Large scale behaviour of the spatial Λ-Fleming-Viot process

N. Berestycki*
Statistical Laboratory
University of Cambridge
DPMMS
Wilberforce Road
Cambridge CB3 0WB, UK

A.M. Etheridge†
Department of Statistics
University of Oxford
1 South Parks Road
Oxford OX1 3TG, UK

A. Véber‡
CMAP -École Polytechnique
Route de Saclay
91128 Palaiseau Cedex
France

November 22, 2011

*NB supported in part by EPSRC grants EP/G055068/1 and EP/I03372X/1

†AME supported in part by EPSRC Grants EP/E065945/1 and EP/G052026/1

‡AV supported in part by the *chaire Modélisation Mathématique et Biodiversité* of Veolia Environnement-École Polytechnique-Museum National d'Histoire Naturelle-Fondation X and by the ANR project MANEGE (ANR-090BLAN-0215).

Abstract

We consider the spatial Λ-Fleming-Viot process model ([BEV10]) for frequencies of genetic types in a population living in $\mathbb{R}^d$, in the special case in which there are just two types of individuals, labelled 0 and 1. At time zero, everyone in a given half-space has type 1, whereas everyone in the complementary half-space has type 0. We are concerned with patterns of frequencies of the two types at large space and time scales. We consider two cases, one in which the dynamics of the process are driven by purely 'local' events and one incorporating large-scale extinction recolonisation events. We choose the frequency of these events in such a way that, under a suitable rescaling of space and time, the ancestry of a single individual in the population converges to a symmetric stable process of index $\alpha \in (1, 2]$ (with $\alpha = 2$ corresponding to Brownian motion). We consider the behaviour of the process of allele frequencies under the same space and time rescaling. For $\alpha = 2$, and $d \geq 2$ it converges to a deterministic limit. In all other cases the limit is random and we identify it as the indicator function of a random set. In particular, there is no local coexistence of types in the limit. We characterise the set in terms of a dual process of coalescing symmetric stable processes, which is of interest in its own right. The complex geometry of the random set is illustrated through simulations.

On étudie le processus Λ-Fleming-Viot spatial ([BEV10]) modélisant les fréquences locales de types génétiques dans une population évoluant dans $\mathbb{R}^d$. On considère le cas particulier où il n'y a que deux types possibles, notés 0 et 1. Initialement, tous les individus présents dans le demi-espace des points dont la première coordonnée est négative sont de type 1, tandis que les individus présents dans le demi-espace complémentaire sont de type 0. On s'intéresse au comportement des fréquences locales sur des échelles de temps et d'espace très grandes. On considère deux cas : dans le premier, l'évolution du processus est due uniquement à des événements 'locaux'; dans le second, on incorpore des événements d'extinction et recolonisation de grande ampleur. On choisit la fréquence de ces événements de sorte qu'après une renormalisation spatiale et temporelle appropriée, la lignée ancestrale d'un individu de la population converge vers un processus α-stable symétrique, d'indice $\alpha \in (1, 2]$ (où $\alpha = 2$ correspond au mouvement brownien). On étudie l'évolution du processus des fréquences alléliques aux mêmes échelles spatio-temporelles. Lorsque $\alpha = 2$ et $d \geq 2$, celui-ci converge vers un processus déterministe. Dans tous les autres cas, le processus limite est aléatoire et on l'identifie comme la fonction indicatrice d'un ensemble aléatoire évoluant au cours du temps. En particulier, les deux types ne coexistent pas à la limite. On caractérise chaque ensemble en termes d'un processus dual constitué de mouvements stables symétriques coalescents ayant un intérêt en eux-mêmes. La géométrie complexe des ensembles limites est illustrée par des simulations.

AMS 2010 subject classifications. *Primary:* 60G57, 60J25, 92D10 ; *Secondary:* 60J75, 60G52.

Key words and phrases: Generalised Fleming-Viot process, limit theorems, duality, symmetric stable processes, population genetics.

1 Introduction

In this article, we are interested in the behaviour over large space and time scales of the *spatial Λ-Fleming-Viot process* (or SLFV) on $\mathbb{R}^d$. This process arises as a particular instance of the framework introduced in [Eth08, BEV10, BKE10] for modelling allele frequencies (that is frequencies of different genetic types) in a population that evolves in a spatial continuum. From the modelling perspective, this framework is interesting as it overcomes an obstruction to modelling biological populations in continua, dubbed 'the pain in the torus' by Felsenstein ([Fel75]), which is typified by the 'clumping and extinction' seen in spatial branching process models in low dimensions. The key idea of the SLFV framework is to base reproduction events on a space-time Poisson process rather than on individuals in the population. In this way one can define what can be thought of as a continuum version of the Kimura stepping stone model ([Kim53]) which is a widely accepted model for evolution of allele frequencies in spatially *subdivided* populations. Moreover, one can incorporate large-scale extinction-recolonisation events through a series of 'local' population bottlenecks, each affecting substantial portions of the species range. Such events dominate the demographic history of many species and, as we shall see in our results here, can have a very significant influence on patterns of allele frequencies.

From a mathematical perspective, the SLFV process is a natural extension to the spatial context of the generalised Fleming-Viot processes which can be traced to §3.1.4 of [DK99] but were first studied in detail by Bertoin & Le Gall ([BLG03]). These processes are dual to the so-called Λ-coalescents which have been the subject of intensive study since their introduction over a decade ago by Donnelly & Kurtz, Pitman and Sagitov ([DK99, Pit99, Sag99]). The duality with the generalised Fleming-Viot processes extends that between the Kingman coalescent and the Wright-Fisher diffusion and our work here will exploit a similar duality between spatial versions of the Λ-coalescents and the SLFV. One of the attractions of the SLFV processes is that they allow us to capture many of the features of Wright-Fisher noise, but in *any* spatial dimension (whereas stochastic partial differential equations driven by Wright-Fisher noise only make sense in one dimension). Thus, although they were originally motivated by purely biological considerations, we believe that these models are also of intrinsic mathematical interest.

1.1 The spatial Λ-Fleming-Viot process

First we describe the model. Each individual in the population is assigned a genetic type, from a compact space K, and a location, in $\mathbb{R}^d$. At time t, the population is represented by a measurable function $\rho_t : \mathbb{R}^d \to \mathcal{M}_1(K)$, where $\mathcal{M}_1(K)$ is the set of all probability measures on K. (In fact, as explained in §3, in defining the state space, Ξ, of the process we identify any two such functions that are equal for Lebesgue-a.e. $x \in \mathbb{R}^d$.) The interpretation of the model is as follows: the population density is uniform across $\mathbb{R}^d$ and, for each $x \in \mathbb{R}^d$, if we sample an individual from x, then its genetic type is determined by sampling from the probability measure $\rho_t(x)$.

The dynamics of the population are driven by a Poisson point process, Π, on $\mathbb{R}\times\mathbb{R}^d\times(0,\infty)\times[0,1]$, each point of which specifies a (local) extinction-recolonisation event. If $(t,x,r,u) \in \Pi$, then, at time t:

1. An extinction-recolonisation event affects the closed ball $B(x,r) \subseteq \mathbb{R}^d$, and nothing happens outside this region.

2. A parent is chosen uniformly in the ball; that is, we sample a location z uniformly at random over $B(x,r)$ and a type k according to the distribution $\rho_{t-}(z)$.

3. For each $y \in B(x,r)$ (including z), a fraction u of the local population is replaced by offspring, whose type is that of the chosen parent. That is,

$$\rho_t(y) := (1-u)\rho_{t-}(y) + u\,\delta_k.$$

Here, we are thinking of reproduction events as equivalent to (frequent) small-scale extinction-recolonisation events.

In [BEV10], the intensity measure of the Poisson point process Π has the form $dt \otimes dx \otimes \zeta(dr, du)$, thus allowing the 'impact', u, of a reproduction event to depend on its radius, r. For instance, small-scale reproduction events may affect only a tiny fraction of individuals, compared to massive extinction-recolonisation events which could wipe out most of the population in a large geographical region. Of course, we require some conditions on the intensity of Π if our process is to be well-defined: according to Theorem 4.2 of [BEV10] (stated for $d = 2$, but the proof is identical for any dimension $d \geq 1$), the corresponding spatial Λ-Fleming-Viot process is well-defined whenever

$$\int_{(0,\infty)\times[0,1]} \zeta(dr, du)\; uV_r < \infty, \tag{1}$$

where V_r denotes the volume of a d-dimensional ball of radius r.

1.2 Main results

Our previous mathematical analysis of the SLFV process ([BEV10, EV11]) has been concerned with understanding the genealogical relationships between individuals sampled from the population. Here, although studying the lineages ancestral to a sample from the population will be fundamental to our analysis, we are interested in understanding the patterns of allele frequencies that result from such a model.

Saadi ([Saa11]) considers a closely related model (which differs from ours only in that the location of the 'parent' in a reproduction event is always taken to be the centre of the event). He considers the most biologically interesting case of two spatial dimensions and, for simplicity, takes all reproduction events to have fixed size r and fixed impact $u \in (0,1]$. Notice that if a particular genetic type is present in a region at some time t, then, unless $u = 1$, it will also be there at all later times. Saadi shows that if a particular genetic type is only present in a bounded region at time zero, then, with probability one, its *range*, that is the region in which it is *ever* seen is bounded. On the other hand, the *shape* of this region will be complex. In order to try to gain some understanding of the boundary of the range, he has also simulated a simpler situation. The idea is to consider just two 'competing' types on a two-dimensional torus which we can identify with $(-L, L]^2 \subseteq \mathbb{R}^2$. At time zero, all points of the torus with a non-positive first coordinate are of one type and all with a strictly positive first coordinate are of the other type. The region in which both types coexist develops in a rather complicated way, but it is natural to ask whether if one 'stands back' and views the process over large spatial scales (at sufficiently large times) a simpler pattern emerges. Saadi's simulations were the starting point for our work here.

We shall concentrate our attention on two special cases of the SLFV model, in both of which individuals can be one of only two genetic types, labelled 0 and 1. Evidently it is then enough to consider the proportion of type-1 individuals at each site and so we define, for every $x \in \mathbb{R}^d$ and $t \geq 0$,

$$w(t,x) := \rho_t(x)(\{1\}). \tag{2}$$

For the sake of clarity, we shall also take the fraction $u \in (0,1]$ to be the same for all events. In our previous notation, this corresponds to taking $\zeta(dr, dv) = \mu(dr)\delta_u(dv)$, for a measure μ on $(0,\infty)$. We shall allow the measure μ to take two forms:

Case A (fixed radius): We fix $r \in (0,\infty)$, and choose μ to be the Dirac mass at r.

Case B (heavy-tailed distribution): We fix $\alpha \in (1,2)$ and define the measure μ by

$$\mu(dr) = r^{-\alpha-d-1}\mathbf{1}_{\{r\geq 1\}}\, dr, \tag{3}$$

where we recall that d is the dimension of the geographical space.

It is easy to check that the condition (1) which guarantees existence of the SLFV process is satisfied in both cases.

Case A bears some similarity to the nearest-neighbour voter model, in that an individual spreads its type (/opinion) in a 'close' neighbourhood. Case B incorporates some large-scale events and consequently, as we shall see, behaves very differently. The particular form of μ is motivated by the fact that with this choice, under a suitable rescaling of space and time, the motion of an ancestral lineage will converge to a symmetric α-stable Lévy process (and, more generally, the ancestry of finitely many individuals converges to a system of coalescing dependent α-stable processes, see §5). Combined with duality, this will imply that with the same space-time rescaling, the forwards in time process of allele frequencies will also converge to a non-trivial limit.

Suppose that the initial condition of the process is

$$w(0,x) = \mathbf{1}_{\{x_{(1)}\leq 0\}},$$

where here again $x_{(1)}$ denotes the first coordinate of x. In words, we start from a half-space H of 1's. Let us set $\alpha = 2$ in Case A, and, for a given $\alpha \in (1,2]$ and any $n \in \mathbb{N}$, define the rescaled density w^n by

$$w^n(t,x) := w(nt, n^{1/\alpha}x), \qquad t \geq 0,\ x \in \mathbb{R}^d.$$

We denote by ρ^n the Ξ-valued process whose local density of 1's at time t is $w^n(t,\cdot)$. Our main results are the following two theorems, which describe the asymptotic behaviour of ρ^n as n tends to infinity. In Case A, σ^2 is the variance of the displacement, after one unit of time, of a single ancestral lineage from its starting point (see (7)).

Theorem 1.1. (Case A) *There exists a Ξ-valued process $\{\rho_t^{(2)},\, t \geq 0\}$ such that*

$$\rho^n \longrightarrow \rho^{(2)} \qquad \text{as } n \to \infty,$$

in the sense of weak convergence of the (temporal) finite-dimensional distributions.

Furthermore, at every time $t \geq 0$, the local density $w^{(2)}(t,\cdot) := \rho_t^{(2)}(\{1\})$ of type-1 individuals can be described as follows. If X denotes standard d-dimensional Brownian motion and

$$p^2(t,x) := \mathbb{P}_x[X_{\sigma^2 t} \in H], \qquad t \geq 0,\ x \in \mathbb{R}^d,$$

then:

1. *If $d = 1$, for every $t \geq 0$ and a.e. $x \in \mathbb{R}$, $w^{(2)}(t,x)$ is a Bernoulli random variable with parameter $p^2(t,x)$. The correlations between their values at distinct sites of $\mathbb{R}$ are non-trivial and are described in (15).*

2. If $d \geq 2$, for every $t \geq 0$ and a.e. $x \in \mathbb{R}^d$, $w^{(2)}(t,x)$ is deterministic and equal to $p^2(t,x)$.

Remark 1.2. Note that, in one dimension, the two types almost surely do not coexist at any given point, since $w^{(2)}(t,x)$ is a Bernoulli random variable. However, in higher dimensions, the two types 0 and 1 do coexist at every site instantaneously.

Remark 1.3. Although we have expressed everything in terms of densities, the convergence in Theorem 1.1, which we define explicitly in §3, is equivalent the convergence for all $j \geq 1$ of the measures

$$\rho_t^n(x_1, d\kappa_1) \dots \rho_t^n(x_j, d\kappa_j) dx_1 \dots dx_j.$$

(This is the vague convergence of nonnegative Radon measures on $(\mathbb{R}^d \times \{0,1\})^j$.) See [VW11] for a measure-valued formulation of the SLFV and for a proof of this equivalence.

Remark 1.4. The quantity $p^2(t,x)$ implicitly depends on the dimension. Also, since u and r are fixed, substituting in (7),

$$\sigma^2 = \frac{u}{dV_r} \int_{\mathbb{R}^d} dz \; |z|^2 L_r(z) \qquad \left(= \frac{4ur^3}{3} \text{ when } d = 1 \right)$$

is finite and proportional to u. Indeed, $L_r(z) := \mathrm{Vol}(B(0,r) \cap B(0,z)) = (2r - |z|)_+$ in dimension 1 and, more generally, $L_r(z) \leq \mathbf{1}_{\{|z| \leq 2r\}} V_r$ for any $d \geq 1$.

In contrast to the case of fixed radii, in Case B, in the limit as $n \to \infty$ types are always segregated, irrespective of dimension.

Theorem 1.5. (Case B) *There exists a Ξ-valued process $\{\rho_t^{(\alpha)}, t \geq 0\}$ such that*

$$\rho^n \longrightarrow \rho^{(\alpha)} \qquad \text{as } n \to \infty,$$

in the sense of weak convergence of the (temporal) finite-dimensional distributions.

Furthermore, there exists a symmetric α-stable process X^α such that if

$$p^\alpha(x,t) := \mathbb{P}_x\big[X^\alpha_{ut} \in H\big], \qquad t \geq 0, \; x \in \mathbb{R}^d,$$

then for every $t > 0$ and a.e. $x \in \mathbb{R}^d$, $w^{(\alpha)}(t,x)$ is a Bernoulli random variable with parameter $p^\alpha(t,x)$. The correlations between the values of the densities at different sites (and at the same time t) are again given by (15) (or (9)), where the process ξ^∞ is now the system of coalescing α-stable processes obtained in Proposition 5.1.

