19). As a result the uniform measure on $[-r_+, r_+]$ is invariant for Y. Since Y is irreducible with respect to the Lebesgue measure and π defined in (18) is absolutely continuous with respect to the Lebesgue measure, π is the uniform probability measure on $[-r_+, r_+]$. $\square$

5.6 Proof of Lemma 5.5

Now that we have built the stationary probability measure for Y, we can prove Lemma 5.5, adapting the arguments of [IP16, Lemma 2.3]. The proof is an application of the law of large numbers to the renewal process $(Z(t))_{t\geq 0}$.

Recall that $Y(t) = Z(\lambda(t))$ with λ defined in (17). From the definition of λ, for $t \geq 0$,

$$\lambda^{-1}(t) = \int_0^t \mathbb{1}_{\{Z_s \neq \partial\}} ds.$$

For $k \geq 0$, set

$$\tilde{R}_k = \lambda^{-1}(R_k) = \lambda^{-1}(E_{k+1}),$$

and

$$V_k^{\pm} = V^{\pm}(\tilde{R}_{k+1}) - V^{\pm}(\tilde{R}_k).$$

Then $V_0^{\pm}, V_1^{\pm}, \ldots$ are independent and for all $k \geq 1$, $V_k^{\pm}$ is distributed as $V_1^{\pm}$ under $\mathbb{E}_{\partial}$. Recall the definition of the operator E in (20) and set $\iota(x) = x$ for $x \in \mathbb{R}$. We prove the following lemma at the end of this subsection.

Lemma 5.8.

$$\mathbb{E}_{\partial}\left[V_1^{\pm}\right] = \frac{1}{2r_+\varepsilon} \int_{-r_+}^{r_+} \int_{\mathbb{R}} \Phi(x,y)(y - E\iota(y))^{\pm} dy dx,$$

where $(\cdot)^+$ (resp $(\cdot)^-$) denotes the positive (resp. negative) part.

Proof of Lemma 5.5. Setting $N(t) = \max\{k \geq 0 : \tilde{R}_{k+1} \leq t\}$, by the strong law of large numbers for renewal processes, we have

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{N(t)} V_k^{\pm} = \lim_{t\to\infty} \frac{N(t)}{t} \frac{1}{N(t)} \sum_{k=1}^{N(t)} V_k^{\pm} = u\varepsilon \mathbb{E}_{\partial}\left[V_1^{\pm}\right] \quad \text{a.s.} \tag{24}$$

Moreover,

$$\begin{aligned} V^{\pm}(t) - \sum_{k=0}^{N(t)} V_k^{\pm} &= V^{\pm}(t) - V^{\pm}(\tilde{R}_{N(t)+1}) \\ &= \pm(Y(t) - Y(\tilde{R}_{N(t)+1})) \\ &\quad \pm \sum_{\tilde{R}_{N(t)+1} \leq s < t} (\xi(\bar{\alpha}(s)) - \xi(\alpha(s)))\mathbb{1}_{\{\pm\xi(\bar{\alpha}(s)) > r_+\}}. \end{aligned}$$

Taking absolute values on both sides, we have

$$\left|V^{\pm}(t)-\sum_{k=0}^{N(t)}V_k^{\pm}\right| \leq \left|Y(t)+Y(\tilde{R}_{N(t)+1})\right| \\ + \sum_{\tilde{R}_{N(t)+1}\leq s<t} \left|\xi(\bar{\alpha}(s))-\xi(\alpha(s))\right| \mathbb{1}_{\{\pm\xi(\bar{\alpha}(s))>r_+\}}$$

Since when $\pm\xi(\bar{\alpha}(s)) > r_+$, $\pm(\xi(\bar{\alpha}(s))-\xi(\alpha(s))) \geq 0$, we can add the terms for which $t \leq s < \tilde{R}_{N(t)+2}$ on the right-hand-side,

$$\left|V^{\pm}(t)-\sum_{k=0}^{N(t)}V_k^{\pm}\right| \leq \left|Y(t)+Y(\tilde{R}_{N(t)+1})\right| \\ + \left|\sum_{\tilde{R}_{N(t)+1}\leq s<\tilde{R}_{N(t)+2}} (\xi(\bar{\alpha}(s))-\xi(\alpha(s)))\mathbb{1}_{\{\pm\xi(\bar{\alpha}(s))>r_+\}}\right|$$

