The Annals of Applied Probability
2012, Vol. 22, No. 6, 2165–2209
DOI: 10.1214/12-AAP842

THE SPATIAL Λ-FLEMING–VIOT PROCESS ON A LARGE TORUS: GENEALOGIES IN THE PRESENCE OF RECOMBINATION

BY A. M. ETHERIDGE[1] AND A. VÉBER[2]

University of Oxford and CMAP—École Polytechnique

We extend the spatial Λ-Fleming–Viot process introduced in [*Electron. J. Probab.* **15** (2010) 162–216] to incorporate recombination. The process models allele frequencies in a population which is distributed over the two-dimensional torus $\mathbb{T}(L)$ of sidelength L and is subject to two kinds of reproduction events: *small events* of radius $\mathcal{O}(1)$ and much rarer *large events* of radius $\mathcal{O}(L^{\alpha})$ for some $\alpha \in (0, 1]$. We investigate the correlation between the times to the most recent common ancestor of alleles at two linked loci for a sample of size two from the population. These individuals are initially sampled from "far apart" on the torus. As L tends to infinity, depending on the frequency of the large events, the recombination rate and the initial distance between the two individuals sampled, we obtain either a complete decorrelation of the coalescence times at the two loci, or a sharp transition between a first period of complete correlation and a subsequent period during which the remaining times needed to reach the most recent common ancestor at each locus are independent. We use our computations to derive approximate probabilities of identity by descent as a function of the separation at which the two individuals are sampled.

1. Introduction.

1.1. *Background*. In the 30 years since its introduction, Kingman's coalescent has become a fundamental tool in population genetics. It provides an elegant description of the genealogical trees relating individuals in a sample from a highly idealized biological population, in which it is assumed that all individuals are selectively neutral and experience identical conditions, and that population size is constant. Spurred on by the flood of DNA sequence data, theoreticians have successfully extended the classical coalescent to incorporate more realistic biological assumptions such as varying population size, natural selection and genetic structure. However, it has proved surprisingly difficult to produce satisfac-

Received June 2011; revised January 2012.

[1]Supported in part by EPSRC Grant EP/E065945/1.

[2]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-09-BLAN-0215).

MSC2010 subject classifications. Primary 60J25, 92D10, 60J75; secondary 60F05.

Key words and phrases. Genealogy, recombination, coalescent, spatial continuum, generalized Fleming–Viot process.

tory extensions for populations living (as many do) in continuous two-dimensional habitats—a problem dubbed *the pain in the torus* by Felsenstein [9].

In the classical models of population genetics, it is customary to assume that populations are either panmictic, meaning, in particular, that they have no spatial structure, or that they are subdivided into "demes." The demes sit at the vertices of a graph which is chosen to caricature the geographic region in which the population resides. Thus, for example, for a population living in a two-dimensional spatial continuum one typically takes the graph to be (a subset of) $\mathbb{Z}^2$. Reproduction takes place within demes and interaction between the subpopulations is through migration along the edges of the graph. Models of this type are collectively known as stepping stone models.

However, in order to apply a stepping stone model to populations that are distributed across continuous space, one is forced to make an artificial subdivision. Moreover, the predictions of stepping stone models fail to match observed patterns of genetic variation. For example, they overestimate genetic diversity (often by many orders of magnitude) and they fail to predict the long-range correlations in allele frequencies seen in real populations.

In recent work [1, 2, 8] we introduced a new framework in which to model populations evolving in a spatial continuum. The key idea, which enables us to overcome the pain in the torus, is that reproduction is driven by a Poisson process of events which are based on geographical space rather than on individuals. This leads, in particular, to a class of models that could reasonably be called *continuum stepping stone models*, but it also allows one to incorporate large-scale extinction/recolonization events. Such events dominate the demographic history of many species. They appear in our framework as "local population bottlenecks." In [2], we show (numerically) how the inclusion of such events can lead to long-range correlations in allele frequencies. In [1] a rigorous mathematical analysis of a class of models on a torus in $\mathbb{R}^2$ illustrates the reduction in genetic diversity that can result from such large-scale demographic events. We expand further on this in Section 2. Thus, large-scale events provide one plausible explanation of the two deficiencies of stepping stone models highlighted above, but of course they are not the only possible explanation.

A natural question now arises: how could we infer the existence of these large-scale events from data? One possible answer is through correlations in patterns of variation at different genetic loci. Recall that in a diploid population (in which chromosomes are carried in pairs) correlations between linked genes (i.e., genes occurring on the same chromosome) are broken down over time by recombination (which results in two genes on the same chromosome being inherited from different chromosomes in the parent). We say that genes are *loosely* linked, if the rate of recombination events is high [e.g., if the chance of a recombination in a single generation is $\mathcal{O}(1)$]. In the Kingman coalescent, genealogies relating loosely linked genes evolve independently. This is because on the timescale of the coalescent, the states in which lineages ancestral to both loci are in the same individual

vanish instantaneously. It is well known that if a population experiences a bottleneck, this is no longer the case. As we trace backward in time, when we reach the bottleneck, we expect to see a significant proportion of surviving lineages coalesce *at the same time* and so we see correlations in genealogies even at *unlinked* loci. With local bottlenecks we can expect a rather more complicated picture. The degree of correlations across loci will depend upon the spatial separation of individuals in the sample.

The purpose of this paper is to extend the model of [1] to diploid populations, to incorporate recombination, and to provide a first rigorous analysis of the correlations in genealogies at different loci in the presence of local extinction/recolonization events. Since the questions we shall address and some of the methods we shall use here are related to those of [1], the reader may find it useful to have some familiarity with the results of that paper.

1.2. *The model.* In [1], we introduced the *spatial* Λ*-Fleming–Viot process* as a model of a haploid population evolving in a spatial continuum. It is a Markov process taking its values in the set of functions which associate to each point of the geographical space a probability measure on a compact space, K, of genetic types. If Φ is the current state of the population and x is a spatial location, the measure $\Phi(x)$ can be interpreted as the distribution of the type of an individual sampled from location x. The dynamics are driven by a Poisson point process of *events*. An event specifies a spatial region, A, say, and a number $u \in (0, 1]$. As a result of the event, a proportion u of individuals within A are replaced by offspring of a parent sampled from a point picked uniformly at random from A. In [1] the regions A are chosen to be discs of random radius (whose centers fall with intensity proportional to the Lebesgue measure) and the distribution of u can depend on the radius of the disc. Under appropriate conditions, existence and uniqueness in law of the process were established.