Here again, one should notice that the speed of evolution of the limiting process is proportional to the parameter u.

Comparing the results of Theorem 1.1 and Theorem 1.5, one can see that very large extinction-recolonisation events create correlations between local genetic diversities over a much larger spatial scale ($n^{1/\alpha} \gg \sqrt{n}$) than purely local reproduction events. This is because an ancestral lineage can move a distance $\mathcal{O}(n^{1/\alpha})$ over the course of n generations. One might initially guess that, since the motion of a single ancestral lineage under our rescaling converges to a symmetric stable process, two distinct ancestral lineages would (asymptotically) only meet (and thus have a chance to coalesce) in dimensions where the stable process hits points. This is precisely what we see in Case A and, in that case, lies behind the deterministic limit in $d \geq 2$. However, this is where the dependence between ancestral lineages in the SLFV process (see §3.2) comes into play. The detailed analysis of the ancestral process for Case B (which we present in §5) reveals that

'very large' events are frequent enough to capture lineages that have moved to arbitrarily large separations. In particular, Lemma 5.3 shows that, in Case B, any finite sample of individuals will find its most recent common ancestor in finite time a.s. (see also Remark 5.4). The large events will, momentarily, create extensive areas in which the two genetic types coexist. Our analysis will also show that, under our rescaling, 'small' events then occur sufficiently quickly to instantaneously restore the allele frequencies in each infinitesimal region to 0 or 1 (see also the simulations presented in §2).

The rest of the paper is laid out as follows. In §2, we present some simulations that illustrate the results and the mechanisms underlying them. In §3, we are explicit about the meaning of 'weak convergence of the (temporal) finite-dimensional distributions' and we describe the duality between allele frequencies and ancestral processes that provides the main tool in our proofs. It is then used to find conditions, expressed in terms of the genealogical trees relating individuals in a sample from the population, under which $w^{(\alpha)}(t,x)$ (at each time $t > 0$ and a.e. point $x \in \mathbb{R}^d$) takes the particular forms seen in our main theorems (see Lemma 3.2). Theorems 1.1 and 1.5 are then proved in §4 and §5 respectively. This last section also contains some results (Lemma 5.3 and the accompanying remark), of independent interest, on the system of coalescing (dependent) Lévy processes that generates the genealogical trees relating a sample of individuals from the limiting population.

2 Simulations

Our results show that in the cases where the rescaled density of type 1 individuals converges to a *random* limit, at any fixed time that limit takes the form of the indicator function of a random set. In one dimension, provided that either $u = 1$ or $\alpha = 2$ (the radius of events is fixed), the set takes a simple form, but for $\alpha \in (1,2)$ this is no longer the case. In this section we present some simulations that illustrate the complex geometry of the limiting random sets and the mechanism that leads to their creation. We are extremely grateful to Jerome Kelleher from the University of Edinburgh for performing these simulations and producing the figures.

First suppose that we are in one spatial dimension. If $u = 1$, then at every stage of the rescaling we will have $w^n(t,x) = \mathbf{1}_{I_t^n}(x)$ where I_t^n is a half-line with right endpoint R_t following a random walk on $\mathbb{R}$. Under our rescaling, as $n \to \infty$, the process R_t will converge to a Brownian motion if $\alpha = 2$ and to a symmetric stable process of index α for $\alpha \in (1,2)$. If $\alpha = 2$, and $d = 1$, then the same is true for $u < 1$. This can be understood via the dual process of ancestral lineages. As we shall see, this converges to a system of independent Brownian motions which coalesce instantaneously on meeting. The type of an individual sampled at x at time t is determined by the type at time t before the present of the corresponding ancestral lineage. Since the Brownian motions are continuous, and they coalesce as soon as they meet, it is impossible for two lineages to 'cross over'. Consequently, asymptotically, if a lineage started from x traces back to a point to the left of the origin at time t before the present, then so must all lineages started from points to the left of x. As a result, at time t, the density of type 1s will still be the indicator function of a half-line. The boundary, R_t, moves in the same way as a single ancestral lineage, that is as a Brownian motion with a clock that runs at a rate proportional to u. Figure 1 shows the results of a simulation of the process of allelic types in this case. In two dimensions, two Brownian motions won't meet and so for $\alpha = 2$, asymptotically, the ancestral lineages will just look like independent Brownian motions and forwards in time, asymptotically, allele frequencies are smeared out by the deterministic heat flow.

The case $\alpha \in (1,2)$ is much more interesting. Now, even in the limit, ancestral lineages evolve in a series of jumps and if $u < 1$ they *can* 'cross over'. Thus although our results show that the

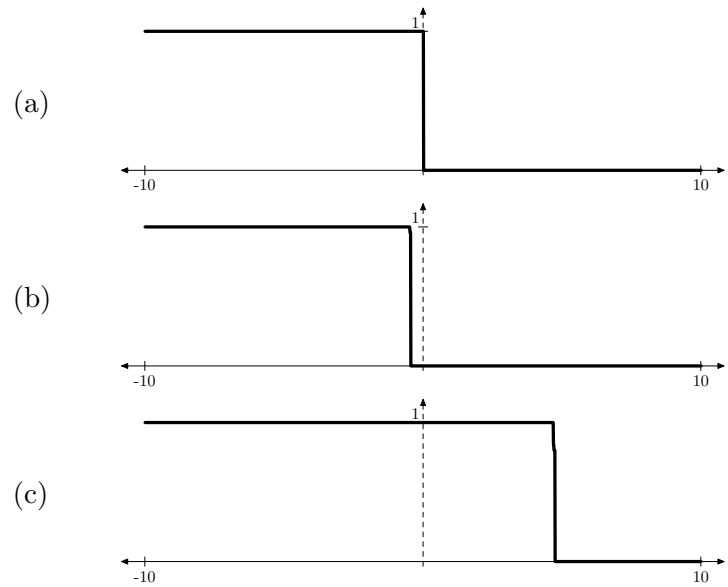

Figure 1: Fixed radius in $d = 1$ on a line of length 20. (a) initial conditions; (b) after 10^5 events; (c) after 10^7 events. The model parameters are $u = 0.8$, $r = 0.033$, $n = 10^3$.

limiting allele frequencies always look like the indicator function of a random set, even in $d = 1$ we can no longer expect that set to be connected. Forwards in time what our results suggest, and simulations confirm, is that a large event can create a region in which allele frequencies are strictly between zero and one, but these frequencies are rapidly (and asymptotically instantaneously) 'resolved' by 'small' events so that the state is restored to being the indicator function of a set. Figure 2 shows how on the line this mechanism leads to allele frequencies that look like a series of 'crenellations'. Even in one spatial dimension, our methods are not powerful enough to allow us to capture detailed information about the random sets observed in the limit.

Figure 3 illustrates the same mechanism in two spatial dimensions. To isolate the effect in which we are interested, we suppose that a large event covers a previously unblemished portion of the interface and observe the resolution of the resulting patch of coexistence.

3 Convergence and duality

3.1 State-space and form of convergence

In order to make the convergence in Theorems 1.1 and 1.5 explicit, let us recall some facts about the state space of the SLFV from [BEV10]. In §1.1, we described the process as taking its values in the set $\tilde{\Xi}$ of all measurable functions $\rho : \mathbb{R}^d \to \mathcal{M}_1(K)$ (where the compact type space K is now $\{0, 1\}$). In fact, we need to define an equivalence relation on this space by setting

$$\rho \sim \rho' \qquad \Leftrightarrow \qquad \mathrm{Vol}\big(\{x \in \mathbb{R}^d \,:\, \rho(x) \neq \rho'(x)\}\big) = 0.$$

The state-space Ξ of the SLFV is then defined as the quotient space $\tilde{\Xi}/\sim$ of equivalence classes of $\sim$.

Ξ is endowed with a natural topology described in §3 of [Eva97]. As mentioned earlier, it is possible to check that this topology is equivalent to the topology of vague convergence of the measures $\rho(x_1, d\kappa_1) \dots \rho(x_j, d\kappa_j) dx_1 \dots dx_j$ for all $j \geq 1$ as Radon measures on $(\mathbb{R}^d \times \{0, 1\})^j$.

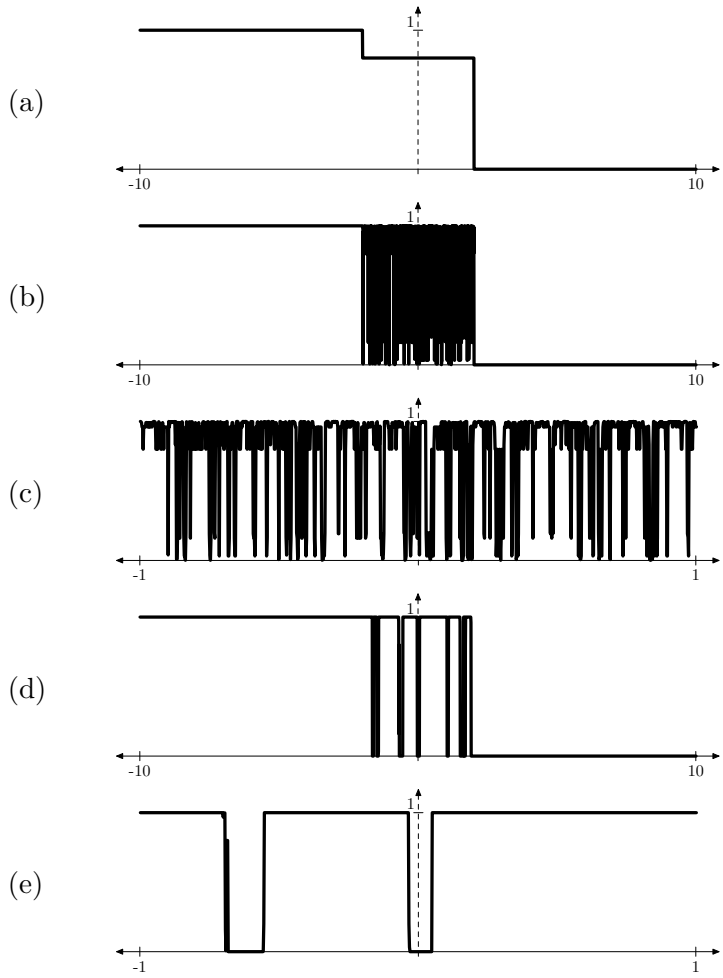

Figure 2: Variable radius in $d = 1$ on a line of length 20. (a) initial conditions; (b) after 100 events, full range; (c) after 100 events, zooming in; (d) after 10^6 events, full range; (e) after 10^6 events, zooming in. The model parameters are $u = 0.8$, $n = 10^4$ and $\alpha = 1.3$.

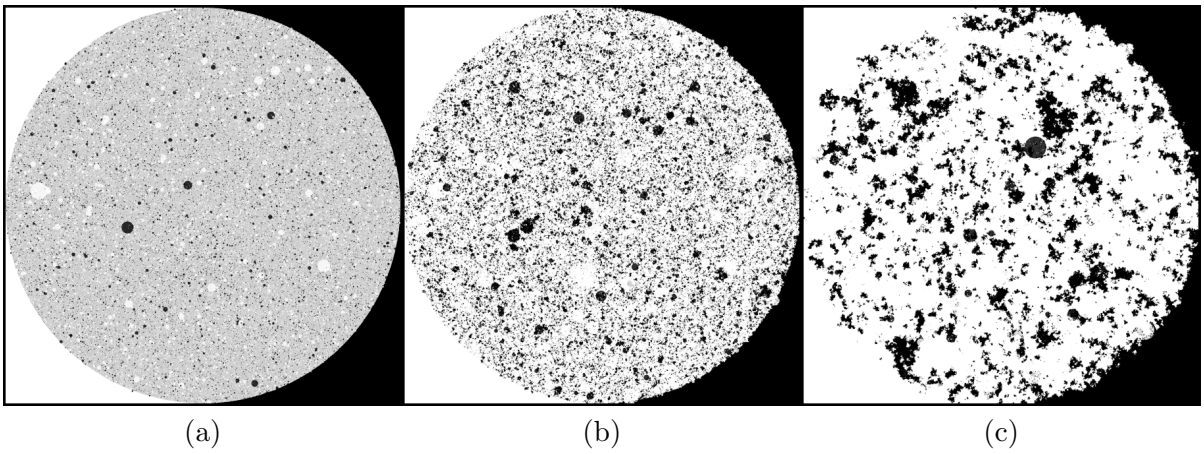

Figure 3: Model in $d = 2$ after (a) 10^5; (b) 10^6; and (c) 10^7 events. We have a square range of edge 8, and the initial patch is a circle of radius 4 with frequency 0.8 (white is frequency 1, black is 0). The model parameters are $u = 0.8$, $\alpha = 1.3$ and $n = 10^3$.

In practice, Lemma 4.1 in [BEV10] provides us with a family of functions which is dense in the set of continuous functions on Ξ. To introduce this class of functions, for a space E, let $C(E)$ denote the set of all continuous functions on E and, for a measure ν, let $L^1(\nu)$ be the set of all functions which are integrable with respect to ν. For every $j \in \mathbb{N}$, $\psi \in C((\mathbb{R}^d)^j) \cap L^1(dx^{\otimes j})$ and $\chi_1, \dots, \chi_j \in C(K)$, we define the function $I_j(\cdot\,;\, \psi, (\chi_i)_{1\le i\le j})$ as follows. For every $\rho \in \Xi$,

$$I_j(\rho\,;\, \psi, (\chi_i)_{1\le i\le j}) := \int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j \;\psi(x_1, \dots, x_j) \left(\prod_{i=1}^{j} \langle \chi_i, \rho(x_i) \rangle \right),$$

where $\langle f, \nu \rangle$ is the integral of the function f with respect to the measure ν.

Since in our setting $K = \{0, 1\}$, we have, for every χ,

$$\begin{aligned} \langle \chi, \rho(x) \rangle &= \chi(1) w(x) + \chi(0)(1 - w(x)) \\ &= \big(\chi(1) - \chi(0)\big) w(x) + \chi(0), \end{aligned}$$

where, as before, $w(x) := \rho(x)(\{1\})$ denotes the mass of 1's at site x. We can therefore restrict our attention to the set of functions I_j such that $\chi_i = \mathrm{Id}$ for every $i \in \{1, \dots, j\}$, that is

$$I_j(\rho\,;\, \psi) = \int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j \;\psi(x_1, \dots, x_j) \left(\prod_{i=1}^{j} w(x_i) \right). \tag{4}$$

Indeed, any $I_j(\cdot\,;\, \psi, (\chi_i)_{1\le i\le j})$ can be written as a finite linear combination of functions of the form (4). The convergence stated in Theorems 1.1 and 1.5 can now be expressed for a single time $t \ge 0$ as: for every $j \in \mathbb{N}$ and $\psi \in C((\mathbb{R}^d)^j) \cap L^1(dx^{\otimes j})$,

$$\lim_{n\to\infty} \mathbb{E}\big[I_j\big(\rho_t^n\,;\, \psi\big)\big] = \mathbb{E}\big[I_j\big(\rho_t^{(\alpha)}\,;\, \psi\big)\big]. \tag{5}$$

The extension of this definition of convergence to joint convergence at several times $t_1, \dots, t_k$ is straightforward.

3.2 Duality between the SLFV and its genealogies

The proofs of Theorems 1.1 and 1.5 rely on a duality relation between the SLFV process, and the system of coalescing jump processes that we call the *genealogical process of a sample of individuals* from the population. We recall this relation in the particular form in which we shall need it. In particular we restrict our attention to $K = \{0, 1\}$. A more general statement (and proofs) can be found in §4 of [BEV10].