Adding and subtracting $Y(\tilde{R}_{N(t)+2})-Y(\tilde{R}_{N(t)+1})$ inside the absolute value, we obtain,

$$\left|V^{\pm}(t)-\sum_{k=0}^{N(t)}V_k^{\pm}\right| \leq 4r_+ + \left|V_{N(t)+1}^{\pm}\right|.$$

Hence, since the $V_k^{\pm}$ are identically distributed for $k \geq 1$,

$$\left|\frac{1}{t}V^{\pm}(t)-\frac{1}{t}\sum_{k=0}^{N(t)}V_k^{\pm}\right| \leq \frac{4r_+}{t} + \frac{1}{t}\left|V_{N(t)+1}^{\pm}\right|.$$

The right-hand-side converges to zero almost surely as $t \to \infty$ since the $V_k^{\pm}$ are identically distributed for $k \geq 1$. As a result, from (24)

$$\lim_{t\to\infty}\frac{1}{t}V^{\pm}(t) = u\varepsilon\mathbb{E}_{\partial}\left[V_1^{\pm}\right]. \tag{25}$$

The statement of Lemma 5.5 now follows from Lemma 5.8 by taking

$$\beta = \frac{\int_{-r_+}^{r_+}\int_{\mathbb{R}}\Phi(x,y)(y-E\iota(y))dydx}{\int_{-r_+}^{r_+}\int_{\mathbb{R}}\Phi(x,y)\left|y-E\iota(y)\right|dydx}. \tag{26}$$

□

We now prove Lemma 5.8.

Proof of Lemma 5.8. Define

$$h^{\pm}(x) = \pm u \int_{\mathbb{R}} \Phi(x,y) \mathbb{1}_{\{\pm y \leq r_+\}} (E\iota(y) - x) dy \\ \pm u \int_{\mathbb{R}} \Phi(x,y) \mathbb{1}_{\{\pm y > r_+\}} (y - x) dy. \quad (27)$$

Writing

$$V^{\pm}(t) = \pm \sum_{0<s\leq t} (Y(s) - Y(s^-)) \mathbb{1}_{\{\pm \xi(\bar{\alpha}(s)) \leq r_+\}} \\ \pm \sum_{0<s\leq t} (\xi(\bar{\alpha}(s)) - Y(s^-)) \mathbb{1}_{\{\pm \xi(\bar{\alpha}(s)) > r_+\}},$$

it follows that

$$V^{\pm}(t) - \int_0^t h^{\pm}(Y(s)) ds$$

is a martingale with respect to the filtration associated with $(Y(t))_{t\geq 0}$. As a result,

$$\begin{aligned} u\varepsilon \mathbb{E}_{\partial} \left[V_1^{\pm} \right] &= u\varepsilon \mathbb{E}_{\partial} \left[\int_{\lambda^{-1}(E_1)}^{\lambda^{-1}(R_1)} h^{\pm}(Y(s)) ds \right] \\ &= u\varepsilon \mathbb{E}_{\partial} \left[\int_{E_1}^{R_1} h^{\pm}(Z(s)) ds \right] \\ &= \left\langle h^{\pm}, \pi \right\rangle, \end{aligned} \quad (28)$$

by (18). Note that since $E\iota(y) = y$ when $|y| \leq r_+$, $h^{\pm}$ can be written as

$$\begin{aligned} h^{\pm}(x) &= \pm u \int_{\mathbb{R}} \Phi(x,y)(E\iota(y) - E\iota(x)) dy \\ &\qquad \pm u \int_{\mathbb{R}} \Phi(x,y) \mathbb{1}_{\{\pm y > r_+\}} (y - E\iota(y)) dy \\ &= \pm \mathcal{L} E\iota(x) + u \int_{\mathbb{R}} \Phi(x,y)(y - E\iota(y))^{\pm} dy, \end{aligned}$$

Besides, we noted above that $\mathcal{L}Ef = \mathcal{L}^Y f$, hence $\langle \mathcal{L}E\iota, \pi \rangle = 0$. Furthermore, from Lemma 5.7,

$$\left\langle h^{\pm}, \pi \right\rangle = \frac{u}{2r_+} \int_{-r_+}^{r_+} \int_{\mathbb{R}} \Phi(x,y)(y - E\iota(y))^{\pm} dy dx.$$