Here we wish to extend this framework in a number of directions. First, whereas in [1] a single parent was chosen from the region A, here we allow A to be repopulated by the offspring of a finite (random) number of its inhabitants. Second, we assume that the population is diploid. We shall follow (neutral) genes at two distinct (linked) loci, with recombination acting between them. Writing K_1 and K_2 for the possible types at the two loci, the type of an individual is an element of $K_1 \times K_2$ (which we can identify with $[0, 1] \times [0, 1]$). As in [1], we work on the torus $\mathbb{T}(L)$ of side L in $\mathbb{R}^2$ and we suppose that there are two types of events: small events, affecting regions of radius $\mathcal{O}(1)$, which might be thought of as "ordinary" reproduction events; and "large" events, representing extinction/recolonization events, affecting regions of radius $\mathcal{O}(L^\alpha)$ where $\alpha \in (0, 1]$ is a fixed parameter. In order to keep the notation as simple as possible, we shall only allow two different radii for our events, R_s corresponding to "small" reproduction events and $R_B L^\alpha$ corresponding to "large" local bottlenecks. We shall also suppose that the corresponding proportions u_s and u_B are fixed. Neither of these assumptions is essential to the

results, which would carry over to the more general setting in which each of R_s, R_B, u_s and u_B is sampled (independently) from given distributions each time an event occurs.

Let us specify the dynamics of the process more precisely. Let:

- $R_s, R_B \in (0, \infty)$, $u_s, u_B \in (0, 1)$ and $\alpha \in (0, 1]$;
- λ_s, λ_B be two distributions on $\mathbb{N} = \{1, 2, \ldots\}$ with bounded support and such that $\lambda_s(\{1\}) < 1$;
- $(\rho_L)_{L\in\mathbb{N}}$ be an increasing sequence such that $\rho_L \geq \log L$ for all $L \in \mathbb{N}$, and $L^{-2\alpha}\rho_L$ tends to a finite limit (possibly zero) as $L \to \infty$; and
- $(r_L)_{L\in\mathbb{N}}$ be a nonincreasing sequence with values in $(0, 1]$.

For $L \in \mathbb{N}$, we denote by Π_s^L a Poisson point process on $\mathbb{R} \times \mathbb{T}(L)$ with intensity measure $dt \otimes dx$, and by Π_B^L another Poisson point process on $\mathbb{R} \times \mathbb{T}(L)$, independent of Π_s^L, with intensity measure $(\rho_L L^{2\alpha})^{-1} dt \otimes dx$. The spatial Λ-Fleming–Viot process Φ^L on $\mathbb{T}(L)$ evolves as follows.

Small events: If (t, x) is a point of Π_s^L, a reproduction event takes place at time t within the closed ball $B(x, R_s)$:

- a number j is sampled according to the measure λ_s;
- j sites, $z_1, \ldots, z_j$ are selected uniformly at random from $B(x, R_s)$; and,
- for each $i = 1, \ldots, j$, a type (a_i, b_i) is sampled according to $\Phi_{t-}^L(z_i)$.

If $j > 1$, then for all $y \in B(x, R_s)$,

$$\Phi_t^L(y) := (1 - u_s)\Phi_{t-}^L(y) + \frac{u_s(1 - r_L)}{j} \sum_{i=1}^{j} \delta_{(a_i,b_i)} + \frac{u_s r_L}{j(j-1)} \sum_{i_1 \neq i_2} \delta_{(a_{i_1},b_{i_2})}.$$

If $j = 1$, for each $y \in B(x, R_s)$,

$$\Phi_t^L(y) := (1 - u_s)\Phi_{t-}^L(y) + u_s \delta_{(a_1,b_1)}.$$

In both cases, sites outside $B(x, R_s)$ are not affected.

Large events: If (t, x) is a point of Π_B^L, an extinction/recolonization event takes place at time t within the closed ball $B(x, L^\alpha R_B)$:

- a number j is sampled according to the measure λ_B;
- j sites, $z_1, \ldots, z_j$ are selected uniformly at random from $B(x, L^\alpha R_B)$; and
- for each $i = 1, \ldots, j$, a type (a_i, b_i) is sampled according to $\Phi_{t-}^L(z_i)$.

For each $y \in B(x, L^\alpha R_B)$,

$$\Phi_t^L(y) := (1 - u_B)\Phi_{t-}^L(y) + \frac{u_B}{j} \sum_{k=1}^{j} \delta_{(a_k,b_k)}.$$

Again, sites outside the ball are not affected.

REMARK 1.1. (1) The scheme of choosing j parental locations and then sampling a parental type at each of those locations is convenient when one is interested in tracing lineages ancestral to a sample from the population. It can be thought of as sampling j individuals, uniformly at random from the ball affected by the reproduction (or extinction/recolonization) event, to reproduce. Of course this scheme allows for the possibility of more than one parent contributing offspring so that we should more correctly call this model a spatial Ξ-Fleming–Viot process, but to emphasize the close link with previous work we shall abuse terminology and use the name Λ-Fleming–Viot process.

(2) The recombination scheme mirrors that generally employed in Moran models. The quantity r_L is the proportion of offspring who, as a result of recombination, inherit the types at the two loci from different parental chromosomes. We have chosen to sample the types of those two chromosomes from different points in space. The result of this is that provided the individuals sampled from the current population are in distinct geographic locations, if two ancestral lineages are at spatial distance zero, then they are necessarily in the same individual. This is mathematically convenient (cf. Remark 3.2) but, arguably, not terribly natural biologically. However, changing the sampling scheme, for example, so that the two recombining chromosomes are sampled from the same location, would not materially change our results.

(3) We are assuming that recolonization is so rapid after an extinction event that the effects of recombination during recolonization are negligible.

(4) Since $\mathbb{T}(L)$ is compact, the overall rate at which events fall is finite for any L and the corresponding *spatial Λ-Fleming–Viot process with recombination* is well-defined. Notice that a given site, x, say, is affected by a small event at rate $\pi R_s^2 = \mathcal{O}(1)$ (since the center of the event must fall within a distance R_s of x), whereas it is hit by a large event at rate $\pi R_B^2 \rho_L^{-1} = \mathcal{O}(\rho_L^{-1})$. So reproduction events are frequent, but massive extinction/recolonization events are rare.

1.3. *Genealogical relationships*. Having established the (forward in time) dynamics of allele frequencies in our model, we now turn to the genealogical relationships between individuals in a sample from the population.

First suppose that we are tracing the lineage ancestral to a single locus on a chromosome carried by just one individual in the current population. Recombination does not affect us and we see that the lineage will move in a series of jumps: if its current location is z, then it will jump to $z + x$ (resp., $z + L^\alpha x$) due to a small (resp., large) event with respective intensities

$$L_{R_s}(0, x) u_s \frac{dx}{\pi R_s^2} \quad \text{and} \quad \frac{L_{R_B}(0, x)}{\rho_L} u_B \frac{dx}{\pi R_B^2}, \tag{1}$$

where $L_R(x, y)$ denotes the volume of the intersection $B(x, R) \cap B(y, R)$ [viewed as a subset of $\mathbb{T}(L)$ for the first intensity measure, and of $\mathbb{T}(L^{1-\alpha})$ for the second].