First suppose that we wish to trace the ancestry of a single individual alive in the current population. Let us, for now, work in a general setting as it will allow us to understand condition (1) a little better. Since the model is translation invariant, without loss of generality we may suppose that the individual is currently at the origin in $\mathbb{R}^d$. Tracing backwards in time, at the first time in the past when 0 is in the area $B(x, r)$ affected by a reproduction event, our individual has probability u of being an offspring of that event, in which case the ancestral lineage jumps to the position of the parent (which is uniformly distributed on $B(x, r)$). Since the Poisson process driving events is reversible, we see that the rate at which our ancestral lineage experiences a jump is

$$\int_{\mathbb{R}^d} \int_{(0,\infty)\times[0,1]} \zeta(dr, du) dx \; \mathbf{1}_{\{0 \in B(x,r)\}} u = \int_{(0,\infty)\times[0,1]} \zeta(dr, du) \; u V_r.$$

By translation invariance in time and space of the law of Π, this tells us that the quantity in (1) is just the instantaneous jump rate of an ancestral lineage (at any time and any location), and we are requiring it to be finite. We refer to §4 in [BEV10] for an explanation of why this guarantees existence and uniqueness of the process $(\rho_t)_{t\geq 0}$.

We will need to be more precise about the law of the compound Poisson process followed by an ancestral lineage and so we now establish the rate at which it jumps from 0 to z (or, by translation invariance, from y to $y+z$). In order for such a jump to occur, first 0 and z must both belong to the area hit by the event; second our lineage at 0 must belong to the fraction u of individuals replaced; and third the parent must be chosen from site z. The intensity measure of the jump process is therefore equal to

$$\begin{aligned} m(dz) &:= \int_{\mathbb{R}^d}\int_{(0,\infty)\times[0,1]} \zeta(dr,du)dx\, \mathbf{1}_{\{x\in B(0,r)\cap B(z,r)\}} u\,\frac{dz}{V_r} \\ &= \left(\int_{(0,\infty)\times[0,1]} \zeta(dr,du)\frac{uL_r(z)}{V_r}\right) dz, \end{aligned} \tag{6}$$

where $L_r(z)$ denotes the volume of the intersection $B(0,r)\cap B(z,r)$. (To see this, note that for an event of radius r to affect both 0 and z, its centre, x, must lie in the region $B(0,r)\cap B(z,r)$ and that since the parent is chosen uniformly from the region, the factor $1/V_r$ arises as the density of the uniform distribution on $B(x,r)$.) In particular, by rotational symmetry, in the special case where the variance of the displacement of a lineage over one unit of time is finite, its covariance matrix is of the form $\sigma^2\mathrm{Id}$, with

$$\sigma^2 := \int_{\mathbb{R}^d} m(dz)\,\left(z_{(1)}\right)^2 = \frac{1}{d}\int_{\mathbb{R}^d} m(dz)\,|z|^2 \tag{7}$$

(here, $z_{(1)}$ denotes the first coordinate of z, and $|z|$ its L^2-norm).

Much of our analysis will rest upon understanding the ancestry of (larger) samples from the population, and these can be established in much the same way as the motion of a single ancestral lineage. If we sample k individuals (possibly from the same location), the ancestry is given by a system of (finite-rate) jump processes, which are *a priori* correlated, since their jumps are generated by the same Poisson point process of events. Furthermore, if at least two of them are encompassed by the same event and lie within the fraction of the local population replaced, then these lineages trace back to the same parent and thus merge into a single lineage during the event. Tracing further back in time, that single lineage and all other remaining lineages continue to evolve in the same manner. Note that if $u<1$, there may be other lineages in the ball where the event takes place, but not in the sub-population replaced. Such lineages neither jump nor coalesce during the event.

Let $(\mathcal{A}_t)_{t\geq 0}$ be a system of finitely many (the initial number will always be specified explicitly) ancestral lineages as described above. That is, each lineage follows a finite-rate jump process with jump intensity (6), and two or more lineages coalesce whenever they are affected by the same event. See Equation (24) in §5 for an expression for the generator of this process, in the particular case where u is fixed. For every $t\geq 0$, let us write N_t for the number of distinct lineages at time t, and $\xi_t^1,\ldots,\xi_t^{N_t}\in\mathbb{R}^d$ for their spatial locations at that time. The weak duality relation we shall use in the sequel is also based on the family of functions in (4), and states that

for every $j \in \mathbb{N}$ and $\psi \in C((\mathbb{R}^d)^j) \cap L^1(dx^{\otimes j})$, we have, for every $t \geq 0$,

$$\int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j \ \psi(x_1, \dots, x_j) \ \mathbb{E}\big[w(t, x_1) \dots w(t, x_j) \,\big|\, w(0, \cdot) = w_0\big] \tag{8}$$
$$= \int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j \ \psi(x_1, \dots, x_j) \ \mathbb{E}\big[w_0(\xi_t^1) \cdots w_0\big(\xi_t^{N_t}\big) \,\big|\, N_0 = j, \ \xi_0^1 = x_1, \dots, \xi_0^j = x_j\big].$$

Since (8) is valid for all functions ψ as above, we also have for Lebesgue-a.e. $(x_1, \dots, x_j)$,

$$\mathbb{E}\big[w(t, x_1) \dots w(t, x_j) \,\big|\, w(0, \cdot) = w_0\big] = \mathbb{E}\big[w_0(\xi_t^1) \cdots w_0\big(\xi_t^{N_t}\big) \,\big|\, N_0 = j, \ \xi_0^1 = x_1, \dots, \xi_0^j = x_j\big]. \tag{9}$$

Remark 3.1. The weak duality in (9) is very similar to the duality between the Kimura stepping stone model and a system of coalescing random walks (see e.g. Chap.6 of [Eth11]). Here, however, in contrast to the discrete space setting, we cannot deduce an expression for the second or higher order moments of the $w(t, x)$'s since (9) only holds for Lebesgue-a.e j-tuple $(x_1, \dots, x_j)$ (and the x_i's are pairwise distinct for Lebesgue-a.e. vector). The problem stems from the fact the actual object with which we are dealing is the random measure $w(t, x)dx$ and not the collection $\{w(t, x)\}_{x \in \mathbb{R}^d}$. The topology on Ξ is too weak to consider the evolution of the density of 1's at every single point, and we are obliged to characterize this density through a local averaging procedure, see (12).

Thanks to (8), proving the convergence of $\rho_t^n \equiv \{w(nt, n^{1/\alpha}x)\}_{x \in \mathbb{R}^d}$ boils down to showing that the genealogical process relating a finite sample of individuals converges, and to transferring the result to the forwards-in-time process. In addition, these duality relations enable us to obtain an explicit description of the local densities $w^{(\alpha)}(t, x)$. Indeed, (5) and (8) lead us to an implicit characterisation of the limiting random field $\rho^{(\alpha)}$ through the values of

$$\mathbb{E}\big[I_j(\rho_t^{(\alpha)}\,; \psi)\big] = \mathbb{E}\bigg[\int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j \ \psi(x_1, \dots, x_j) \bigg(\prod_{i=1}^{j} w^{(\alpha)}(t, x_j)\bigg)\bigg].$$

However, the following result gives us more information on the form of the $w^{(\alpha)}(t, x)$'s.

Lemma 3.2. *Suppose that $(\rho_t)_{t \geq 0}$ is a Ξ-valued process dual to an exchangeable and consistent system of coalescing Markov processes $(\mathcal{A}_t)_{t \geq 0}$ through the relations (8). Let $(\xi_t)_{t \geq 0}$ denote the Markov process followed by a single lineage, and suppose that the initial condition of ρ is such that for every $t > 0$, the map $z \mapsto \mathbb{E}_z[w(0, \xi_t)]$ is continuous on $\mathbb{R}^d$ (where as usual $\mathbb{E}_z$ denotes expectation under $\mathbb{P}_z$).*

(i) If for every $\varepsilon > 0$ we have

$$\lim_{|y-x| \to 0} \mathbb{P}\big[\text{lineages 1 and 2 have not coalesced by time } \varepsilon \,\big|\, \xi_0^1 = x, \xi_0^2 = y\big] = 0, \tag{10}$$

where the convergence is uniform with respect to $x \in \mathbb{R}^d$, then for every $t > 0$ and a.e. $x \in \mathbb{R}^d$, $w(t, x)$ is a Bernoulli random variable with parameter $\mathbb{E}_x[w(0, \xi_t)]$.

(ii) If $(\mathcal{A}_t)_{t \geq 0}$ is a system of independent Markov processes which never coalesce whenever they start from distinct locations, then for every $t > 0$ and a.e. $x \in \mathbb{R}^d$, $w(t, x)$ is deterministic and equal to $\mathbb{E}_x[w(0, \xi_t)]$.

Here, by 'exchangeable' we mean that the law of $(\mathcal{A}_t)_{t\geq 0}$ is invariant under relabelling of the initial lineages; 'consistent' means that for every $j \in \mathbb{N}$, if $\mathcal{A}$ starts with $j+1$ lineages but we only follow the evolution of the first j of them, we obtain a system of coalescing Markov processes that has the same law as $\mathcal{A}$ started with only j lineages. In other words, the evolution of the $(j+1)$-st lineage does not influence that of the other j. It is not difficult to see that the system $(\mathcal{A}_t)_{t\geq 0}$ introduced at the beginning of this section is indeed exchangeable and consistent (since each lineage present in the area hit by an event is affected with probability u independently of all others). The limiting genealogies we shall obtain will inherit these properties.

Proof of Lemma 3.2. Let us fix $t \geq 0$, and consider the random measure $\ell(dx)$ on $\mathbb{R}^d$ defined by: for every nonnegative measurable function ψ,

$$\int_{\mathbb{R}^d} \ell(dx)\ \psi(x) := \int_{\mathbb{R}^d} dx\ \psi(x) w(t,x). \tag{11}$$

Notice that, according to the description of Ξ given in §3.1, $w(t,\cdot) := \rho_t(\cdot)(\{1\})$ is in fact an equivalence class of functions of the form $\tilde{w} : \mathbb{R}^d \to [0,1]$. Two representatives of $w(t,\cdot)$ differ only on a Lebesgue negligible subset of $\mathbb{R}^d$. For the rest of this proof we assume that for every ω in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which ρ_t is defined, we have fixed a representative $\tilde{w}(\omega) : \mathbb{R}^d \to [0,1]$ of $w(\omega,t,\cdot)$ and define $\ell(\omega, dx)$ as in (11), with $w(\omega,t,\cdot)$ replaced by $\tilde{w}(\omega,\cdot)$.

Let $(\varphi_m)_{m\in\mathbb{N}}$ be a sequence of continuous functions on $\mathbb{R}^d$ such that for every m, $0 \leq \varphi_m \leq 1$, $\varphi_m \equiv 1$ on $B(0,1/m)$ and $\varphi_m \equiv 0$ outside $B(0,2/m)$. Let us write $\varphi_m(\mathbb{R}^d)$ for the integral $\int_{\mathbb{R}^d} dz\ \varphi_m(z)$. Since $\tilde{w}$ is locally integrable (it has values in $[0,1]$), the Lebesgue Differentiation Theorem guarantees that for every $\omega \in \Omega$, there exists a Lebesgue null set $\mathcal{N}(\omega)$ such that for every $x \notin \mathcal{N}(\omega)$,

$$\lim_{m\to\infty} \frac{1}{\varphi_m(\mathbb{R}^d)} \int_{\mathbb{R}^d} \ell(\omega, dz)\ \varphi_m(x+z) = \tilde{w}(\omega, x). \tag{12}$$

Consequently, by Fubini's theorem there exists a Lebesgue null set $\mathcal{O}$ such that for every $x \notin \mathcal{O}$, the convergence in (12) occurs with $\mathbb{P}(d\omega)$-probability one. Evidently, if we can show that the random variable $\tilde{w}(x)$ is as in the statement of Lemma 3.2 for every $x \notin \mathcal{O}$, we shall obtain the desired result for $w(t,\cdot)$.

Now fix $x \in \mathbb{R}^d \setminus \mathcal{O}$, so that (12) holds $\mathbb{P}$-a.s. We show that $\tilde{w}(x)$ is a Bernoulli random variable under the condition stated in (i), and a deterministic constant under the condition given in (ii). Let $j \in \mathbb{N}$. On the one hand, the Dominated Convergence Theorem yields that

$$\lim_{m\to\infty} \mathbb{E}\left[\left(\varphi_m(\mathbb{R}^d)^{-1} \int_{\mathbb{R}^d} \ell(dz)\ \varphi_m(x+z)\right)^j\right] = \mathbb{E}\big[\tilde{w}(x)^j\big]. \tag{13}$$

On the other hand, by Fubini's theorem and (8), we have that for every $m \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{E}\left[\left(\varphi_m(\mathbb{R}^d)^{-1} \int_{\mathbb{R}^d} \ell(dz)\ \varphi_m(x+z)\right)^j\right] \\
&\quad = \varphi_m(\mathbb{R}^d)^{-j} \int_{(\mathbb{R}^d)^j} dz_1 \dots dz_j\, \varphi_m(x+z_1)\cdots\varphi_m(x+z_j)\ \mathbb{E}\big[w(t,z_1)\dots w(t,z_j)\big] \\
&\quad = \varphi_m(\mathbb{R}^d)^{-j} \int_{(\mathbb{R}^d)^j} dz_1 \dots dz_j\, \mathbb{E}\big[w(0,\xi_t^1)\cdots w\big(0,\xi_t^{N_t}\big) \mid N_0 = j,\ \xi_0^1 = z_1, \dots, \xi_0^j = z_j\big] \times \\
&\qquad\qquad \varphi_m(x+z_1)\cdots\varphi_m(x+z_j).
\end{aligned} \tag{14}$$

Since the Lebesgue measure of the set of j-tuples with at least two identical coordinates is 0, under the condition of (ii) the quantity in the right-hand side of (14) is equal to

$$\int_{(\mathbb{R}^d)^j} dz_1 \dots dz_j \prod_{i=1}^{j} \left\{ \frac{\varphi_m(x+z_i)}{\varphi_m(\mathbb{R}^d)} \mathbb{E}_{z_i}[w(0,\xi_t)] \right\} = \left(\int_{\mathbb{R}^d} dz \, \frac{\varphi_m(x+z)}{\varphi_m(\mathbb{R}^d)} \mathbb{E}_z[w(0,\xi_t)] \right)^j.$$

By our continuity assumption, this quantity tends to $\mathbb{E}_x[w(0,\xi_t)]^j$ as $m \to \infty$. Combined with (13), this gives us that $\tilde{w}(x)$ is a.s. equal to the constant $\mathbb{E}_x[w(0,\xi_t)]$ under the condition stated in (ii).