This, together with (28) concludes the proof of Lemma 5.8. □

A Inequalities for hitting times

Proof of Proposition 4.1. We first prove the inequality for T_O^n. Suppose that $\limsup T_O^n > T_O$ and fix $\varepsilon > 0$ such that $T_O + \varepsilon \leq \limsup T_O^n$. There exists a subsequence $(n_k)_k$ such that for all $k \in \mathbb{N}$, $T_O^{n_k} \geq T_O + \varepsilon$. By the definition of T_O, there exists $t \in [T_O, T_O + \varepsilon)$ such that $f(t) \in O$. By the convergence of f_n to f, $f_{n_k}(t)$ converges to $f(t)$ as $k \to \infty$. Since $f(t) \in O$ which is open, for k large enough, $f_{n_k}(t) \in O$ and $T_O^{n_k} \leq t$, leading to a contradiction.

For the second inequality, suppose that $\liminf T_F^n < T_F$ and take $\varepsilon > 0$ such that $\liminf T_F^n \leq T_F - 2\varepsilon$. There exists a subsequence $(n_k)_k$ such that for all $k \in \mathbb{N}$, $T_F^{n_k} \leq T_F - 2\varepsilon$. Since f is continuous, the image of $[0, T_F - \varepsilon]$ by f is a compact set which does not intersect F, hence there exists $\eta > 0$ such that its η-neighbourhood is in $\mathbb{R}^d \setminus F$. By the locally uniform convergence of f_n to f, $\sup\{|f_{n_k}(t) - f(t)| : t \in [0, T_F - \varepsilon]\}$ converges to zero as $k \to \infty$. Taking k large enough that this quantity is smaller that η, we have that $f_{n_k}(t) \notin F$ for $t \in [0, T_F - \varepsilon]$. Hence $T_F^{n_k} \geq T_F - \varepsilon$, which is a contradiction. □

References

[ABT+11] Thilanka Appuhamillage, Vrushali Bokil, Enrique Thomann, Edward Waymire, and Brian Wood. Occupation and local times for skew Brownian motion with applications to dispersion across an interface. *The Annals of Applied Probability*, 21(1):183–214, 2011.

[ADN99] Bruce P Ayati, Todd F Dupont, and Thomas Nagylaki. The Influence of Spatial Inhomogeneities on Neutral Models of Geographical Variation IV. Discontinuities in the Population Density and Migration Rate. *Theoretical Population Biology*, 56(3):337–347, 1999.

[Ald78] David Aldous. Stopping times and tightness. *The Annals of Probability*, 6(2):335–340, 1978.

[BDE02] Nick H. Barton, Frantz Depaulis, and Alison M. Etheridge. Neutral evolution in spatially continuous populations. *Theoretical population biology*, 61(1):31–48, 2002.

[BEV10] Nick H. Barton, Alison M. Etheridge, and Amandine Véber. A new model for evolution in a spatial continuum. *Electronic Journal of Probability*, 15(7):162–216, 2010.

[BEV13a] Nick H. Barton, Alison M. Etheridge, and Amandine Véber. Modeling evolution in a spatial continuum. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01002, 2013.

[BEV13b] Nathanaël Berestycki, Alison M. Etheridge, and Amandine Véber. Large scale behaviour of the spatial Lambda-Fleming–Viot process. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 49(2):374–401, 2013.

[BGT89] Nicholas H. Bingham, Charles M. Goldie, and Jef L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge university press, Cambridge, 1989.

[BP87] Richard F. Bass and Etienne Pardoux. Uniqueness for diffusions with piecewise constant coefficients. *Probability Theory and Related Fields*, 76(4):557–572, 1987.

[Dob40] Wolfgang Doblin. Éléments d'une théorie générale des chaînes simples constantes de Markoff. *Annales scientifiques de l'École Normale Supérieure*, 57:61–111, 1940.

[DR08] Richard Durrett and Mateo Restrepo. One-dimensional stepping stone models, sardine genetics and Brownian local time. *The Annals of Applied Probability*, 18(1):334–358, 2008.