To see this, note first that by translation invariance of the model we may suppose that $z = 0$. In order for the lineage to experience a small jump, say, from the origin to x, the origin and the position x must be covered by the same event. This means that the center of the event must lie in both $B(0, R_s)$ and $B(x, R_s)$. The rate at which such events occur is $L_{R_s}(0, x)$. The lineage will only jump if it is sampled from the portion u_s of the population that are offspring of the event and then it will jump to the position of its parent, which is uniformly distributed on a ball of area πR_s^2. Combining these observations gives the first intensity in (1). A lineage ancestral to a single locus in a single individual thus follows a compound Poisson process on $\mathbb{T}(L)$.

Suppose now that we sample a single individual, but trace back its ancestry at *both* loci. We start with a single lineage which moves, as above, in a series of jumps as long as it is in the fraction $u_s(1 - r_L)$ of "nonrecombinants" in the population. However, every time it is hit by a small event, there is a probability $u_s r_L$ that it was created by recombination from two parental chromosomes, whose locations are sampled uniformly at random from the region affected by the event. If this happens, we must follow two distinct lineages, one for each locus, which jump around $\mathbb{T}(L)$ in an a priori correlated manner (since they may be hit by the same events), until they coalesce again. This will happen if they are both affected by an event (small or large) and are both derived from the same parent (which for a given event has probability $1/j$ in our notation above).

Thus, the ancestry of the two loci from our sampled individual is encoded in a system of splitting and coalescing lineages. If we now sample two individuals, (A, B) and (a, b), we represent their genealogical relations at the two loci by a process $\mathcal{A}^L$ taking values in the set of partitions of $\{A, a, B, b\}$ whose blocks are labeled by an element of $\mathbb{T}(L)$. As in [1], at time $t \geq 0$ each block of $\mathcal{A}_t^L$ contains the labels of all the lineages having a common ancestor (i.e., carried by the same individual) t units of time in the past, and the mark of the block records the spatial location of this ancestor. The only difference with the ancestral process defined in [1] is that blocks can now split due to a recombination event.

Of course, if r_L is small, then the periods of time when the lineages are in a single individual, that is, during which they have coalesced and not split apart again, can be rather extensive. This has the potential to create strong correlations between the two loci. The other source of correlation is the large events which can cause coalescences between lineages even when they are geographically far apart. To gain an understanding of these correlations, we ask the following question:

The problem: Given α, ρ_L and r_L, is there a minimal distance D_L^* such that, asymptotically as $L \to \infty$,

- if we sample two individuals (A, B) and (a, b) at distance at least D_L^* from each other, then the coalescence time of the ancestral lineages of A and a is independent of that of the ancestral lineages of B and b (in other words, genealogies at the two loci are completely decorrelated);

- if two individuals are sampled at a distance less than D_L^*, then the genealogies at the two loci are correlated (i.e., the lineages ancestral to A and B, and, similarly, those of a and b, remain sufficiently "close together" for a sufficiently long time that there is a significant chance that the coalescence of A and a implies that of B and b at the same time or soon after)?

1.4. *Main results*. Before stating our main results, we introduce some notation. We shall always denote the types of the two individuals in our sample by (A, B) and (a, b). The same letters will be used to distinguish the corresponding ancestral lineages. As we briefly mentioned in the last section, the genealogical relationships between the two loci at time $t \geq 0$ before the present are represented by a marked partition of $\{A, a, B, b\}$, in which each block corresponds to an individual in the ancestral population at time t who carries lineages ancestral to our sample. The labels in the block are those of the corresponding lineages and the mark is the spatial location of the ancestor. For any such marked partition a_L, we write $\mathbb{P}_{a_L}$ for the probability measure under which the genealogical process starts from a_L, with the understanding that marks then evolve on the torus $\mathbb{T}(L)$. Typically, our initial configuration will be of the form

$$a_L := \{(\{A, B\}, x_L^1), (\{a, b\}, x_L^2)\},$$

where the separation $x_L := x_L^1 - x_L^2$ between the two sampled individuals will be assumed to be large. The coalescence times of the ancestral lineages at each locus are denoted by τ_{Aa}^L and τ_{Bb}^L. Finally, we write $|x|$ for the Euclidean norm of $x \in \mathbb{R}^2$ (or in a torus of any size) and $\sigma^2 > 0$ is a constant, whose value is given just after (7). (It corresponds, after a suitable space–time rescaling, to the limit as $L \to \infty$ of the variance of the displacement of a lineage during a time interval of length one.)

For later comparison, we first record the asymptotic behaviour of the coalescence time at a single locus. The proof of the following result is in Section 3.

PROPOSITION 1.2. *Suppose that for each $L \in \mathbb{N}$ the two individuals comprising our initial configuration a_L are at separation $x_L \in \mathbb{T}(L)$. Suppose also that $\frac{\log |x_L|}{\log L} \to \beta \in (\alpha, 1]$ as $L \to \infty$. (In particular, $\alpha < 1$ here.) Then*:

(a) *For all $t \in [\beta, 1]$,*

$$\lim_{L\to\infty} \mathbb{P}_{a_L}[\tau_{Aa}^L > \rho_L L^{2(t-\alpha)}] = \frac{\beta - \alpha}{t - \alpha}.$$

(b) *For all $t > 0$,*

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau_{Aa}^L > \frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)} \log Lt\right] = \frac{\beta - \alpha}{1 - \alpha} e^{-t}.$$

REMARK 1.3. Observe that the timescale considered in case (b) above coincides with the quantity ϖ_L defined in Theorem 3.3 of [1]. Indeed, using the notation of [1], the variance σ^2 is given by the following limit:

$$\sigma^2 = \lim_{L\to\infty} \frac{\rho_L}{L^{2\alpha}}\sigma_s^2 + \sigma_B^2,$$

where σ_s^2 and σ_B^2 are defined in equation (20) of [1]. Now, if $\rho_L L^{-2\alpha} \to 0$ as in case (a) of Theorem 3.3, we obtain

$$\frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)}\log L \approx \frac{1-\alpha}{2\pi\sigma_B^2}\rho_L L^{2(1-\alpha)}\log L,$$

while if $\rho_L L^{-2\alpha} \to 1/b > 0$, we have

$$\begin{aligned}\frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)}\log L &\approx \frac{1-\alpha}{2\pi((1/b)\sigma_s^2+\sigma_B^2)}\frac{1}{b}L^2\log L\\ &= \frac{1-\alpha}{2\pi(\sigma_s^2+b\sigma_B^2)}L^2\log L.\end{aligned}$$

In both cases, the timescale considered in Proposition 1.2 is the same as the quantity ϖ_L of Theorem 3.3 of [1].