To see (i), consider the case $j = 2$ (i.e., $\mathcal{A}$ consists of two ancestral lineages) and let us write τ for the time at which they coalesce, with the convention that $\tau = \infty$ if $\mathcal{A}$ always contains two lineages. Since $\varphi_m(x+\cdot)$ is concentrated on $B(x, 2/m)$, using (10) we obtain that for every $\varepsilon > 0$,

$$\lim_{m\to\infty} \frac{1}{\varphi_m(\mathbb{R}^d)^2} \int_{(\mathbb{R}^d)^2} dz_1 dz_2 \, \varphi_m(x+z_1)\varphi_m(x+z_2) \mathbb{P}\left[\tau > \varepsilon \,|\, N_0 = 2,\ \xi_0^1 = z_1, \xi_0^2 = z_2\right] = 0.$$

Hence, for $j = 2$ and $\varepsilon < t$, the quantity on the right-hand side of (14) can be written

$$\int_{(\mathbb{R}^d)^2} dz_1 dz_2 \, \mathbb{E}\big[w(0,\xi_t^1)\mathbf{1}_{\{\tau\le\varepsilon\}} \,\big|\, N_0 = 2,\ \xi_0^1 = z_1, \xi_0^2 = z_2\big] \frac{\varphi_m(x+z_1)\varphi_m(x+z_2)}{\varphi_m(\mathbb{R}^d)^2} + \delta(\varepsilon, m),$$

where $\delta(\varepsilon, m) \to 0$ as $m \to \infty$ for every fixed ε. By the same argument, we have

$$\begin{aligned}
&\int_{(\mathbb{R}^d)^2} dz_1 dz_2 \, \mathbb{E}\big[w(0,\xi_t^1)\mathbf{1}_{\{\tau\le\varepsilon\}} \,\big|\, N_0 = 2,\ \xi_0^1 = z_1, \xi_0^2 = z_2\big] \frac{\varphi_m(x+z_1)\varphi_m(x+z_2)}{\varphi_m(\mathbb{R}^d)^2} \\
&= \int_{(\mathbb{R}^d)^2} dz_1 dz_2 \, \mathbb{E}\big[w(0,\xi_t^1) \,\big|\, N_0 = 2,\ \xi_0^1 = z_1, \xi_0^2 = z_2\big] \frac{\varphi_m(x+z_1)\varphi_m(x+z_2)}{\varphi_m(\mathbb{R}^d)^2} + \delta'(\varepsilon, m) \\
&= \int_{(\mathbb{R}^d)^2} dz_1 dz_2 \, \mathbb{E}_{z_1}[w(0,\xi_t)] \frac{\varphi_m(x+z_1)\varphi_m(x+z_2)}{\varphi_m(\mathbb{R}^d)^2} + \delta'(\varepsilon, m) \\
&= \int_{\mathbb{R}^d} dz_1 \, \mathbb{E}_{z_1}[w(0,\xi_t)] \frac{\varphi_m(x+z_1)}{\varphi_m(\mathbb{R}^d)} + \delta'(\varepsilon, m),
\end{aligned}$$

where $\delta'(\varepsilon, m)$ also tends to 0 as $m \to \infty$ for every $\varepsilon > 0$, and the third line is justified by the consistency of $(\mathcal{A}_t)_{t\ge0}$. Using again our continuity assumption on $z \mapsto \mathbb{E}_z[w(0,\xi_t)]$, we obtain that under the condition stated in (i), the quantity on the right-hand side of (14) converges to $\mathbb{E}_x[w(0,\xi_t)]$ as $m \to \infty$. Hence, coming back to (13), we arrive at

$$\mathbb{E}\big[\tilde{w}(x)^2\big] = \mathbb{E}_x[w(0,\xi_t)] = \mathbb{E}\big[\tilde{w}(x)\big].$$

Since $\tilde{w}(x) \in [0,1]$ almost surely, we deduce that $\tilde{w}(x) \in \{0,1\}$ almost surely, whence $\tilde{w}(x)$ is a Bernoulli random variable. This completes the proof of Lemma 3.2 (i). □

Note that (ii) corroborates a remark at the beginning of §5 in [Eva97]. In Evans' construction, all the genealogical processes used as duals are made up of independent Hunt processes that coalesce instantaneously upon meeting. Evans points out that, in this case, if ξ and ξ' are two independent processes having the same law as the motion of a single lineage, then the corresponding Ξ-valued process evolves deterministically iff

$$\mathrm{Vol}\Big(\Big\{(z_1,z_2) \in (\mathbb{R}^d)^2 \,:\, \mathbb{P}_{z_1,z_2}\big[\exists\, t \ge 0 \,:\, \xi_t = \xi'_t\big] > 0\Big\}\Big) = 0.$$

That is, if the set of pairs of starting points (z_1, z_2) such that ξ and ξ' have a positive chance to meet in finite time is negligible with respect to Lebesgue measure, then for every $t > 0$, ρ_t is a deterministic function of its initial value (and so is $w(t, \cdot)$). Our proof of Lemma 3.2 gives an alternative proof of Evans's remark when the type-space K is $\{0, 1\}$.

4 Proof of Theorem 1.1

Let us start by proving the convergence stated in Theorem 1.1 for a single time $t \geq 0$. Since we start from $w(0, \cdot) = \mathbf{1}_H(\cdot)$ (where $H \subset \mathbb{R}^d$ is the half-space of all points whose first coordinate is non-positive), for every $n \in \mathbb{N}$ we have $w(0, \cdot\sqrt{n}) = \mathbf{1}_H(\cdot)$. Hence, we need only prove the result for $t > 0$.

From our definition of convergence (see (5)), our aim is to show that for every $j \in \mathbb{N}$ and $\psi \in C((\mathbb{R}^d)^j) \cap L^1(dx^{\otimes j})$,

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j\, \psi(x_1, \dots, x_j) w(tn, x_1\sqrt{n}) \cdots w(tn, x_j\sqrt{n})\Big]$$
$$= \mathbb{E}\Big[\int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j\, \psi(x_1, \dots, x_j) w^{(2)}(t, x_1) \cdots w^{(2)}(t, x_j)\Big].$$

As we explained in §3.2, this question boils down to establishing the asymptotic behaviour of

$$\int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j\, \psi(x_1, \dots, x_j) \mathbb{E}\big[w(0, \xi^1_{tn}) \cdots w(0, \xi^{N_{tn}}_{tn}) \,\big|\, N_0 = j, \xi^1_0 = x_1\sqrt{n}, \dots, \xi^j_0 = x_j\sqrt{n}\big].$$

This will be achieved in Lemmas 4.1 and 4.2 below, but first we need some notation. Recall that ξ represents the motion of a single ancestral lineage, that is ξ is a compound Poisson process in which jumps from x to $x + z$ have intensity

$$m(dz) = \frac{uL_r(z)}{V_r}\, dz.$$

Observe in passing that this intensity is 0 whenever $|z| \geq 2r$ (since the start and end points of a jump must belong to the same ball of radius r and so the size of this jump is bounded by $2r$). For every $n \in \mathbb{N}$, let ξ^n be the process on $\mathbb{R}^d$ defined by

$$\xi^n_t := \frac{1}{\sqrt{n}}\, \xi_{tn}, \qquad t \geq 0,$$

and let $\mathcal{A}^n$ be the corresponding rescaling of $\mathcal{A}$ in which time is multiplied by n and spatial locations are scaled down by $\sqrt{n}$. More formally, here we view $\mathcal{A}^n$ and $\mathcal{A}$ as having values in the state-space $\bigcup_{m\geq 0}\{m\} \times (\mathbb{R}^d)^m$; thus the first coordinate of $\mathcal{A}$ or $\mathcal{A}_n$ indicates the number of distinct lineages, while the remaining coordinates give their respective positions. (In Section 5 we will find it convenient to enrich the state space to also record the whole genealogical information).

Lemma 4.1. *If $d = 1$, for every $j \in \mathbb{N}$ and $x_1, \dots, x_j \in \mathbb{R}^d$, the process $\mathcal{A}^n$ starting from j lineages at locations $x_1, \dots, x_j$ converges, in the sense of finite-dimensional distributions, to a system $\mathcal{A}^\infty$ of independent Brownian motions with clock speed σ^2 that coalesce instantaneously upon meeting.*

More generally, let $k \in \mathbb{N}$ and $0 < t_1 < \ldots < t_k$. Suppose that we start $\mathcal{A}^n$ with j_0 lineages at distinct locations $x_{0,1}, \ldots, x_{0,j_0}$, let the process evolve until time t_1, add to the surviving lineages j_1 lineages at distinct locations $x_{1,1}, \ldots, x_{1,j_1}$, let all resulting lineages evolve until time t_2 when we add j_2 further lineages, and so on. Call the corresponding process $\hat{\mathcal{A}}^n$. Define $\hat{\mathcal{A}}^\infty$ analogously. Then for any $t \geq 0$, the law of $\hat{\mathcal{A}}^n_t$ converges to that of $\hat{\mathcal{A}}^\infty_t$ as n tends to infinity.

Lemma 4.2. *If $d \geq 2$, for every $j \in \mathbb{N}$ and distinct $x_1, \ldots, x_j \in \mathbb{R}^d$, the process $\mathcal{A}^n$ starting from j lineages at locations $x_1, \ldots, x_j$ converges to a system of independent Brownian motions with speed σ^2. In particular, the limiting lineages never coalesce.*

More generally, define $\hat{\mathcal{A}}^n$ and $\hat{\mathcal{A}}^\infty$ as in Lemma 4.1. Then for any $t \geq 0$, the law of $\hat{\mathcal{A}}^n_t$ converges to that of $\hat{\mathcal{A}}^\infty_t$ as n tends to infinity.

We postpone the proofs of Lemmas 4.1 and 4.2 until the end of this section.

Since the boundary of H has zero Lebesgue measure, Portmanteau's Lemma and the first part of Lemma 4.1 give us that if $d = 1$, (using the obvious generalisation to $\mathcal{A}^\infty$ of our previous notation)

$$\begin{aligned}
&\lim_{n\to\infty} \int_{(\mathbb{R}^d)^j} dx_1 \ldots dx_j\, \psi(x_1, \ldots, x_j) \mathbb{E}\big[\mathbf{1}_H(\xi^1_{tn}) \cdots \mathbf{1}_H\big(\xi^{N_{tn}}_{tn}\big) \,\big|\, N_0 = j, \xi^1_0 = x_1\sqrt{n}, \ldots, \xi^j_0 = x_j\sqrt{n}\big] \\
&= \lim_{n\to\infty} \int_{(\mathbb{R}^d)^j} dx_1 \ldots dx_j\, \psi(x_1, \ldots, x_j) \mathbb{E}\big[\mathbf{1}_H(\xi^{n,1}_t) \cdots \mathbf{1}_H\big(\xi^{n,N^n_t}_t\big) \big| N^n_0 = j, \xi^{n,1}_0 = x_1, \ldots, \xi^{n,j}_0 = x_j\big] \\
&= \int_{(\mathbb{R}^d)^j} dx_1 \ldots dx_j\, \psi(x_1, \ldots, x_j) \mathbb{E}\big[\mathbf{1}_H(\xi^{\infty,1}_t) \cdots \mathbf{1}_H\big(\xi^{\infty,N^\infty_t}_t\big) \,\big|\, N^\infty_0 = j, \xi^{\infty,1}_0 = x_1, \ldots, \xi^{\infty,j}_0 = x_j\big].
\end{aligned}$$

Now, Theorem 4.1 in [Eva97] guarantees that there exists a unique Ξ-valued Markov process starting from (the equivalence class of) $\mathbf{1}_H(x)$ and dual to $\mathcal{A}^\infty$ through the relations (8). Let us call this process $\rho^{(2)}$. Using the more compact notation of §3.1, we obtain that for every $j \in \mathbb{N}$ and $\psi \in C((\mathbb{R}^d)^j) \cap L^1(dx^{\otimes j})$,

$$\lim_{n\to\infty} \mathbb{E}\big[I_j(\rho^n_t; \psi)\big] = \mathbb{E}\big[I_j(\rho^{(2)}_t; \psi)\big].$$

Since this family of test functions in dense in $C(\Xi)$ (c.f. §3.1), we can conclude that $\rho^n_t \xrightarrow{\mathcal{L}} \rho^{(2)}_t$ as $n \to \infty$. It is then straightforward to check that the conditions of Lemma 3.2 (i) are satisfied, and so for a.e. $x \in \mathbb{R}^d$, $w^{(2)}(t, x)$ is a Bernoulli random variable with parameter

$$\mathbb{P}_x\big[\xi^\infty_t \in H\big] = \mathbb{P}_x\big[X_{\sigma^2 t} \in H\big] = p^2(t, x).$$

Moreover, by Lemma 4.1 and (9), the correlations between the values of $w^{(2)}(t, \cdot)$ at different sites can be described as follows. For every $j \in \mathbb{N}$ and Lebesgue-a.e. $(x_1, \ldots, x_j)$,

$$\begin{aligned}
&\mathbb{E}\big[\, w^{(2)}(t, x_1) \ldots w^{(2)}(t, x_j) \,\big] \\
&\quad = \mathbb{E}\Big[w^{(2)}(0, \xi^{\infty,1}_t) \cdots w^{(2)}\big(0, \xi^{\infty,N^\infty_t}_t\big) \,\Big|\, N^\infty_0 = j,\ \xi^{\infty,1}_0 = x_1, \ldots, \xi^{\infty,j}_0 = x_j\Big] \\
&\quad = \mathbb{P}\Big[\xi^{\infty,i}_t \in H,\ \forall i \in \{1, \ldots, N^\infty_t\} \,\Big|\, N^\infty_0 = j,\ \xi^{\infty,1}_0 = x_1, \ldots, \xi^{\infty,j}_0 = x_j\Big]. \qquad (15)
\end{aligned}$$

Since we are dealing with Bernoulli random variables, equation (15) completely characterizes these correlations.

If $d \geq 2$, by the same chain of arguments (using this time Lemma 4.2), we obtain

$$\lim_{n\to\infty} \int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j\, \psi(x_1,\dots,x_j)\mathbb{E}\big[\mathbf{1}_H(\xi^1_{tn}) \cdots \mathbf{1}_H(\xi^{N_{tn}}_{tn}) \,\big|\, N_0 = j, \xi^1_0 = x_1\sqrt{n}, \dots, \xi^j_0 = x_j\sqrt{n}\big]$$
$$= \int_{(\mathbb{R}^d)^j} dx_1 \dots dx_j\, \psi(x_1,\dots,x_j)\mathbb{P}_{x_1}\big[\xi^{\infty,1}_t \in H\big] \cdots \mathbb{P}_{x_j}\big[\xi^{\infty,j}_t \in H\big].$$

Here again, these equalities guarantee the convergence in law of ρ^n_t towards the value at time t of the unique Ξ-valued Markov process $\rho^{(2)}$ starting from $\mathbf{1}_H(x)$ and dual to the system $\mathcal{A}^\infty$ of independent Brownian motions which never coalesce. Lemma 3.2 (ii) then applies and gives us that for a.e. $x \in \mathbb{R}^d$, $w^{(2)}(t,x)$ is the deterministic constant $p^2(t,x)$.

So far, we have obtained the desired convergence at a given time $t > 0$, and the form of the local densities of 1's in the limit. It remains to show that the convergence holds true for finitely many times $0 \leq t_1 \leq \cdots \leq t_k$. Because functions of the form $I_j(\cdot\,;\,\psi)$ are dense in $C(\Xi)$, we need only show that for every $j_1,\dots,j_k$ and $\psi_1,\dots,\psi_k$,

$$\lim_{n\to\infty} \mathbb{E}\Big[\prod_{i=1}^k I_{j_i}\big(\rho^n_{t_i}\,;\,\psi_i\big)\Big] = \mathbb{E}\Big[\prod_{i=1}^k I_{j_i}\big(\rho^{(2)}_{t_i}\,;\,\psi_i\big)\Big]. \tag{16}$$

Therefore, let us fix $j_1,\dots,j_k$ and $\psi_1,\dots,\psi_k$ such that $\psi_i \in C((\mathbb{R}^d)^{j_i}) \cap L^1(dx^{\otimes j_i})$. To simplify notation, we write $\underline{x}^i$ for the vector $(x^i_1,\dots,x^i_{j_i})$ and $W^n_i(\underline{x}^i)$ for the product $\prod_{l=1}^{j_i} w^n(t_i,x^i_l)$. We will occasionally abuse notations and write $\underline{x}^i \cup \underline{x}^j$ for the concatenation of the vectors $\underline{x}^i$ and $\underline{x}^j$. Our strategy is to use duality again, but now with the genealogical process described in the second part of Lemmas 4.1 and 4.2. Once again, to simplify our notation, let us denote the law of $\mathcal{A}^n$ (resp., $\mathcal{A}^n_t$) starting from j lineages at locations $\underline{x} = (x_1,\dots,x_j)$ by $\mathbb{P}^n_{\underline{x}}$ (resp., $\mathbb{P}^n_{\underline{x},t}$). Hence recall that $\mathbb{P}^n_{\underline{x}}$ is a distribution over $\bigcup_{m\geq 0}\{m\} \times (\mathbb{R}^d)^m$. Using the Markov property of w at time $t_{k-1}n$ and the duality property (8), we can write

$$\mathbb{E}\Big[\prod_{i=1}^k I_{j_i}\big(\rho^n_{t_i}\,;\,\psi_i\big)\Big]$$
$$= \int \dots \int d\underline{x}^1 \cdots d\underline{x}^k\ \psi_1(\underline{x}^1)\cdots\psi_k(\underline{x}^k)$$
$$\times\mathbb{E}\Big[\Big\{\prod_{i=1}^{k-1} W^n_i\big(\underline{x}^i\big)\Big\}\ \mathbb{E}^n_{\underline{x}^k}\Big[w^n\big(t_{k-1},\xi^{n,1}_{t_k-t_{k-1}}\big)\cdots\ w^n\Big(t_{k-1},\xi^{n,N^n_{t_k-t_{k-1}}}_{t_k-t_{k-1}}\Big)\Big]\Big]$$
$$= \int \dots \int d\underline{x}^1 \cdots d\underline{x}^k\ \psi_1(\underline{x}^1)\cdots\psi_k(\underline{x}^k) \int d\mathbb{P}^n_{\underline{x}^k,t_k-t_{k-1}}\big(m_{k-1},y^{k-1}_1,\dots,y^{k-1}_{m_{k-1}}\big)$$
$$\mathbb{E}\Big[\Big\{\prod_{i=1}^{k-2} W^n_i\big(\underline{x}^i\big)\Big\}\ w^n\big(t_{k-1},x^{k-1}_1\big)\cdots w^n\big(t_{k-1},x^{k-1}_{j_{k-1}}\big)w^n\big(t_{k-1},y^{k-1}_1\big)\cdots\ w^n\big(t_{k-1},y^{k-1}_{m_{k-1}}\big)\Big].$$