[Dur10] Richard Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge university press, Cambridge, fourth edition, 2010.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, Inc., New York, 1986.

[Eth08] Alison M. Etheridge. Drift, draft and structure: some mathematical models of evolution. In *Stochastic models in biological sciences*, volume 80 of *Banach Center Publ.*, pages 121–144. Polish Acad. Sci. Inst. Math., Warsaw, 2008.

[Eva97] Steven N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 33(3):339–358, 1997.

[EVY14] Alison Etheridge, Amandine Veber, and Feng Yu. Rescaling limits of the spatial Lambda-Fleming-Viot process with selection. *arXiv preprint arXiv:1406.5884*, 2014.

[GCR+07] Tenzin Gayden, Alicia M. Cadenas, Maria Regueiro, Nanda B. Singh, Lev A. Zhivotovsky, Peter A. Underhill, Luigi L. Cavalli-Sforza, and Rene J. Herrera. The Himalayas as a Directional Barrier to Gene Flow. *The American Journal of Human Genetics*, 80(5):884–894, 2007.

[HS81] John Michael Harrison and Lawrence A. Shepp. On skew Brownian motion. *The Annals of Probability*, 9(2):309–313, 1981.

[IM63] K. Itô and H. P. McKean. Brownian motions on a half line. *Illinois Journal of Mathematics*, 7:181–231, 1963.

[IP16] Alexander Iksanov and Andrey Pilipenko. A functional limit theorem for locally perturbed random walks. *Probability and Mathematical Statistics*, 36(2):353–368, 2016.

[KW64] Motoo Kimura and George H. Weiss. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(4):561, 1964.

[Lej06] Antoine Lejay. On the constructions of the skew Brownian motion. *Probability Surveys*, 3:413–466, 2006.

[LG84] Jean-François Le Gall. One-dimensional stochastic differential equations involving the local times of the unknown process. In *Stochastic analysis and applications*, volume 1095 of *Lecture Notes in Math.*, pages 51–82. Springer, Berlin, 1984.

[Mal48] Gustave Malécot. *Les Mathématiques de l'Hérédité*. Masson et Cie., Paris, 1948.

[MH13] Stéphanie Manel and Rolf Holderegger. Ten years of landscape genetics. *Trends in Ecology & Evolution*, 28(10):614–621, 2013.

[Nag76] Thomas Nagylaki. Clines with Variable Migration. *Genetics*, 83(4):867–886, 1976.

[Nag88] Thomas Nagylaki. The influence of spatial inhomogeneities on neutral models of geographical variation. I. Formulation. *Theoretical Population Biology*, 33(3):291–310, 1988.

[NB88] Thomas Nagylaki and Victor Barcilon. The influence of spatial inhomogeneities on neutral models of geographical variation. II. The semi-infinite linear habitat. *Theoretical Population Biology*, 33(3):311–343, 1988.

[Num78] Esa Nummelin. A splitting technique for Harris recurrent Markov chains. *Probability Theory and Related Fields*, 43(4):309–318, 1978.

[Num04] Esa Nummelin. *General irreducible Markov chains and non-negative operators*, volume 83. Cambridge University Press, 2004.

[Por79a] N. I. Portenko. Diffusion processes with generalized drift coefficients. *Theory of Probability & Its Applications*, 24(1):62–78, 1979.

[Por79b] N. I. Portenko. Stochastic differential equations with generalized drift vector. *Theory of Probability & Its Applications*, 24(2):332–347, 1979.

[RPS+06] Seth PD Riley, John P. Pollinger, Raymond M. Sauvajot, Eric C. York, Cassity Bromley, Todd K. Fuller, and Robert K. Wayne. FAST-TRACK: A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology*, 15(7):1733–1741, 2006.

[Sla87] Montgomery Slatkin. Gene flow and the geographic structure of natural populations. *Science*, 236(4803):787–792, 1987.

[SQH+03] H. Su, L. J. Qu, K. He, Z. Zhang, J. Wang, Z. Chen, and H. Gu. The Great Wall of China: a physical barrier to gene flow? *Heredity*, 90(3):212–219, 2003.

[Wal78] John B. Walsh. A diffusion with a discontinuous local time. *Astérisque*, 52(53):37–45, 1978.