In the case $\alpha = 0$, Proposition 1.2 precisely matches corresponding results of [6] and [12] for coalescing random walks on a torus in $\mathbb{Z}^2$. For $\alpha > 0$, we see that if lineages start at a separation of $\mathcal{O}(L^\beta)$, with $\beta > \alpha$, then the small events don't affect the asymptotic coalescence times; they are the same as those for a random walk with bounded jumps on $\mathbb{T}(L^{1-\alpha})$ started at separation $\mathcal{O}(L^{\beta(1-\alpha)})$. In particular, the first statement tells us that the chance that coalescence occurs at a time $\ll \rho_L L^{2(1-\alpha)}\log L$ is $(1-\beta)/(1-\alpha)$. If this does not happen, then since the time taken for the random walks to reach their equilibrium distribution is $\mathcal{O}(\rho_L L^{2(1-\alpha)}\log L)$, in these units, the additional time that we must wait to see a coalescence is asymptotically exponential.

When we consider the genealogies at two loci, several regimes appear depending on the recombination rate and the initial distance between the individuals sampled.

THEOREM 1.4. *Suppose $(a_L)_{L\in\mathbb{N}}$ is as in Proposition 1.2. If*

$$\limsup_{L\to\infty} \frac{\log(1+\log\rho_L/(r_L\rho_L))}{2\log L} \le \beta - \alpha, \tag{2}$$

then we have the following:

(a) *For all $t \in [\beta, 1]$,*

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\big[\tau_{Aa}^L \wedge \tau_{Bb}^L > \rho_L L^{2(t-\alpha)}\big] = \frac{(\beta-\alpha)^2}{(t-\alpha)^2}.$$

(b) *For all* $t > 0$,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau^L_{Aa} \wedge \tau^L_{Bb} > \frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)} \log Lt\right] = \frac{(\beta-\alpha)^2}{(1-\alpha)^2} e^{-2t}.$$

Under the conditions of Theorem 1.4, the individuals are initially sampled at a distance much larger than the radius of the large events, and recombination is fast enough for the coalescence times at the two loci to be asymptotically independent (see Remark 1.7). For slower recombination rates this is no longer the case. When Condition (2) is not satisfied, we have instead:

THEOREM 1.5. *Suppose* $(a_L)_{L\in\mathbb{N}}$ *is as in Proposition 1.2. Assume there exists* $\gamma \in (\beta, 1)$ *such that*

$$\lim_{L\to\infty} \frac{\log(1 + \log \rho_L/(r_L\rho_L))}{2\log L} = \gamma - \alpha. \tag{3}$$

Then:

(a) *For all* $t \in [\beta, \gamma]$,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}[\tau^L_{Aa} \wedge \tau^L_{Bb} > \rho_L L^{2(t-\alpha)}] = \frac{\beta-\alpha}{t-\alpha}.$$

(b) *For all* $t \in (\gamma, 1]$,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}[\tau^L_{Aa} \wedge \tau^L_{Bb} > \rho_L L^{2(t-\alpha)}] = \frac{(\beta-\alpha)(\gamma-\alpha)^2}{(\gamma-\alpha)(t-\alpha)^2}.$$

(c) *For all* $t > 0$,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau^L_{Aa} \wedge \tau^L_{Bb} > \frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)} \log Lt\right] = \frac{(\beta-\alpha)(\gamma-\alpha)^2}{(\gamma-\alpha)(1-\alpha)^2} e^{-2t}.$$

This time, we observe a "phase transition" at time $\rho_L L^{2(\gamma-\alpha)}$. Asymptotically, coalescence times are completely correlated for times of $\mathcal{O}(\rho_L L^{2(\gamma-\alpha)})$, but conditional on being greater than this "decorrelation threshold" they are independent. To understand this threshold, recall from Proposition 1.2 that, initially, coalescence of lineages ancestral to a single locus happens on the exponential timescale $\rho_L L^{2(t-\alpha)}$, $t \in [\beta, 1]$ and is driven by large events. This tells us that the effect of recombination will be felt only if exactly one of the lineages ancestral to A and B (or to a and b) is "hit" by a large event. Since recombination events between A and B result in only a small separation of the corresponding ancestral lineages, we can expect that many of them will rapidly be followed by coalescence of the corresponding lineages (due to small events). This leads us to the idea of an "effective" recombination event, which is one following which at least one of the lineages ancestral to A and B is affected by a large event *before* they coalesce due

to small events. We shall see in Proposition 4.1 that recombination is "effective" on the linear timescale $\rho_L(1 + (\log \rho_L)/(r_L\rho_L))t$, $t \geq 0$. Under condition (3) the timescales of coalescence and effective recombination cross over precisely at time $\rho_L L^{2(\gamma-\alpha)}$.

Two cases remain:

The case $\alpha < \beta \leq 1$, $\gamma \geq 1$: If $\gamma > 1$, the arguments of the proof of Theorem 1.5 show that the recombination is too slow to be effective on the timescale of coalescence and so the coalescence times at the two loci are completely correlated and are given by Proposition 1.2. For $\gamma = 1$, the result depends on the precise form of $(\log \rho_L)/(r_L\rho_L)$. If it remains close enough to $L^{2(1-\alpha)}$ (or smaller), the proof of Theorem 1.5 shows that lineages are completely correlated on the timescale $\rho_L L^{2(t-\alpha)}$, $t \leq 1$, and then, conditional on not having coalesced before $\rho_L L^{2(1-\alpha)}$, they evolve independently on the timescale $\rho_L L^{2(1-\alpha)} \log L\, t$. On the other hand, if $(\log \rho_L)/(r_L\rho_L L^{2(1-\alpha)})$ grows to infinity sufficiently fast, then, just as in the case $\gamma > 1$, recombination is too slow to be effective.

The case $\beta \leq \alpha \leq 1$: If we drop our assumption that the separation of the individuals in our sample is much greater than the radius of the largest events, then we can no longer make such precise statements. Proposition 6.4(a) in [1] (with $\psi_L = L^\alpha$) shows that the coalescence time for lineages ancestral to a single locus will now be at most $\mathcal{O}(\rho_L)$. This does tell us that if $r_L\rho_L \to 0$ as $L \to \infty$, then asymptotically we will not see any recombination before coalescence and the coalescence times τ^L_{Aa} and τ^L_{Bb} are identical. However, in contrast to the setting of Proposition 1.2, even asymptotically, their common value depends on the exact separation of the individuals sampled. The same reasoning is valid when $r_L \ll (\log \rho_L)/\rho_L$. In this case, although we may see some recombination events before any coalescence occurs, a closer look at the proof of Proposition 4.1 reveals that the time spent in distinct individuals by the two lineages ancestral to A, B, say, in $\mathcal{O}(\rho_L)$ units of time, is negligible compared to ρ_L. Thus, with high probability, any large event affecting lineages ancestral to our sample will occur at a time when the lineages ancestral to A and B are in the same individual, (as are those ancestral to a and b). As a result, once again $\tau^L_{Aa} = \tau^L_{Bb}$ with probability tending to 1.