Since the law of the locations at time $t_k - t_{k-1}$ of the $N^n_{t_k-t_{k-1}}$ lineages is absolutely continuous with respect to Lebesgue measure, we can carry on the recursion and use the Markov property

(this time at time t_{k-2}) and duality to write the quantity above as

$$\int \dots \int d\underline{x}^1 \cdots d\underline{x}^k \ \psi_1(\underline{x}^1) \cdots \psi_k(\underline{x}^k) \int d\mathbb{P}^n_{\underline{x}^k, t_k - t_{k-1}}\big(m_{k-1}, y_1^{k-1}, \dots, y_{m_{k-1}}^{k-1}\big)$$
$$\times \int d\mathbb{P}^n_{\underline{x}^{k-1} \cup \underline{y}^{k-1}, t_{k-1} - t_{k-2}}\big(m_{k-2}, y_1^{k-2}, \dots, y_{m_{k-2}}^{k-2}\big) \mathbb{E}\Big[\Big\{ \prod_{i=1}^{k-3} W_i^n\big(\underline{x}^i\big) \Big\}$$
$$\times w^n\big(t_{k-2}, x_1^{k-2}\big) \cdots \ w^n\big(t_{k-2}, x_{j_{k-2}}^{k-2}\big) w^n\big(t_{k-2}, y_1^{k-2}\big) \cdots \ w^n\big(t_{k-2}, y_{m_{k-2}}^{k-2}\big)\Big]$$
$$= \int \dots \int d\underline{x}^1 \cdots d\underline{x}^k \ \psi_1(\underline{x}^1) \cdots \psi_k(\underline{x}^k) \int d\mathbb{P}^n_{\underline{x}^k, t_k - t_{k-1}}\big(m_{k-1}, y_1^{k-1}, \dots, y_{m_{k-1}}^{k-1}\big)$$
$$\times \int \cdots \int d\mathbb{P}^n_{\underline{x}^1 \cup \underline{y}^1, t_1}\big(m_0, y_1^0, \dots, y_{m_0}^0\big) \mathbb{E}\Big[w^n\big(0, y_1^0\big) \cdots w^n\big(0, y_{m_0}^0\big)\Big]. \tag{17}$$

Now, recall the family of processes $\hat{\mathcal{A}}^n$ introduced in the second part of Lemmas 4.1 and 4.2. Let us denote the times of appearance and the locations of the additional lineages in the form $(\tau_1, \underline{z}^1), \dots, (\tau_k, \underline{z}^k)$. Using (recursively) the Markov property of $\hat{\mathcal{A}}^n$, we obtain that the quantity on the right-hand side of (17) is equal to

$$\int \dots \int d\underline{x}^1 \cdots d\underline{x}^k \ \psi_1(\underline{x}^1) \cdots \psi_k(\underline{x}^k)$$
$$\times \mathbb{E}\Big[w^n\big(0, \hat{\xi}_{t_k}^{n,1}\big) \cdots \ w^n\Big(0, \hat{\xi}_{t_k}^{n, \hat{N}^n_{t_k}}\Big) \ \Big| \ \big(0, \underline{x}^k\big), \big(t_k - t_{k-1}, \underline{x}^{k-1}\big), \dots, \big(t_k - t_1, \underline{x}^1\big)\Big].$$

Let us now conclude when $d = 1$ (the reasoning is exactly the same when $d \geq 2$). Recall that for every $n \in \mathbb{N}$, $w^n(0, \cdot) = \mathbf{1}_H(\cdot) = w^{(2)}(0, \cdot)$. By the second part of Lemma 4.1 and the Dominated Convergence Theorem (and the fact that the boundary of H has zero Lebesgue measure), we obtain that

$$\lim_{n \to \infty} \mathbb{E}\left[\prod_{i=1}^k I_{j_i}\big(\rho^n_{t_i} \, ; \, \psi_i\big)\right]$$
$$= \int \dots \int d\underline{x}^1 \cdots d\underline{x}^k \ \psi_1(\underline{x}^1) \cdots \psi_k(\underline{x}^k)$$
$$\times \mathbb{E}\Big[w^{(2)}\big(0, \hat{\xi}_{t_k}^{\infty,1}\big) \cdots w^{(2)}\Big(0, \hat{\xi}_{t_k}^{\infty, \hat{N}^\infty_{t_k}}\Big) \ \Big| \ \big(0, \underline{x}^k\big), \big(t_k - t_{k-1}, \underline{x}^{k-1}\big), \dots, \big(t_k - t_1, \underline{x}^1\big)\Big].$$

Analogous calculations using the duality between $\mathcal{A}^\infty$ and $\rho^{(2)}$ lead to (16). This completes the proof of Theorem 1.1. □

It remains to prove Lemmas 4.1 and 4.2. Let us start with the latter, which is somewhat simpler, but contains the main ingredients of both proofs.

Proof of Lemma 4.2. Let $x_1, \dots, x_k$ be k distinct points of $\mathbb{R}^d$. Suppose that $\mathcal{A}^n$ starts from k lineages at locations $x_1\sqrt{n}, \dots, x_k\sqrt{n}$. First, since a single lineage ξ follows a finite-rate homogeneous jump process whose jumps are uniformly bounded by $2r$, standard arguments guarantee that $\xi^n = (n^{-1/2}\xi_{tn})_{t \geq 0}$ converges in distribution to Brownian motion with clock speed σ^2 given in (7).

Second, observe that two lineages can be hit by the same event (and possibly coalesce) only if they lie at distance at most $2r$ of each other. Consequently, as long as they are at distance greater than $2r$ they evolve independently, according to the law of the motion of a single lineage.

Hence, let us define $n\tau_n$ to be the first time at which at least two of the k initial lineages are within distance at most $2r$ of one another. Equivalently, τ_n is the first time at which at least two lineages of $\mathcal{A}^n$ are at separation at most $2r/\sqrt{n}$. We wish to show that for any $t \geq 0$, $\mathbb{P}^n_{\underline{x}}[\tau_n \leq t] \to 0$ as $n \to \infty$.

To this end, note that until time τ_n, the motions of the rescaled lineages $\xi^{n,1}, \ldots, \xi^{n,k}$ can be embedded in the paths of independent standard Brownian motions $X^1, \ldots, X^k$ starting from $x_1, \ldots, x_k$ (we use the same Brownian motions for all n). Indeed, for each path i we proceed as follows (this construction is in the spirit of the one-dimensional Skorokhod Embedding Theorem, see e.g. [Bil95]). Let $(R^{n,i}_j)_{j\geq 1}$ be a sequence of i.i.d. random variables (independent of X^i) distributed according to the law of the radius of a typical jump of ξ^n, and let us define a sequence $\{s^n_{i,j}, j \geq 0\}$ of random times, recursively, by

1. $s^n_{i,0} := 0$,
2. for every $j \geq 1$, $s^n_{i,j}$ is the first time greater than $s^n_{i,j-1}$ at which X^i exits the ball $B\big(X^i_{s^n_{i,j-1}}, R^{n,i}_j\big)$.

By rotational symmetry of the law of a jump of $\xi^{n,i}$, conditional on its radius being γ the location of $\xi^{n,i}$ just after the jump is uniformly distributed over the sphere $\partial B(\xi^{n,i}_{t-}, \gamma)$. Likewise, conditional on the variable $R^{n,i}_j$ being equal to γ, the location of $X^i_{s^n_{i,j}}$ is uniformly distributed over $\partial B(X^i_{s^n_{i,j-1}}, \gamma)$. Consequently, by comparing their jump rates and their jump distributions, one can show that for every $i \in \{1, \ldots, k\}$ the processes $(\xi^{n,i}_t)_{t\geq 0}$ and $\big(X^i_{s^n_{i,j(n,i,t)}}\big)_{t\geq 0}$ have the same laws, where $(j(n,i,t))_{t\geq 0}$ is a Poisson process with intensity nuV_r (recall from (1) that uV_r is the jump rate of an unrescaled lineage under the conditions of Case A, where V_r is the volume of a ball of radius r). Since the lineages $\xi^{n,1}, \ldots, \xi^{n,j}$ evolve independently until time τ_n, we can ask that the Poisson processes $\{j(n,1,\cdot), \ldots, j(n,k,\cdot)\}$ should be independent and the embedding holds for all $i \in \{1, \ldots, k\}$ simultaneously until the first time t such that

$$\big|X^i_{s^n_{i,j(n,i,t)}} - X^m_{s^n_{m,j(n,m,t)}}\big| \leq 2r/\sqrt{n} \qquad \text{for some } i \neq m.$$

Now, each rescaled lineage makes jumps of size at most $2rn^{-1/2}$ at rate $\mathcal{O}(n)$. Hence, each difference $s^n_{i,j} - s^n_{i,j-1}$ is the exit time of Brownian motion from a ball of radius $\mathcal{O}(n^{-1/2})$, and $s^n_{i,j(n,i,t\wedge\tau_n)}$ is the sum of (morally) $\mathcal{O}(n)$ such quantities, all independent of one another. More formally, if we write R for the (random) radius of a typical jump of an unrescaled lineage and if we notice that the exit time of Brownian motion starting at 0 from a ball $B(0,\gamma)$ is bounded by the first time that one of its coordinates leaves the interval $[-\gamma, \gamma]$, then for all $n \in \mathbb{N}$ and all $1 \leq i \leq k$ we can write

$$\mathbb{E}\big[s^n_{i,1}\big] \leq d\,\mathbb{E}\Big[\big(R^{n,i}_1\big)^2\Big] = \frac{d}{n}\,\mathbb{E}\big[R^2\big] \leq \frac{4dr^2}{n},$$

where the first inequality uses the property that the exit time from $[-\gamma, \gamma]$ of one-dimensional Brownian motion starting at 0 has expectation γ^2. By the independence of X_i and the Poisson processes, this yields that for all n and i,

$$\mathbb{E}\big[s^n_{i,j(n,i,t\wedge\tau_n)}\big] = \mathbb{E}\big[j(n,i,t\wedge\tau_n)\big].\mathbb{E}\big[s^n_{i,1}\big] \leq 4duV_r r^2.$$

To conclude our proof, let us observe that $\mathbb{P}^n_{\underline{x}}[\tau_n \leq t]$ is bounded by the probability that at least two of the k independent Brownian motions $X^1, \ldots, X^k$ come within distance $2rn^{-1/2}$

before time $\min\{s^n_{i,j(n,i,t)},\ 1 \le i \le k\}$. But if $\tilde{\tau}_n$ denotes the first time at which two independent Brownian motions starting at $x_1 \neq x_2$ come within distance $2rn^{-1/2}$, for every $T \ge 0$ we have

$$\lim_{n\to\infty} \mathbb{P}_{x_1,x_2}\big[\tilde{\tau}_n \le T\big] = 0.$$

Hence, the probability that at least two out of k independent Brownian motions come within distance $2rn^{-1/2}$ before any given time T also tends to 0, and thanks to the uniform bound on the expectation of $s^n_{i,j(n,i,t\wedge\tau_n)}$ (together with the Markov inequality), it is straightforward to obtain that for any $t \ge 0$

$$\lim_{n\to\infty} \mathbb{P}^n_{\underline{x}}\big[\tau_n \le t\big] = 0.$$

We have thus shown that with probability growing to 1 as $n \to \infty$, until a given time $t \ge 0$ the k ancestral lineages evolve as if they were independent. Since the law of each $\xi^{n,i}$ converges to that of Brownian motion with clock speed σ^2, the convergence of the one-dimensional distributions of $\mathcal{A}^n$ to those of a collection of k independent Brownian motions is proved.

The proofs of the convergence of the finite-dimensional distributions and that of the second part of Lemma 4.2 follow the same lines, using the Markov property of each $\mathcal{A}^n$ at suitable times. Details are left to the reader. □

Proof of Lemma 4.1. Once again we start with the one-dimensional distributions, and proceed by recursion on the number m of lineages of $\mathcal{A}^n$. As in the proof of Lemma 4.2, before rescaling each lineage follows a homogeneous symmetric (finite rate) jump process, whose jumps have length at most $2r$, and so $\xi^n = (n^{-1/2}\xi_{nt})_{t\ge0}$ converges in distribution to Brownian motion with clock speed σ^2 as n tends to infinity.

Let us consider the case $m = 2$. As we saw in the proof of Lemma 4.2, the two rescaled lineages evolve independently until they come within distance $2rn^{-1/2}$ of one another. Let us first show that this 'meeting' time converges to the meeting time (at distance 0) of two independent Brownian motions starting at x_1 and x_2 and with clock speed σ^2, and secondly that coalescence is quasi-instantaneous once the lineages are gathered at this distance.

For the first claim, let us write τ_n for the time at which $\xi^{n,1}$ and $\xi^{n,2}$ first come within distance at most $4rn^{-1/2}$ of one another (note the constant 4 instead of 2, which we shall need later for purely technical reasons). Because the motion of a single lineage is a symmetric jump process, until τ_n the law of the difference $\xi^{n,1} - \xi^{n,2}$ is the same as that of the motion of a single rescaled lineage, run at speed 2. Let X be a standard one-dimensional Brownian motion, starting from $x_1 - x_2$ and independent of all ξ^n's. Using anew the construction introduced in the proof of Lemma 4.2, for every n we can find a sequence of random times $\{s^n_j, j \ge 0\}$ such that $(\xi^{n,1}_t - \xi^{n,2}_t)_{t\ge0}$ has the same law as $(X_{s^n_{j(n,t)}})_{t\ge0}$, where $j(n,\cdot)$ is a Poisson process, independent of X and with intensity $4nru$ (that is, twice the jump rate of a single rescaled lineage). Recall from the proof of Lemma 4.2 that for every $n \in \mathbb{N}$, the random variables $s^n_j - s^n_{j-1}$, $j \ge 1$, are i.i.d and if R is distributed like the radius of a typical jump of ξ, we have $\mathbb{E}[ns^n_1] = \mathbb{E}[R^2] < \infty$. Let $t \ge 0$, and, as a first step, let us show that $s^n_{j(n,t)}$ converges in probability towards $2\sigma^2 t$ as n grows to infinity. The second step will then consist of proving that, for every $t \ge 0$, the probability that $\tau_n > t$ tends to the probability that the hitting time of 0 by X is greater than $2\sigma^2 t$. This will give us the desired result.

By definition, $j(n,t)$ is a Poisson random variable with parameter $(4nur)t$. By the Central Limit Theorem, we therefore have that

$$n^{-1/2}\big(j(n,t) - 4nurt\big) \xrightarrow{(d)} \mathcal{N}(0, 4urt). \tag{18}$$

Now, recalling the properties of the $s_i^n - s_{i-1}^n$'s expounded above, by the Strong Law of Large Numbers we have

$$s_{\lfloor 4nurt \rfloor}^n = \frac{1}{n} \sum_{i=1}^{\lfloor 4nurt \rfloor} n\big(s_i^n - s_{i-1}^n\big) \xrightarrow{\text{a.s.}} 4urt \times \mathbb{E}[R^2] \qquad \text{as } n \to \infty, \tag{19}$$

where $\lfloor z \rfloor$ denotes the integer part of z. But σ^2 is defined in (7) as the variance of the displacement at time 1 of a single unrescaled lineage, and so

$$\sigma^2 = 2ur\,\mathbb{E}[R^2],$$

which shows that the limit in (19) is equal to $2\sigma^2 t$. To conclude the first step, observe that $|s_{j(n,t)}^n - s_{\lfloor 4nurt \rfloor}^n|$ is the sum of $|j(n,t) - \lfloor 4nurt \rfloor|$ i.i.d. terms of the form $s_i^n - s_{i-1}^n$, all of them independent of $j(n,t)$, so that for every $\varepsilon > 0$ and every $n \geq 1$ we have

$$\mathbb{P}\big[\big|s_{j(n,t)}^n - s_{\lfloor 4nurt \rfloor}^n\big| > \varepsilon\big] \leq \mathbb{P}\big[|j(n,t) - 4nurt| > n^{3/4}\big] + \mathbb{P}\left[\sum_{i=1}^{n^{3/4}} \big(s_i^n - s_{i-1}^n\big) > \varepsilon\right].$$

As $n \to \infty$, the first term on the right-hand side tends to 0 by (18), while Markov's inequality gives us that

$$\mathbb{P}\left[\sum_{i=1}^{n^{3/4}} \big(s_i^n - s_{i-1}^n\big) > \varepsilon\right] \leq \frac{1}{\varepsilon}\,\mathbb{E}\left[\sum_{i=1}^{n^{3/4}} \big(s_i^n - s_{i-1}^n\big)\right] = \frac{C}{\varepsilon n^{1/4}} \longrightarrow 0.$$

Since this is true for any $\varepsilon > 0$, $s_{j(n,t)}^n - s_{\lfloor 4nurt \rfloor}^n$ converges in probability to 0. But we have shown that $s_{\lfloor 4nurt \rfloor}^n$ converges a.s. to $2\sigma^2 t$, and so we obtain that $s_{j(n,t)}^n$ converges in probability to $2\sigma^2 t$, as required.

As explained above, we can now use this result to show that τ_n converges in distribution to the hitting time of 0 by $(X_{2\sigma^2 t})_{t \geq 0}$. Indeed, by construction of the random times s_i^n and the fact that the rescaled jumps of a lineage are bounded by $2r/\sqrt{n}$, for any $i \geq 1$ the Brownian motion X cannot move to a distance greater than $2r/\sqrt{n}$ from $X_{s_{i-1}^n}$ before time s_i^n. Thus, if τ_0 denotes the hitting time of 0 by X, we have

$$\mathbb{P}_{x_1 - x_2}[\tau_n > t] \leq \mathbb{P}_{x_1 - x_2}\big[\tau_0 > s_{j(n,t)}^n\big].$$

But we showed that $s_{j(n,t)}^n$ converges in probability towards $2\sigma^2 t$ as $n \to \infty$, and so

$$\limsup_{n \to \infty}\ \mathbb{P}_{x_1 - x_2}[\tau_n > t] \leq \mathbb{P}_{x_1 - x_2}\big[\tau_0 > 2\sigma^2 t\big]. \tag{20}$$

On the other hand, for every $\varepsilon \in (0, |x_1 - x_2|/2)$ and every n large enough, we can write

$$\begin{aligned}\mathbb{P}_{x_1 - x_2}[\tau_n > t] &\geq \mathbb{P}_{x_1 - x_2}\big[X \text{ does not enter } B(0, 4r/\sqrt{n}) \text{ before } s_{j(n,t)}^n\big] \\ &\geq \mathbb{P}_{x_1 - x_2}\big[X \text{ does not enter } B(0, \varepsilon) \text{ before } s_{j(n,t)}^n\big].\end{aligned}$$

Again, we can deduce from the convergence in probability of $s_{j(n,t)}^n$ to $2\sigma^2 t$ that

$$\liminf_{n \to \infty}\ \mathbb{P}_{x_1 - x_2}[\tau_n > t] \geq \mathbb{P}_{x_1 - x_2}\big[X \text{ does not enter } B(0, \varepsilon) \text{ before } 2\sigma^2 t\big].$$

This inequality holds for every small $\varepsilon > 0$, and by the point recurrence of one-dimensional Brownian motion, we can conclude that

$$\liminf_{n\to\infty} \mathbb{P}_{x_1-x_2}[\tau_n > t] \geq \mathbb{P}_{x_1-x_2}\big[\tau_0 > 2\sigma^2 t\big]. \tag{21}$$

Together with (20), we obtain that for every $t > 0$

$$\lim_{n\to\infty} \mathbb{P}_{x_1-x_2}[\tau_n > t] = \mathbb{P}_{x_1-x_2}\big[\tau_0 > 2\sigma^2 t\big], \tag{22}$$

from which we can conclude that the 'meeting time at distance $4r/\sqrt{n}$' of two rescaled lineages starting at x_1 and x_2 converges in distribution to the hitting time of 0 by Brownian motion with clock speed $2\sigma^2$, or equivalently to the meeting time of 2 independent Brownian motions each of clock speed σ^2.

Let us now prove our second claim; that is, let us show that once at distance at most $4r/\sqrt{n}$, the additional time the two lineages need to merge becomes negligible as n tends to infinity. Because the proof is highly reminiscent of that of Proposition 6.4(b) in [BEV10], we only outline the main steps here. Let us work with the unrescaled lineages, and suppose they start at distance at most $4r$ of each other. First, it is not difficult to convince oneself that the first time at which the two lineages are at separation less than $2r$ is of order $\mathcal{O}(1)$, 'uniformly' over all initial locations which are at separation at most $4r$. Once close together, they become correlated, because they can be hit by the same reproduction event. But for the same reason, they have a positive probability of being affected by the same event and of coalescing before separating again to distance at least $2r$. If they do coalesce, the additional time they had to wait for this event is also of order $\mathcal{O}(1)$. If they separate rather than coalescing, then again the time they need to come back to separation less than $2r$ is of order $\mathcal{O}(1)$, and once 'gathered' they have a positive chance to coalesce before separating, and so on. In the end, the number of excursions of $\xi^1 - \xi^2$ out of $B(0, 2r)$ before the two lineages merge can be stochastically bounded by a geometric random variable, and each of the finitely many excursions and incursions lasts a time of order $\mathcal{O}(1)$. This tells us that for every $\varepsilon > 0$, one can find $T(\varepsilon) > 0$ such that

$$\sup_{|y_1-y_2|\leq 4r} \mathbb{P}_{(y_1,y_2)}\big[\xi^1 \text{ and } \xi^2 \text{ do not coalesce before } T(\varepsilon)\big] \leq \varepsilon.$$

Rephrasing the above inequality in terms of the rescaled lineages, we obtain that, for every $n \geq 1$,

$$\sup_{|z_1-z_2|\leq 4r/\sqrt{n}} \mathbb{P}_{(z_1,z_2)}\big[\xi^{n,1} \text{ and } \xi^{n,2} \text{ do not coalesce before } T(\varepsilon)/n\big] \leq \varepsilon. \tag{23}$$

Finally, if τ_n^c denotes the coalescence time of $\xi^{n,1}$ and $\xi^{n,2}$, using the strong Markov property of $(\xi^{n,1}, \xi^{n,2})$ at time τ_n, we have, for every $t > 0$,

$$\mathbb{P}_{(x_1,x_2)}[\tau_n^c - \tau_n > t] = \mathbb{E}_{(x_1,x_2)}\Big[\mathbf{1}_{\{\tau_n<\infty\}}\mathbb{P}_{(\xi^{n,1}_{\tau_n},\xi^{n,2}_{\tau_n})}[\tau_n^c > t]\Big].$$

By (23), the probability inside the expectation tends to 0 as $n \to \infty$, and so does the quantity on the left-hand side (by dominated convergence). Hence, $\tau_n^c - \tau_n$ converges to 0 in probability. This concludes the proof of the first part of Lemma 4.1 when $m = 2$: in the limit, the two lineages follow independent Brownian motions run at clock speed σ^2 until the first time at which they meet, which is also the time at which they coalesce by the convergence of $\tau_n^c - \tau_n$ to 0.

We now proceed by induction. Suppose we know that the result of Lemma 4.1 holds true for a system of $m-1$ lineages. Let $x_1 < \ldots < x_m$ be m distinct points of $\mathbb{R}$ and suppose that m

lineages start from these locations. Because the lineages 'choose' to take part in an event that encompasses them independently of one another, the law of the restriction of the system started from m lineages to that started from $m-1$ lineages at $x_1, \ldots, x_{m-1}$ is the same as that of the $(m-1)$-system starting from $x_1, \ldots, x_{m-1}$. (This is the 'consistency' of the genealogical process described below Lemma 3.2). Hence, our inductive hypothesis tells us that the restricted process converges to a system of (initially) $m-1$ independent Brownian motions with clock speed σ^2, that coalesce instantaneously upon meeting. Now, as we explained several times already, the motion of the m-th lineage, starting at the right-most location x_m, is independent of that of the others until the first time, τ_n, at which it comes to within distance $2r/\sqrt{n}$ of another lineage. But with probability tending to 1, the right-most lineage among those that started from $x_1, \ldots, x_{m-1}$ is the lineage ancestral to the individual sampled in x_{m-1}. Indeed, our inductive hypothesis guarantees that the probability that the lineage starting from x_{m-1} jumps over a lineage on its left without coalescing with it tends to 0 as n tends to infinity. Again by consistency of the genealogical process, when singled out, the motion of lineage $m-1$ has the same law as the process ξ^n (that is, a typical single lineage), and so we can focus on the two right-most lineages and use the results obtained for $m = 2$ to conclude: their meeting time at distance at most $4r/\sqrt{n}$ converges in distribution to the meeting time of two independent Brownian motions run at clock speed σ^2, and in the limit this meeting time is also the coalescence time of the two lineages. But this is precisely the evolution of a system of (initially) m independent Brownian motions which coalesce instantaneously when they meet, and so the desired convergence also holds for a system starting with m lineages.

As in the proof of Lemma 4.2, the other points of Lemma 4.1 are obtained by using the convergence of the one-dimensional distributions and the Markov property at suitable times. □

5 Heavy-tailed case

In this section, we prove Theorem 1.5 and give some properties of the limiting genealogical process, which are of independent interest. Recall that the fraction of individuals affected by an event is set constant, equal to $u \in (0, 1]$, and the radii of the events are sampled according to the intensity measure

$$\mu(dr) = r^{-\alpha-d-1}\mathbf{1}_{\{r\geq 1\}}\, dr,$$

where d is the dimension of the geographical space.

As in the proof of Theorem 1.1, due to the duality relations (8) we need only establish the asymptotic behaviour of the rescaled genealogical process $(\mathcal{A}_t^n)_{t\geq 0}$ of a finite sample of individuals, defined in our previous notation by

$$\mathcal{A}_t^n \equiv \big(\xi_t^{n,1}, \ldots, \xi_t^{n,N_t^n}\big) := \big(n^{-1/\alpha}\xi_{nt}^1, \ldots, n^{-1/\alpha}\xi_{nt}^{N_{nt}}\big).$$

In words, we speed up time by a factor n and scale down the spatial locations of the lineages by $n^{1/\alpha}$. Indeed, if we can show that the finite-dimensional distributions of $\mathcal{A}^n$ converge to those of a system of coalescing processes $\mathcal{A}^\infty$ that has sufficiently nice properties (i.e., which can be used to construct a dual Ξ-valued process $\rho^{(\alpha)}$ using the technique of [Eva97]), then the same arguments as those used in the proof of Theorem 1.1 will grant us the convergence of the finite-dimensional distributions of ρ^n to those of $\rho^{(\alpha)}$. Then it will remain to show that $\mathcal{A}^\infty$ satisfies the conditions of Lemma 3.2(i) to obtain the desired form for the local densities of 1's, $w^{(\alpha)}(t, x)$, and to use (9) to characterize the correlations between these Bernoulli random variables. Hence, the crucial step is to prove the following proposition.

Proposition 5.1. *There exists a system $\mathcal{A}^\infty$ of coalescing symmetric α-stable Lévy processes such that*

$$\mathcal{A}^n \to \mathcal{A}^\infty, \qquad \text{as } n \to \infty,$$

in the sense of weak convergence of the finite-dimensional distributions. Moreover, if we define the process $\hat{\mathcal{A}}^n$ and $\hat{\mathcal{A}}^\infty$ in an analogous way to the corresponding processes in Lemmas 4.1 and 4.2, we also have convergence of the one-dimensional distributions of $\hat{\mathcal{A}}^n$ to those of $\hat{\mathcal{A}}^\infty$.

Proof of Proposition 5.1. Our aim is to write down the generator $\mathcal{G}^n$ of $\mathcal{A}^n$, and to show that it converges to the generator of a system of coalescing symmetric α-stable processes. Up to now, we were able to be rather vague about the precise representation of the ancestral lineages, but in order to write down a sensible generator we now need to be more precise. Suppose we start with k lineages. The system at any time $t \geq 0$ is represented by a marked partition of $\{1, \dots, k\}$. Each block of $\mathcal{A}^n_t$ contains the labels of all individuals in the initial sample which have the same ancestor at time t in the past (that is, whose ancestral lineages merged before t), and the mark associated to the block gives the spatial location of this ancestor at time t.

Since only the lineages present in the area hit by an event can be affected by this event, for every $y \in \mathbb{R}^d$, $r > 0$ and every marked partition A let us write $J(y, r, A)$ for the set of indices of lineages (blocks) of A whose mark belongs to $B(y, r)$ (to index the blocks of A, we rank them in increasing order of the smallest label that each contains). For convenience, we shall also use the notation $J_n(y, r, A) := J(n^{-1/\alpha}y, n^{-1/\alpha}r, A)$. Next, if A contains m blocks and $I \subset \{1, \dots, m\}$, then for every $z \in \mathbb{R}^d$ we write $\Phi_I(A, z)$ for the marked partition obtained by merging all blocks of A indexed by $i \in I$ and by assigning the mark z to this new block (the other blocks and marks remain unchanged). For instance, if $A = \{(\{1,5\}, x_1), (\{2,3\}, x_2), (\{4,6\}, x_3), (\{7\}, x_4)\}$ and $I = \{1, 4\}$, then

$$\Phi_I(A, z) = \big\{(\{1,5,7\}, z), (\{2,3\}, x_2), (\{4,6\}, x_3)\big\}.$$

Finally, we write $|I|$ for the cardinality of the set I, and we recall that V_r denotes the volume of a ball of radius r.

Because lineages jump and merge at finite rate, the generator $\mathcal{G}$ of the system of unrescaled lineages $(\mathcal{A}_t)_{t\geq 0}$ can be expressed as follows. For every bounded measurable function f and every marked partition A (of some finite set $\{1, \dots, k\}$),

$$\mathcal{G}f(A) = \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A)} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A, z)) - f(A)\big], \tag{24}$$

where in the above and what follows we write J as a shorthand notation for $J(y, r, A)$. Indeed, if an event occurs in $B(y, r)$ and the parent is chosen at location z, then every lineage present in this area is affected by the event with probability u, independently of each other, and all lineages that are affected merge and jump onto the location z of their parent.