On the other hand, suppose r_L remains large enough that lineages ancestral to A and B have a chance to be hit by a large event while they are in different individuals and thus jump to a separation $\mathcal{O}(L^\alpha)$ (the *effective recombination* of Section 4.1). We are still unable to recover precise results. The reason is that even after such an event, we may be in a situation in which all lineages could be hit by the same large event, or at least remain at separations $\mathcal{O}(L^\alpha)$. But we shall see that a key to the proofs of Theorems 1.4 and 1.5 is the fact that, in the settings considered there, where individuals are sampled from far apart, whenever two lineages come to within $2R_B L^\alpha$ of one another, the other ancestral lineages are still very far from them. This gives the pair time to merge without "interference" from the other lineages. Since lineages at separations $\mathcal{O}(L^\alpha)$ are correlated and their coalescence

times depend strongly on their precise (geographical) paths on this scale, it is difficult to quantify the extent to which the fact that the ancestral lines of A, B and of a, b start within the same individuals makes the coalescence times τ^L_{Aa} and τ^L_{Bb} *more* correlated. Nonetheless, this is an important question and will be addressed elsewhere.

To answer our initial question, we see from Theorems 1.4 and 1.5 (and the subsequent discussion) that D^*_L is informally given by

$$\log\left(1+\frac{\log \rho_L}{r_L \rho_L}\right) \approx 2(\log D^*_L - \alpha \log L), \qquad \text{i.e., } D^*_L \approx L^\alpha \sqrt{1+\frac{\log \rho_L}{r_L \rho_L}}.$$

When the sampling distance is greater than the radius of the largest events, correlated genealogies are only possible when recombination is slow enough, or large events occur rarely enough, that $(\log \rho_L)/(r_L \rho_L) \gg 1$. If, for instance, $r_L \equiv r > 0$, the two loci are always asymptotically decorrelated. On the other hand, if γ is as in (3) [note that γ does not need to exist for condition (2) to hold] and the sampling distance is L^β, Theorem 1.4 shows that if $\beta \geq \gamma$, the genealogies at the two loci are asymptotically independent, whereas Theorem 1.5 tells us that if $\beta \in (\alpha, \gamma)$, there is a first phase of complete correlation. Thus, D^*_L "$=$" L^γ.

Before closing this section, let us make two remarks:

REMARK 1.6 (Bounds on the rates of large events). Recall that we imposed the condition $\log L \leq \rho_L \leq C L^{2\alpha}$. The reason for the upper bound is that in [1], we showed that the coalescence of the ancestral lineages is then driven by the large events and, moreover, is very rapid once lineages are at separation $\mathcal{O}(L^\alpha)$ (see the proof of Theorem 3.3 in [1]). Similar results should hold, although on different timescales, in the other cases presented in [1]. However, to keep the presentation of our results as simple as possible, we have chosen to concentrate on this upper bound. The (rather undemanding) lower bound is needed in the proof of Proposition 4.4.

REMARK 1.7 (Generalization of Theorems 1.4 and 1.5 to distinct coalescence times). In these two theorems, we could also consider the probabilities of events of the form $\{\tau^L_{Aa} > \rho_L L^{2(t-\alpha)}$ and $\tau^L_{Bb} > \rho_L L^{2(t'-\alpha)}\}$, with $t < t'$. However, they can be computed by a simple application of Theorem 1.4 or 1.5 at time t, and the Markov property. Indeed, arguments similar to those of the proofs of Lemma C and Lemma 3.7 in Section 3 tell us that the distance between lineages ancestral to B and b at time $\rho_L L^{2(t-\alpha)}$, conditional on not having coalesced by this time, lies in $[L^t/(\log L), L^t \log L]$. Proposition 1.2 then enables us to conclude. We leave this generalization to the reader.

The rest of the paper is laid out as follows. In Section 2 we provide more detail of the motivation for the question addressed here. In Section 3 we prove Proposition 1.2 and collect several results on genealogies of a sample from a single locus

that we shall need in the sequel. Since most of these results are close to those established in [1], or require techniques used in [6] and [12] for similar questions on the discrete torus, their proofs will only be sketched. Our main results are proved in Section 4: we define an effective recombination rate in Section 4.1, use it to find an upper bound on the time we must wait before the two lineages ancestral to A and B start to evolve independently in Section 4.2 and finally derive the asymptotic coalescence times of our two pairs of lineages in Section 4.3.

2. Biological motivation. In this section we expand on the biological motivation for our work.

It has long been understood that for many models of spatially distributed populations, if individuals are sampled sufficiently far from one another, then the genealogical tree that records the relationships between the alleles carried by those individuals at a single locus is well-approximated by a Kingman coalescent with an "effective population size" capturing the influence of the geographical structure. If the underlying population model is a stepping stone model, with the population residing in discrete demes located at the vertices of $\mathbb{Z}^2$ or $\mathbb{T}(L) \cap \mathbb{Z}^2$, individuals reproducing within demes and migration modeled as a random walk, then the genealogical trees relating individuals in a finite sample from the population are traced out by a system of coalescing random walks. The case in which random walks coalesce instantly on meeting corresponds (loosely) to a single individual living in each deme in which case the stepping stone model reduces to the voter model. In this setting, and with symmetric nearest neighbour migration, convergence to the Kingman coalescent as the separation of individuals in the initial sample tends to infinity was established for $\mathbb{Z}^2$ in [6, 7], and for $\mathbb{T}(L) \cap \mathbb{Z}^2$ in [4]. In [5, 12], Zähle, Cox and Durrett prove the same kind of convergence for coalescing random walks on $\mathbb{T}(L) \cap \mathbb{Z}^2$ with finite variance jumps and delayed coalescence (describing the genealogy for a sample from Kimura's stepping stone model on the discrete torus in which reproduction within each deme is modeled by a Wright–Fisher diffusion). In [10], Limic and Sturm prove the analogous result when mergers between random walks within a deme are not necessarily pairwise. In the same spirit but on the continuous space $\mathbb{T}(L)$ and with additional *large* extinction/recolonization events (similar to those described in Section 1.2), the same asymptotic behavior is obtained in [1] for the systems of coalescing compound Poisson processes describing the genealogy of a sample from the spatial Λ-Fleming–Viot process, under suitable conditions on the frequency and extent of the large events.