Mutiplying time by n and marks by $n^{-1/\alpha}$, we obtain from the expression in (24) that the generator of $\mathcal{A}^n$ is given, for every f and A as above, by

$$\mathcal{G}^n f(A) = n \int_{\mathbb{R}^d} dy \int_0^\infty \mu(dr) \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J_n(y,r,A)} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A, n^{-1/\alpha}z)) - f(A)\big].$$

To see where the sum comes from, observe that an unrescaled mark belongs to $B(y, r)$ iff its rescaled version belongs to $B(n^{-1/\alpha}y, n^{-1/\alpha}r)$, and that the affected (rescaled) lineages jump

onto $n^{-1/\alpha}z$ when their unrescaled counterparts jump to z. Making the change of variables $z' = n^{-1/\alpha}z$, and then $y' = n^{-1/\alpha}y$ and $r' = n^{-1/\alpha}r$, we obtain that $\mathcal{G}^n(A)$ is equal to

$$\begin{aligned}
&n^{1+\frac{d}{\alpha}} \int_{\mathbb{R}^d} dy \int_1^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(n^{-1/\alpha}y, n^{-1/\alpha}r)} \frac{dz}{V_r} \sum_{I \subset J_n(y,r,A)} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A,z)) - f(A)\big] \\
&= \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A)} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A,z)) - f(A)\big] \\
&= \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A), |I|\geq 2} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A,z)) - f(A)\big] \\
&\quad + \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{i \in J(y,r,A)} u(1-u)^{|J|-1}\big[f(\Phi_{\{i\}}(A,z)) - f(A)\big]. \qquad (25)
\end{aligned}$$

Let us define $\delta(A)$ as half of the minimal pairwise distance between marks in A ($\delta(A) := +\infty$ if A contains only one block), and let us show that for every A such that $\delta(A) > 0$ and every f compactly supported and of class C^2 with respect to the marks, $\mathcal{G}^n f(A)$ converges as $n \to \infty$ towards the quantity $\mathcal{G}^\alpha f(A)$ defined by

$$\begin{aligned}
&\mathcal{G}^\alpha f(A) \\
&:= \int_{\mathbb{R}^d} dy \int_0^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A), |I|\geq 2} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A,z)) - f(A)\big] \\
&\quad + u \sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_i \in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \\
&\qquad\qquad \times \int_{B(y,r)} \frac{dz}{V_r} \big[f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z - x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\leq 1\}}\big] \\
&\quad + u \sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_i \in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \int_{B(y,r)} \frac{dz}{V_r} \langle z - x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\leq 1\}}, \qquad (26)
\end{aligned}$$

where $|A|$ denotes the number of blocks of A, x_i is the mark of the i-th block, $\nabla_i f$ is the gradient of f with respect to x_i and $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^d$.

We shall comment on the different terms of $\mathcal{G}^\alpha f(A)$ later. For now, let us show the desired convergence, as well as the finiteness of $\mathcal{G}^\alpha f(A)$. Let us start with the first term on the right-hand side of (25). By definition of $\delta(A)$, a ball of radius $r < \delta(A)$ cannot contain more than 1 lineage (mark), so that the integral over r runs in fact from $n^{-1/\alpha} \vee \delta(A)$ to $+\infty$. For n large enough, this first term is thus equal to

$$\int_{\mathbb{R}^d} dy \int_{\delta(A)}^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A), |I|\geq 2} u^{|I|}(1-u)^{|J\setminus I|}\big[f(\Phi_I(A,z)) - f(A)\big],$$

and so is the first term of $\mathcal{G}^\alpha f(A)$. Since $u \in (0,1]$, f is bounded, the sum over I is finite and since any event location $B(y,r)$ must intersect the compact support of f to have a nonzero contribution to the generator (so that we may restrict the integral over y to some ball $B(0, r + \Delta(f))$ with $\Delta(f)$ depending only on f), there exists a constant $C(f) > 0$, independent of A, such that the absolute value of the first term of $\mathcal{G}^\alpha f(A)$ is bounded by

$$C(f)\, 2^{|A|} \int_{\delta(A)}^\infty \frac{dr}{r^{\alpha+d+1}}\, r^d < \infty. \qquad (27)$$

Now consider the second term on the right-hand side of (25). Let us split it once again into

$$\int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^{\infty} \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{i\in J(y,r,A)} u(1-u)^{|J|-1} \times \big[f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z-x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\le 1\}}\big] \tag{28}$$

$$+ \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^{\infty} \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{i\in J(y,r,A)} u(1-u)^{|J|-1} \langle z-x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\le 1\}}. \tag{29}$$

We rewrite $\sum_{i\in J(y,r,A)}$ as $\sum_{i=1}^{|A|} \mathbf{1}_{\{x_i\in B(y,r)\}}$, and, for n large enough, we split the integral over $r\in[n^{-1/\alpha},\infty)$ in (29) into the integral over $[n^{-1/\alpha},\delta(A))$ and that over $[\delta(A),\infty)$. The second integral is finite for the same reasons as in (27). On the other hand, if $r<\delta(A)$ then $J(y,r,A)\le 1$ for every y, and so the first integral is equal to

$$\begin{aligned} u\sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^{\delta(A)} \frac{dr}{r^{\alpha+d+1}} \mathbf{1}_{\{x_i\in B(y,r)\}} \int_{B(y,r)} \frac{dz}{V_r} \langle z-x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\le 1\}} \\ = u\sum_{i=1}^{|A|} \int_{n^{-1/\alpha}}^{\delta(A)} \frac{dr}{V_r r^{\alpha+d+1}} \int_{B(x_i,1)} dz \int_{\mathbb{R}^d} dy\ \mathbf{1}_{\{|x_i-y|\le r\}} \mathbf{1}_{\{|z-y|\le r\}} \langle z-x_i, \nabla_i f(A)\rangle \\ = u\sum_{i=1}^{|A|} \int_{n^{-1/\alpha}}^{\delta(A)} \frac{dr}{V_r r^{\alpha+d+1}} \int_{B(x_i,1)} dz \ \Big(\mathrm{Vol}\big(B(x_i,r)\cap B(z,r)\big)\Big) \langle z-x_i, \nabla_i f(A)\rangle, \end{aligned}$$

and, by symmetry, the integral over z is equal to 0 for every r. The integral in (29) is thus equal to

$$u\sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_{\delta(A)}^{\infty} \frac{\mathbf{1}_{\{x_i\in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \int_{B(y,r)} \frac{dz}{V_r} \langle z-x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\le 1\}},$$

and if we decompose the range $(0,\infty)$ over which we integrate r in the third term of $\mathcal{G}^\alpha f(A)$ into $(0,\delta(A))$ and $[\delta(A),\infty)$, we find that the integral over the latter is equal to the quantity above.

Finally, let us show that (28) converges to the second term of $\mathcal{G}^\alpha f(A)$. This time, we split (28) into

$$\begin{aligned} u\sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^{\infty} \frac{\mathbf{1}_{\{x_i\in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \int_{B(y,r)} \frac{dz}{V_r} \big(f(\Phi_{\{i\}}(A,z)) - f(A)\big) \mathbf{1}_{\{|z-x_i|>1\}} \\ + u\sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_{n^{-1/\alpha}}^{\infty} \frac{\mathbf{1}_{\{x_i\in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \\ \times \int_{B(y,r)} \frac{dz}{V_r} \big(f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z-x_i, \nabla_i f(A)\rangle\big) \mathbf{1}_{\{|z-x_i|\le 1\}}. \end{aligned}$$

The first term is finite for the same reasons as in (27), since for the parent to be at distance greater than 1 from the affected lineage, one must have $r>1/2$. Now, using the same steps as above, we obtain that the second term is equal to

$$u\sum_{i=1}^{|A|} \int_{B(x_i,1)} dz \int_{n^{-1/\alpha}\vee\frac{|z-x_i|}{2}}^{\infty} \frac{dr}{V_r r^{\alpha+d+1}} \int_{B(z,r)\cap B(x_i,r)} dy\ (1-u)^{|J(y,r,A)|-1} \times \big(f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z-x_i, \nabla_i f(A)\rangle\big). \tag{30}$$

But f is of class C^2 and has compact support, and so we can find a constant $\tilde{C}(f) > 0$, independent of A, such that for every i and every $z \in B(x_i, 1)$,

$$\big|f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z - x_i, \nabla_i f(A)\rangle\big| \leq \tilde{C}(f)|z - x_i|^2.$$

As a consequence, the absolute value of the quantity in (30) is bounded by

$$\begin{aligned}
&u\tilde{C}\sum_{i=1}^{|A|}\int_{B(x_i,1)} dz \int_{n^{-1/\alpha}\vee\frac{|z-x_i|}{2}}^{\infty} \frac{dr}{r^{\alpha+d+1}} \frac{\mathrm{Vol}\big(B(z,r)\cap B(x_i,r)\big)}{V_r}\,|z-x_i|^2 \\
&\leq uC'|A|\int_{B(0,1)} dz\ |z|^2\big(n^{-1/\alpha}\vee(|z|/2)\big)^{-\alpha-d} \\
&= uC'|A|\left\{n^{1+\frac{d}{\alpha}}\int_{B(0,2n^{-1/\alpha})} dz\ |z|^2 + 2^{\alpha+d}\int_{B(0,1)\setminus B(0,2n^{-1/\alpha})} dz\ |z|^{2-\alpha-d}\right\} \\
&\leq C''|A|\big\{n^{-\frac{2-\alpha}{\alpha}} + C'''\big(1 - n^{-\frac{2-\alpha}{\alpha}}\big)\big\},
\end{aligned} \tag{31}$$

where all the constants appearing in this bound depend on f, d and α, but not on A. Since $\alpha < 2$, (30) remains bounded as $n \to \infty$ and (28) converges to

$$\begin{aligned}
u\sum_{i=1}^{|A|}\int_{\mathbb{R}^d} dy \int_0^{\infty} \frac{\mathbf{1}_{\{x_i\in B(y,r)\}}dr}{r^{\alpha+d+1}}(1-u)^{|J(y,r,A)|-1}\int_{B(y,r)}\frac{dz}{V_r}\big[f(\Phi_{\{i\}}(A,z)) - f(A) \\
- \langle z - x_i, \nabla_i f(A)\rangle \mathbf{1}_{\{|z-x_i|\leq 1\}}\big],
\end{aligned}$$

which is precisely the second term of $\mathcal{G}^\alpha f(A)$ (and is finite according to the analysis above). Tracing back our calculations, we see that for n large enough (such that $n^{-1/\alpha} < \delta(A)$) the difference between $\mathcal{G}^n f(A)$ and $\mathcal{G}^\alpha f(A)$ is equal to the difference between the quantity in (30) and its counterpart in $\mathcal{G}^\alpha f(A)$ (that is, the second term of $\mathcal{G}^\alpha f(A)$ in which y is only integrated over $B(x_i, 1)$). Hence, according to (31), for every $n > \delta(A)^{-\alpha}$

$$\big|\mathcal{G}^n f(A) - \mathcal{G}^\alpha f(A)\big| \leq c_f\,|A|\,n^{-\frac{2-\alpha}{\alpha}},$$

where the constant c_f is again independent of A. Consequently, for every f which is compactly supported and of class C^2 with respect to the marks, the function $\mathcal{G}^\alpha f$ is bounded and the convergence

$$\lim_{n\to\infty}\sup_{\delta(A)>\varepsilon,|A|\leq k}\big|\mathcal{G}^n f(A) - \mathcal{G}^\alpha f(A)\big| = 0 \tag{32}$$

holds for any choice of $\varepsilon > 0$ and $k \in \mathbb{N}$.

To conclude the proof of Proposition 5.1, let us use the following result, whose proof we postpone for the sake of clarity. For every $\varepsilon > 0$, let t_ε be the first time at which at least two lineages lie at distance less than $\varepsilon > 0$ without having coalesced.

Lemma 5.2. *For every initial value A_0 such that $\delta(A_0) > 0$, we have*

$$\lim_{\varepsilon\to 0}\mathbb{P}_{A_0}[t_\varepsilon < \infty] = 0. \tag{33}$$

As a consequence, the martingale problem associated to $(\mathcal{G}^\alpha, A_0)$ has a unique solution (with càdlàg paths) for any initial value A_0 satisfying $\delta(A_0) > 0$. Let us denote this solution by $\mathcal{A}^\infty$. Then $\mathcal{A}^\infty$ is a consistent system of coalescing symmetric α-stable processes.

Let us suppose that Lemma 5.2 has been established, and verify that the conditions of Theorem 4.8.2(b) of [EK86] are then fulfilled. First, one can check that the set of functions f considered above is dense in the set of all bounded continuous functions on marked partitions. We can thus restrict our attention to these particular functions. Second, (33) enables us to use (32) and dominated convergence to obtain that Condition (8.7) of Theorem 4.8.2(b) of [EK86] is satisfied, and consequently that the finite-dimensional distributions of $\mathcal{A}^n$ converge weakly to those of $\mathcal{A}^\infty$ as n tends to infinity. The arguments for the convergence of the one-dimensional distributions of $\hat{\mathcal{A}}^n$ are the same as in the case with fixed radii, and so the proof of Proposition 5.1 is now complete. □

Before proving Lemma 5.2, let us study some of properties of the 'genealogical' process $\mathcal{A}^\infty$. Indeed, in order to use Lemma 3.2(i), we need to show that (10) holds. In fact we can be more precise about the way coalescence occurs.

Lemma 5.3. *Sample two individuals at separation x, and consider their ancestral lineages $(X_t, t \geq 0)$, $(Y_t, t \geq 0)$. Let*

$$\tau = \inf\{t \geq 0 : X_s = Y_s \text{ for all } s \geq t\}$$

be their coalescence time. Then $\tau < \infty$ almost surely, and moreover, there exists a random variable Z, a.s. finite and independent of x, such that

$$\tau \preceq x^\alpha Z, \tag{34}$$

where $\preceq$ stands for stochastic domination.

Proof of Lemma 5.3. In essence, the strategy of the proof consists of showing that if the two lineages start at distance $a > 0$, they have some positive chance (independent of a) of coalescing before they either separate to a distance greater than $2a$ or come within distance less than $a/2$ of each other. The dependence on x^α in the lemma then comes from the fact that the time needed to coalesce, or separate, or get closer by a factor of 2, is of the order of x^α when the initial separation is x.

By translation invariance, we may assume without loss of generality that the origin of $\mathbb{R}^d$ sits at the midpoint between X_0 and Y_0. Let $T(x)$ be the first time that *any* point in $B := B(0, x)$ is touched by an event whose radius r is greater than $x/4$. Then we claim that $T(x)$ is an exponential random variable whose rate $\lambda(x)$ is given for every $x > 0$ by

$$\lambda(x) = \int_{x/4}^\infty \frac{d\ell}{\ell^{d+1+\alpha}} \operatorname{Vol}(B(0, x+\ell)). \tag{35}$$

Indeed, recall the intensity measure (3) we introduced before rescaling the process. In the original units of time and space, the rate at which any point of the closed ball $B(0, x)$ ($x \geq 4$) is hit by an event of radius greater than $x/4$ is given by

$$\int_{\mathbb{R}^d} dz \int_{x/4}^\infty \frac{d\ell}{\ell^{d+1+\alpha}} \mathbf{1}_{\{B(0,x)\cap B(z,\ell)\neq\emptyset\}} = \int_{x/4}^\infty \frac{d\ell}{\ell^{d+1+\alpha}} \operatorname{Vol}(B(0, x+\ell)).$$

Multiplying this rate by n and looking at distances of the form $xn^{1/\alpha}$, a simple change of variables gives us that for every $x \geq 4n^{-1/\alpha}$, the rescaled rate of interest is also equal to the expression above, independently of n. Passing to the limit $n \to \infty$ yields (35).

Now, setting $\ell = rx$ we can write

$$\begin{aligned}\lambda(x) &= x^{-d-\alpha} \int_{1/4}^{\infty} \frac{dr}{r^{d+1+\alpha}} \operatorname{Vol}(B(0, x + rx)) \\ &= x^{-\alpha} \int_{1/4}^{\infty} \frac{dr}{r^{d+1+\alpha}} \operatorname{Vol}(B(0, 1 + r)) = Cx^{-\alpha}, \end{aligned} \tag{36}$$

where the constant C is independent of x.

On the other hand, similar calculations enable us to see that the rate at which B is entirely contained within the area $B(z, r)$ of an event is given by

$$\begin{aligned}\int_{\mathbb{R}^d} dz \int_{|z|+x}^{\infty} \frac{d\ell}{\ell^{d+1+\alpha}} &= \int_{x}^{\infty} \frac{d\ell}{\ell^{d+1+\alpha}} \int_{\mathbb{R}^d} dz\, \mathbf{1}_{\{|z| \le \ell - x\}} \\ &= x^{-\alpha} \int_{1}^{\infty} \frac{dr}{r^{d+1+\alpha}} \operatorname{Vol}(B(0, r-1)) = C' x^{-\alpha},\end{aligned}$$

where we used the same change of variable as before and $C' > 0$ is again independent of x. As a consequence, with probability $p_0 := C'/C$ independent of x, the first event of radius greater than $x/4$ that hits at least one point of B actually covers the whole ball. Moreover, (36) also implies that for arbitrary $q \ge 1/4$, the radius $R(x)$ of the event occurring at time $T(x)$ satisfies

$$\mathbb{P}(R(x) > qx) \le cq^{-\alpha}, \tag{37}$$

for some constant c which does not depend on x or q.

Let $\tilde{X}, \tilde{Y}$ be the motion of the lineages as governed by all the events except those that affect some point in B and whose radius is greater than $x/4$. Then by the Poisson point process formulation of the reproduction events, $T(x)$ is independent of $\tilde{X}, \tilde{Y}$ and $(X_t, Y_t, t < T(x))$ coincides with $(\tilde{X}_t, \tilde{Y}_t, t < T(x))$. Let $S(x) := \inf\{t \ge 0 : \tilde{D}_t \le x/2 \text{ or } \tilde{X}_t \notin B \text{ or } \tilde{Y}_t \notin B\}$, where $D_t = |\tilde{X}_t - \tilde{Y}_t|$. Fix $\delta > 0$, and define the following events:

$$E := \{T(x) \le \delta x^{\alpha}\}, \qquad F := \{S(x) \ge \delta x^{\alpha}\}.$$

Then E and F are independent, and by (36) there exists $p(\delta) > 0$ such that $\mathbb{P}(E) = p(\delta)$ for all $x > 0$. A similar property holds for F. Indeed, note first that up until the time $S(x)$, the trajectories $\tilde{X}$ and $\tilde{Y}$ are independent, since the trajectories can only move as a result of events occurring in necessarily disjoint regions of space. Moreover, it is easy to check that

$$\Big(\frac{1}{x}\, \tilde{X}_{tx^{\alpha} \wedge S(x)},\ \frac{1}{x}\, \tilde{Y}_{tx^{\alpha} \wedge S(x)}\Big)_{t \ge 0} \tag{38}$$

has the same distribution as the pair $(\tilde{X}_{t \wedge S(1)}, \tilde{Y}_{t \wedge S(1)})_{t \ge 0}$ obtained by taking $x = 1$: both coordinates of this process perform independent stable Lévy processes where each jump greater than $1/4$ occurring in $B(0, 1)$ is removed, and the process is stopped when either coordinate leaves $B(0, 1)$ or they come within distance $1/2$ of one another. (A formal proof is given by comparing the generators; the generator of the pair $(\tilde{X}, \tilde{Y})$ is the same as $\mathcal{G}^{\alpha} f(A)$ when A has two blocks, as defined in (26), but with the first term equal to 0).

Hence for all $x > 0$, $\mathbb{P}(S(x) \ge \delta x^{\alpha}) = \mathbb{P}(S(1) \ge \delta) =: q(\delta)$, and $q(\delta) > 0$ whenever δ is chosen small enough.

Let us denote the centre, radius and impact parameter of the event taking place at time $T(x)$ by $(Z(x), R(x), u)$. We shall say that a *success* occurs if both E and F occur, and if

(a) $B(0,x) \subset B(Z(x), R(x))$,

(b) both $X_{T(x)}, Y_{T(x)}$ are affected by the event occurring at time $T(x)$ (this is possible since under these assumptions, $X_{T(x)}$ and $Y_{T(x)}$ are still both in $B(0,x)$ which is entirely covered by the event.)

Note that by the above discussion,

$$\wp := \mathbb{P}(\text{ success }) = p(\delta)q(\delta)p_0u^2, \tag{39}$$

independently of $x > 0$.

If a success did not occur, we say that a *failure* has occurred. Since the success probability is independent of x and the waiting time between two attempts is always stochastically bounded by an exponential random variable of the form $T(y)$ (which is a.s. finite), we deduce that after a Geometric($\wp$) number N of attempts, success is guaranteed, hence $\tau < \infty$ almost surely. Moreover, in the case of failure, consider the mutual distance $D_{T(x)\wedge S(x)}$ between the two lineages at time $T(x) \wedge S(x)$. Then $D_{T(x)\wedge S(x)} \leq 2x + R(x)$. From (37) we can deduce that there exists a random variable R, independent of x and a.s. finite, such that $2 + R(x)/x \preceq R$ in the sense of stochastic domination. Let $R_1, R_2, \ldots$ be a sequence of i.i.d. random variables with distribution R. The strong Markov property and (36) then show that

$$\tau \preceq x^\alpha \Big\{ \mathcal{E}[C] + \mathcal{E}\big[CR_1^{-\alpha}\big] + \ldots + \mathcal{E}\big[C(R_1 \cdots R_N)^{-\alpha}\big] \Big\},$$

where $\mathcal{E}[y]$ stands for an exponential random variable with parameter y and all the above exponential random variables are conditionally independent given their arguments. Define Z as the random variable within the curly brackets to conclude. □

Remark 5.4. The system $\mathcal{A}^\infty$ inherits the consistency property from its construction as the limit of $\mathcal{A}^n$ (this property can also be shown directly from the generator of $\mathcal{A}^\infty$). Hence, a notable consequence of Lemma 5.3 is that any finite sample of lineages finds its most recent common ancestor in finite time with probability one. The same kind of behaviour, as well as the convergence of the forwards-in-time process to a field of correlated Bernoulli random variables, was already observed by Evans in the case where the genealogical process of his *continuous sites stepping-stone model* is a system of one-dimensional independent α-stable motions coalescing instantly upon meeting. See §5 in [Eva97] for a full description of his results. However, the underlying mechanisms are quite different here. Not only does Lemma 5.3 hold for any $\alpha \in (1,2)$ and any dimension, which cannot be the case in Evans' framework since two independent stable processes may not meet, but even in dimension 1 the way lineages coalesce is different: the limit in (33) shows that two lineages of $\mathcal{A}^\infty$ have no chance to meet, but their coalescence is due to the fact that large events of the appropriate size are just frequent enough to catch them even when they are very far from each other. As a last consequence, it is then possible to see multiple mergers during the evolution of $\mathcal{A}^\infty$, which is not the case when the α-stable processes move independently of each other and coalesce only when they meet.

Let us now finish with the proof of Lemma 5.2 and of Theorem 1.5. Recall that for any marked partition A, $\delta(A)$ stands for half the minimum distance between two marks in A ($\delta(A) = +\infty$ if A has only one block).

Proof of Lemma 5.2. Because most of the ideas and computations we shall use to establish (33) are developed in detail in the proof of Lemma 5.3, we only present an outline here and refer to that proof for more precise arguments. Since we always deal with partitions of some

finite set, it is sufficient to show the result when A_0 consists of just two blocks starting at some positive separation.

If $x > 0$ denotes the initial distance between our two lineages, let us call $T(x)$ the first time at which any of the lineages is in the geographical area of an event of radius greater than $x/4$, and let us call $S(x)$ the first time at which the distance between the two lineages is greater than $2x$, or less than $x/2$. Notice that the lineages evolve independently until the random time $T(x) \wedge S(x)$, since they are hit by events that are necessarily disjoint until that time. Moreover, they both move according to the law of a symmetric α-stable process whose large jumps have been truncated (see reasoning below (38)). Hence it is not difficult to show that $S(x)$ is of the order of x^α, and so is $T(x)$, while the coalescence rate of two lineages at distance x is commensurate with $x^{-\alpha}$. Using the more careful analysis performed in the proof of Lemma 5.3, we can in fact conclude that the probability p_0 that the two lineages coalesce before their distance doubles or is divided by two is not only positive, but also independent of x. Together with the fact that $T(y) \wedge S(y)$ is a.s. finite for every $y > 0$ (for reasons expounded in Lemma 5.3), the number of attempts before succeeding to coalesce is a geometric random variable with parameter p_0, which we shall denote by N.

As a second step, suppose that the lineages fail to coalesce at time $T(x) \wedge S(x)$. The new location of the lineage which jumps at that time (at most one of them jumps, otherwise they would coalesce) is uniformly distributed over the area of the event, and since the lineages are at distance at least $x/2$ from each other just before $T(x) \wedge S(x)$ a small calculation using the scaling properties of the evolution mechanism shows that the probability $\pi(\eta)$ that their new distance at that time is less than ηx satisfies

(a) $\pi(\eta)$ is independent of x,

(b) $\lim_{\eta \to 0} \pi(\eta) = 0$.

As a consequence, if $\eta \in (0, 1/10)$ and $k \in \mathbb{N}$, we can write

$$\mathbb{P}_{A_0}\big[\text{coal. before distance decreases by } \eta^k\big] \geq \mathbb{E}\big[(1-\pi(\eta))^{N-1}\mathbf{1}_{\{N<k\}}\big]. \tag{40}$$

Note in passing that, by monotonicity, the same inequality holds if we replace η^k by any $\varepsilon \leq \eta^k$.

Let us now draw some conclusions from these observations. We fix $c > 0$, and choose $k(c)$ and $\eta(c)$ such that for every $k \geq k(c)$ and $\eta \leq \eta(c)$,

$$\mathbb{P}[N \geq k] \leq \frac{c}{2} \qquad \text{and} \qquad \mathbb{E}\big[(1-\pi(\eta))^{N-1}\big] \geq 1 - \frac{c}{2}.$$

Then, using the fact that the event described in the left-hand side of (40) implies $t_\varepsilon = +\infty$ for every $\varepsilon \leq \eta^k x$, we have that, for every such ε,

$$\mathbb{P}_{A_0}[t_\varepsilon = \infty] \geq 1 - \frac{c}{2} - \frac{c}{2} = 1 - c.$$

Since c was arbitrary, (33) follows.

As regards the second part of Lemma 5.2, recall from (26) that the operator $\mathcal{G}^\alpha$ is defined, for every function f of class C^2 with compact support and every marked partition A satisfying

$\delta(A) > 0$, by

$$\mathcal{G}^\alpha f(A) = \int_{\mathbb{R}^d} dy \int_{\delta(A)}^\infty \frac{dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \sum_{I \subset J(y,r,A), |I| \geq 2} u^{|I|}(1-u)^{|J \setminus I|} \big[f(\Phi_I(A,z)) - f(A) \big]$$

$$+ u \sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_i \in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1}$$

$$\times \int_{B(y,r)} \frac{dz}{V_r} \big[f(\Phi_{\{i\}}(A,z)) - f(A) - \langle z - x_i, \nabla_i f(A) \rangle \mathbf{1}_{\{|z-x_i| \leq 1\}} \big]$$

$$+ u \sum_{i=1}^{|A|} \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_i \in B(y,r)\}} dr}{r^{\alpha+d+1}} (1-u)^{|J(y,r,A)|-1} \int_{B(y,r)} \frac{dz}{V_r} \langle z - x_i, \nabla_i f(A) \rangle \mathbf{1}_{\{|z-x_i| \leq 1\}}.$$

In particular, if $A = \{(b_1, x_1)\}$ contains only one block and if f is a function of its mark only, then $\mathcal{G}^\alpha f(A)$ is equal to

$$u \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_1 \in B(y,r)\}} dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \big[f(z) - f(x_1) - \langle z - x_1, \nabla f(x_1) \rangle \mathbf{1}_{\{|z-x_1| \leq 1\}} \big]$$

$$+ u \int_{\mathbb{R}^d} dy \int_0^\infty \frac{\mathbf{1}_{\{x_1 \in B(y,r)\}} dr}{r^{\alpha+d+1}} \int_{B(y,r)} \frac{dz}{V_r} \langle z - x_1, \nabla f(x_1) \rangle \mathbf{1}_{\{|z-x_1| \leq 1\}}$$

$$= u \int_{\mathbb{R}^d} dz \left(\int_0^\infty \frac{dr}{r^{\alpha+d+1}} \frac{\mathrm{Vol}\big(B(z,r) \cap B(x_1,r)\big)}{V_r} \right) \big[f(z) - f(x_1) - \langle z - x_1, \nabla f(x_1) \rangle \mathbf{1}_{\{|z-x_1| \leq 1\}} \big]$$

$$+ u \int_{\mathbb{R}^d} dz \left(\int_0^\infty \frac{dr}{r^{\alpha+d+1}} \frac{\mathrm{Vol}\big(B(z,r) \cap B(x_1,r)\big)}{V_r} \right) \langle z - x_1, \nabla f(x_1) \rangle \mathbf{1}_{\{|z-x_1| \leq 1\}}.$$

The second term above is zero (by symmetry) and hence the generator can be written as

$$\mathcal{G}^\alpha f(A) = \int_{\mathbb{R}^d} dz \; \iota(z - x_1) \; \big(f(z) - f(x_1) \big), \tag{41}$$

where the intensity $\iota(w)$ is given by

$$\iota(w) := u \int_0^\infty \frac{dr}{r^{\alpha+d+1}} \frac{\mathrm{Vol}\big(B(w,r) \cap B(0,r)\big)}{V_r}.$$

Now, one can check that for any $k > 0$

$$k \, \iota(zk^{-1/\alpha}) \, d(zk^{-1/\alpha}) = \iota(z) \, dz$$

and so the motion of a single lineage is a symmetric α-stable Lévy process.

When there are at least two blocks, as long as $\delta(\mathcal{A}_t^\infty) > 0$ the first term of $\mathcal{G}^\alpha f(\mathcal{A}_t^\infty)$ is finite and clearly represents the merger and jump at finite rate of several blocks of $\mathcal{A}^\infty$. However, the coalescence rate of two lineages at distance ε is equal to

$$u^2 \int_{\mathbb{R}^d} dy \int_{\varepsilon/2}^\infty \frac{dr}{r^{\alpha+d+1}} \mathbf{1}_{\{x_1, x_2 \in B(y,r)\}} = u^2 \int_{\varepsilon/2}^\infty \frac{dr}{r^{\alpha+d+1}} \mathrm{Vol}\big(B(x_1,r) \cap B(x_2,r)\big) \propto \varepsilon^{-\alpha}$$

as $\varepsilon \to 0$, and so one can prove the existence of the process $\mathcal{A}^\infty$ only up to t_ε, for any $\varepsilon > 0$. Yet (33) is actually more than what is required to invoke Theorem 4.6.3 in [EK86] and complete the proof of existence of $\mathcal{A}^\infty$. □

Proof of Theorem 1.5. There is nothing else to do. Duality and the convergence of $\mathcal{A}^n$ give us the convergence of ρ^n exactly as in the proof of Theorem 1.1. Lemma 5.3 is sufficient to show that (10) holds and so the limiting densities $w^{(\alpha)}(t,x)$ are Bernoulli random variables as stated. $\square$

References

[BEV10] N.H. Barton, A.M. Etheridge and A. Véber (2010). A new model for evolution in a spatial continuum. *Electron. J. Probab.*, 15:162–216.

[BKE10] N. H. Barton, J. Kelleher and A. M. Etheridge (2010). A new model for extinction and recolonization in two dimensions : quantifying phylogeography. *Evolution*, 64:2701–2715.

[BLG03] Bertoin, J. and Le Gall, J.-F. (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields*, 126:261–288.

[Bil95] P. Billingsley (1995). *Probability and Measure*. Wiley.

[DK99] P.J. Donnelly and T.G. Kurtz (1999). Particle representations for measure-valued population models. *Ann. Probab.*, 27:166–205.

[Eth08] A. M. Etheridge (2008). Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.*, 80:121–144.

[Eth11] A.M. Etheridge (2011). *Some mathematical models from population genetics*. Ecole d'été de probabilités de Saint-Flour 2009, Springer.

[EV11] A.M. Etheridge and A. Véber (2011). The spatial Lambda-Fleming-Viot process on a large torus: genealogies in the presence of recombination. arXiv:1106.4050.

[EK86] S.N. Ethier and T.G. Kurtz (1986). *Markov processes: characterization and convergence*. Wiley.

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

[Fel75] J. Felsenstein (1975). A pain in the torus: some difficulties with the model of isolation by distance. *Amer. Nat.*, 109:359–368.

[Kim53] M. Kimura (1953). Stepping stone model of population. *Ann. Rep. Nat. Inst. Genetics Japan*, 3:62–63.

[Pit99] J. Pitman (1999). Coalescents with multiple collisions. *Ann. Probab.*, 27:1870–1902.

[Saa11] H. Saadi. Λ-Fleming-Viot processes and their spatial extensions. *DPhil Thesis, University of Oxford.*

[Sag99] S. Sagitov (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 26:1116–1125.

[VW11] A. Véber and A. Wakolbinger. A flow representation for the spatial Λ-Fleming-Viot process. *In preparation.*