In all of these examples, the result stems from a separation of timescales. For example, in [1] we were concerned with the genealogy of a sample picked uniformly at random from the whole torus. Under this assumption, the time that two lineages need to be "gathered" close enough together that they can both be affected by the same event dominates the additional time the lineages take to coalesce, having being gathered. As explained in Section 1.4, this decomposition does not hold when

lineages start too close together, and so the tools developed for well-separated samples are of no use in the study of local correlations. However, although we still cannot make precise statements about the genealogy of samples which are initially too close together, the work of Sections 4.1 and 4.2, which are concerned with "effective recombination" and "decorrelation," provides a much better understanding than we had before of the local mechanisms that create correlations between nearby lineages, how strong these correlations are, and how to "escape" them.

Our main results in this paper are concerned with samples taken at "intermediate" scales. Individuals are sampled at pairwise distances much larger than the radius of the largest events, but these distances can still be much less than the radius of the torus. In this case, the "gathering time" of two lineages starting at separation x_L depends on that separation, but asymptotically this dependence is only through $\log|x_L|/\log L$. As in the case of a uniform sample, the gathering time dominates the additional time to coalescence. In Theorem 3.3 of [1] we showed that if we sample a finite number of individuals uniformly at random from the geographic range of a population which is subject to small and large demographic events, then measuring time in units of size $\varpi_L = \frac{1-\alpha}{2\pi\sigma^2}(\rho_L/L^{2\alpha})L^2\log L$ (under the assumption on ρ_L used here), their genealogical tree is determined by Kingman's coalescent. In particular, if $\rho_L < L^{2\alpha}$ (i.e., large events are not too rare), one major effect of the presence of large extinction/recolonization events is to reduce the *effective population size* and, consequently, genetic diversity. The assumption of uniform sampling guarantees that initially ancestral lineages are $\mathcal{O}(L/\log L)$ apart. Proposition 1.2 extends the result by showing that, if we sample our individuals from much closer together, then we should consider two timescales. The first is $(\rho_L/L^{2\alpha})L^{2t}, t \in [\beta, 1]$. The second kicks in after $\mathcal{O}(\rho_L L^{2(1-\alpha)})$, when the lineages start to feel the fact that space is limited and their ancestries evolve on the linear timescale $\varpi_L t$. Now, by the same reasoning, if there were no large events, these timescales would be, respectively, $L^{2t}, t \in [\beta, 1]$, and $\frac{1}{2\pi\sigma_s^2}L^2\log Lt$, $t > 0$. Of course, one never observes genealogies directly and so, for illustration, we introduce (infinitely many alleles) mutation into our model and compute the probability that two individuals sampled at a given separation are *identical by descent* (IBD) as a function of the exponent β. In other words, what is the probability that the two individuals carry the same type (at a given locus) because it was inherited from a common ancestor.

Since mutations are generally assumed to occur at a linear rate, while the first phase of the genealogical tree develops on a much slower exponential timescale, for a given time parameter $t \in [\beta, 1]$, asymptotically as $L \to \infty$, we would see either zero or infinitely many mutations on the tree. However, let us suppose that L is large and write θ for the mutation rate at locus A. We denote by c_L the ratio $\rho_L/L^{2\alpha}$. Since IBD is equivalent to our individuals experiencing no mutation

between the time of their most recent common ancestor and the present, the probability of IBD of two individuals sampled at distance L^β is given by

$$\begin{aligned}\mathbb{E}_{L^\beta}[e^{-2\theta\tau^L_{Aa}}] &\approx \mathbb{E}_{L^\beta}\big[e^{-2\theta\tau^L_{Aa}}\mathbf{1}_{\{c_L L^{2\beta}\le\tau^L_{Aa}\le c_L L^2\}}\big] + \mathbb{E}_{L^\beta}\big[e^{-2\theta\tau^L_{Aa}}\mathbf{1}_{\{c_L L^2<\tau^L_{Aa}\}}\big] \\ &= \int_{c_L L^{2\beta}}^{c_L L^2} e^{-2\theta t}\mathbb{P}_{L^\beta}[\tau^L_{Aa}\in dt] + \int_{c_L L^2}^{\infty} e^{-2\theta t}\mathbb{P}_{L^\beta}[\tau^L_{Aa}\in dt] \qquad (4) \\ &\approx (\beta-\alpha)\int_\beta^1 \frac{e^{-2\theta c_L L^{2u}}}{(u-\alpha)^2}\,du + \frac{\beta-\alpha}{1-\alpha}\int_{1/\log L}^{\infty} e^{-2\theta c_L L^2\log Lu}e^{-u}\,du,\end{aligned}$$

where the last line uses a change of variable and the results of Proposition 1.2. The corresponding quantity when there are no large events is given by

$$\beta\int_\beta^1 \frac{e^{-2\theta L^{2u}}}{u^2}\,du + \beta\int_{1/\log L}^{\infty} e^{-2\theta L^2\log Lu}e^{-u}\,du.$$

The leading term in each sum is the first one, and we thus see that if $c_L \ll 1$ (i.e., $\rho_L \ll L^{2\alpha}$), then, as expected, the probability of IBD is higher in the presence of large events and, moreover, as a consequence of shorter genealogies, correlations between gene frequencies persist over longer spatial scales. See Figure 1 for an illustration (in which only the leading terms are plotted). In classical models IBD

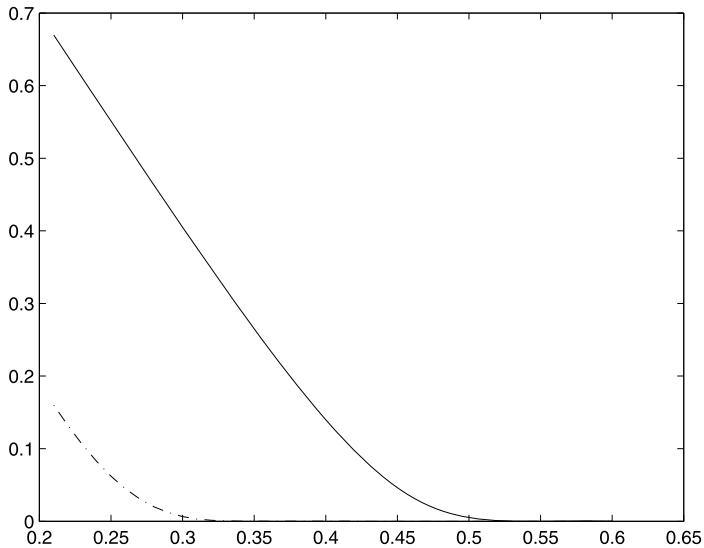

FIG. 1. *Probability of IBD at a single locus, as a function of β. Here, $L = 10^5$, $\alpha = 0.1$, $c_L = 0.01$ and $\theta = 10^{-3}$. The solid line corresponds to the case with small and large events, the dash–dot line to the case with only small events. Geographical correlations vanish around $\beta = 0.32$ without large events, and are positive up to $\beta = 0.52$ when large events occur.*

decays approximately exponentially with the sampling distance, at least over small scales. In [2], a numerical investigation of a similar model to that presented here revealed approximately exponential decay over small scales followed by a transition to a different exponential rate over somewhat larger scales. Since the (rigorous) results of Proposition 1.2 only apply for sufficiently well separated samples, our arguments above cannot capture this. They do, on the other hand, give a clear indication of the reduction of effective population size due to large events.

Local bottlenecks are not the only explanations for a reduced effective population size. For example, selection or fluctuating population sizes can have the same effect, and so we should like to find a more "personal" signature of the presence of demographic events of different orders of magnitude. The idea that we explore here is to consider several loci on the same chromosome, subject to recombination, and to investigate the pattern of *linkage disequilibrium* obtained under the assumptions of Section 1.4. Using the results of Theorems 1.4 and 1.5, we have

$$\begin{aligned}
&\mathbb{P}_{L^\beta}[\text{IBD at both loci}]\\
&\quad = \mathbb{E}_{L^\beta}\big[e^{-2(\theta_1\tau^L_{Aa}+\theta_2\tau^L_{Bb})}\big]\\
&\quad \approx \mathbb{E}_{L^\beta}\big[e^{-2\theta_1\tau^L_{Aa}-2\theta_2\tau^L_{Bb}}\mathbf{1}_{\{c_L L^{2\beta}\le\tau^L_{Aa}=\tau^L_{Bb}\le c_L L^{2\gamma}\}}\big]\\
&\qquad + \mathbb{E}_{L^\beta}\big[e^{-2\theta_1\tau^L_{Aa}}\mathbf{1}_{\{\tau^L_{Aa}>c_L L^{2\gamma}\}}\big]\times\mathbb{E}_{L^\beta}\big[e^{-2\theta_2\tau^L_{Bb}}\mathbf{1}_{\{\tau^L_{Bb}>c_L L^{2\gamma}\}}\big],
\end{aligned}$$

where θ_1 and θ_2 denote the mutation rates at each locus and the first integral is 0 if condition (2) holds (i.e., if there is no first period of complete correlation). By the same computations as in (4), the leading terms in this expression are

$$\begin{aligned}
&(\beta-\alpha)\int_\beta^\gamma \frac{e^{-2(\theta_1+\theta_2)c_L L^{2u}}}{(u-\alpha)^2}\,du\\
&\qquad + (\beta-\alpha)^2\bigg(\int_\gamma^1 \frac{e^{-2\theta_1 c_L L^{2u}}}{(u-\alpha)^2}\,du\bigg)\bigg(\int_\gamma^1 \frac{e^{-2\theta_2 c_L L^{2u}}}{(u-\alpha)^2}\,du\bigg).
\end{aligned}\tag{5}$$

On the other hand, when there are no large events, the analysis of Lemma 4.3 (with *effective recombination* replaced by *recombination* and the separation to attain of the order of L) tells us that the time two lineages initially in the same individual need to "decorrelate" is of the order of $r_L^{-1}\log L$. Here r_L^{-1} is the expected time to wait until we see a recombination event, and $\log L$ is (roughly) the mean number of recombination events before we see one after which the lineages remain separated for a duration $\mathcal{O}(L^t)$ for some $t\in[\beta,1]$. Hence, when there are only small events, the leading terms in the probability of IBD at both loci are

$$\beta\int_\beta^{\gamma_L^*}\frac{e^{-2(\theta_1+\theta_2)L^{2u}}}{u^2}\,du+\beta^2\bigg(\int_{\gamma_L^*}^1\frac{e^{-2\theta_1 L^{2u}}}{u^2}\,du\bigg)\bigg(\int_{\gamma_L^*}^1\frac{e^{-2\theta_2 L^{2u}}}{u^2}\,du\bigg),$$

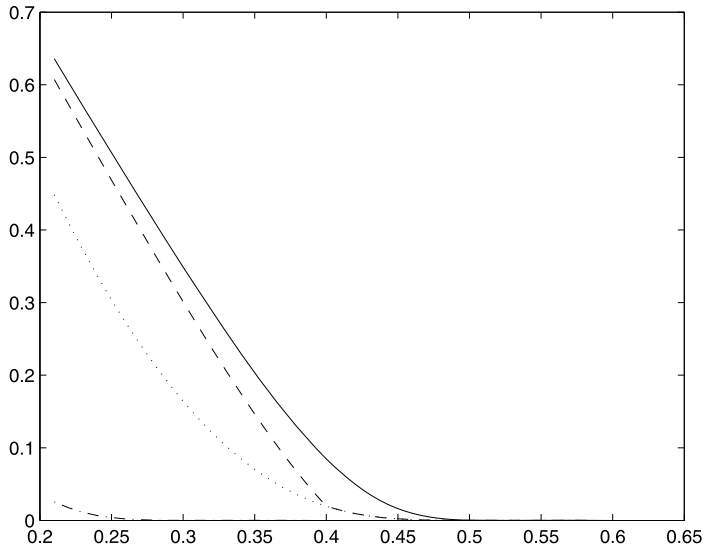

FIG. 2. *Probability of IBD at both loci, as a function of* β. *As in Figure* 1, $L = 10^5$, $\alpha = 0.1$, $c_L = 0.01$ *and* $\theta_1 = \theta_2 = 10^{-3}$. *The solid line corresponds to the case* $\gamma \geq 1$ (*complete correlation for any* β), *the dotted line to the case* $\gamma \leq \alpha$ (*decorrelation for any* β) *and the dashed line to the intermediate case* $\gamma = 0.4$. *The dash–dot line corresponds to the case without large events, for which* γ_L^* *is computed from the same parameter values* (*i.e.*, $\gamma_L^* = 0.2$).

where we have set $\gamma_L^* := \log(r_L^{-1} \log L)/(2 \log L)$ and the first integral is again zero if $\beta > \gamma_L^*$. Figure 2 compares the different curves obtained when (i) we always have decorrelation ($\gamma \leq \alpha$), (ii) we always have complete correlation ($\gamma \geq 1$), or (iii) when we have a transition between these two regimes [$\gamma \in (\alpha, 1)$]. As expected, we see that the probability of IBD at both loci is higher in the presence of large events (when $\rho_L \leq L^{2\alpha}$), and there is correlation between the two loci when individuals are sampled over large spatial distances. Furthermore, (5) gives us an idea of how the correlations between the two loci decay with sampling distance, as this grows from the radius of the large events to the whole population range. Correlations for sampling distances smaller than or equal to the size of the large events will be the object of future work.

3. Genealogies at one locus. In this section we prove Proposition 1.2. In the process we introduce a rescaling of the spatial motion of our ancestral lineages and collect together several results on the time required to "gather" two lineages to within distance $2R_B L^\alpha$ which will also be needed in Section 4. Since the techniques mirror closely those used in previous work, in the interests of brevity, we restrict ourselves to sketching the proofs and providing references where appropriate.

Assume for the rest of this section that $\alpha < 1$.

The following local central limit theorem, corresponding to Lemma 5.4 of [1], is the key to understanding the behavior of two lineages. Suppose that for each $L \in \mathbb{N}$, ℓ^L is a Lévy process on $\mathbb{T}(L)$ such that $\ell^L(1) - \ell^L(0)$ has a covariance matrix of the form $\sigma_L^2 \,\mathrm{Id}$, and that:

(i) there exists $\sigma^2 > 0$ such that $\sigma_L^2 \to \sigma^2$ as $L \to \infty$;
(ii) $\mathrm{E}_0[|\ell^L(1)|^4]$ is bounded uniformly in L.

We shall implicitly suppose that all processes ℓ^L are defined on the same probability space, and that under the probability measure P_x the Lévy process we consider starts at x. Let $(d_L)_{L\geq 1}$ be a sequence of positive reals such that $\liminf_{L\to\infty} d_L > 0$ and $\frac{\log^+(d_L)}{\log L} \to \eta \in [0, 1)$. Finally, let us write $p^L(x, t)$ for $\mathrm{P}_x[\ell^L(t) \in B(0, d_L)]$ and $\lfloor z \rfloor$ for the integer part of $z \in \mathbb{R}$.

LEMMA A (Lemma 5.4 in [1]).

(a) *Let $\varepsilon_L := (\log L)^{-1/2}$. There exists a constant $C_1 < \infty$ such that for every $L \geq 2$,*

$$\sup_{t\geq \lfloor \varepsilon_L L^2 \rfloor} \sup_{x\in\mathbb{T}(L)} \frac{\lfloor \varepsilon_L L^2 \rfloor}{d_L^2} p^L(x, t) \leq C_1.$$

(b) *If $v_L \to \infty$ as $L \to \infty$, then*

$$\lim_{L\to\infty} \sup_{t\geq \lfloor v_L L^2 \rfloor} \sup_{x\in\mathbb{T}(L)} \frac{L^2}{d_L^2} \left| p^L(x, t) - \frac{\pi d_L^2}{L^2} \right| = 0.$$

(c) *If $u_L \to \infty$ as $L \to \infty$ and $I(d_L, x) := 1 + (|x|^2 \vee d_L^2)$, then*

$$\lim_{L\to\infty} \sup_{x\in\mathbb{T}(L)} \sup_{u_L I(d_L,x)\leq t\leq \varepsilon_L L^2} \frac{2\sigma_L^2 t}{d_L^2} \left| p^L(x, t) - \frac{d_L^2}{2\sigma_L^2 t} \right| = 0.$$

(d) *There exists a constant $C_2 < \infty$ such that for every $L \geq 1$,*

$$\sup_{t\geq 0} \sup_{x\in\mathbb{T}(L)} \left(1 + \frac{|x|^2}{d_L^2}\right) p^L(x, t) \leq C_2.$$

What Lemma A shows is that, for times which are large but of order at most $\mathcal{O}(L^2)$, ℓ^L behaves like two-dimensional Brownian motion (case *c*), and, in particular, it has not yet explored the torus enough to "see" that space is limited. On the other hand, $\ell^L(t)$ is nearly uniformly distributed over $\mathbb{T}(L)$ at any time much greater than L^2 (case *b*).

Fix $R > 0$. As a direct corollary of this local central limit theorem, we proved in Lemma 5.5 of [1] that, if $T(R, \ell^L)$ denotes the entrance time of ℓ^L into the ball $B(0, R)$, then the following inequality holds.

LEMMA B (Lemma 5.5 in [1]). *Let $(U_L)_{L\geq 1}$ and $(u_L)_{L\geq 1}$ be two sequences increasing to infinity such that $U_L L^{-2} \to \infty$ as $L \to \infty$ and $2u_L \leq L^2(\log L)^{-1/2}$ for every $L \geq 1$. Then, there exist $C_0 > 0$ and $L_0 \in \mathbb{N}$ such that for every sequence $(U'_L)_{L\geq 1}$ satisfying $U'_L \geq U_L$ for each L, every $L \geq L_0$ and all $x \in \mathbb{T}(L)$,*

$$\mathrm{P}_x\big[T(R, \ell^L) \in [U'_L - u_L, U'_L]\big] \leq \frac{C_0 u_L}{L^2}.$$

Lemma B tells us about the regime in which ℓ^L has already homogenized over $\mathbb{T}(L)$. Using exactly the same method, but employing parts (c) and (d) of Lemma A rather than (b), we obtain the analogous result for the regime in which ℓ^L behaves as Brownian motion on $\mathbb{R}^2$:

LEMMA 3.1. *If $U_L \leq L^2(\log L)^{-1/2}$ for each $L \geq 1$, $U_L, u_L \to \infty$ and $u_L/U_L \to 0$ as $L \to \infty$, then there exist $C_1 > 0$ and $L_1 \in \mathbb{N}$ such that for every sequence $(U'_L)_{L\geq 1}$ satisfying $U_L \leq U'_L \leq L^2(\log L)^{-1/2}$ for each L, for every $L \geq L_1$ and $x \in \mathbb{T}(L)$,*

$$\mathrm{P}_x\big[T(R, \ell^L) \in [U'_L - u_L, U'_L]\big] \leq \frac{C_1 u_L}{U'_L}.$$

Let us now introduce the processes to which we wish to apply these results. For each $L \in \mathbb{N}$, let $\{\tilde{X}^L_{Aa}(t), t \geq 0\}$ be the process recording the difference between the locations on $\mathbb{T}(L)$ of the ancestral lineages of A and a (i.e., the first locus of each of the two individuals sampled). The process $\tilde{X}^L_{Aa}$ is the difference between two dependent compound Poisson processes. Under the probability measures we shall use, it is a Markov process (see Remark 3.2). Observe that, because the largest events have radius $R_B L^\alpha$, the lineages have to be within a distance less than $2R_B L^\alpha$ of each other to be hit by the same event. As a consequence, the law of $\tilde{X}^L_{Aa}$ outside $B(0, 2R_B L^\alpha)$ is equal to that of the difference $\tilde{Y}^L$ of two i.i.d. Lévy processes, each of which follows the evolution given in (1), and thus is also equal to the law of the motion of a single lineage run at twice the speed. We define the processes X^L_{Aa} and Y^L by

$$X^L_{Aa}(t) = \frac{1}{L^\alpha}\tilde{X}^L_{Aa}(\rho_L t) \quad \text{and} \quad Y^L(t) = \frac{1}{L^\alpha}\tilde{Y}^L(\rho_L t), \qquad t \geq 0, \tag{6}$$

both evolving on $\mathbb{T}(L^{1-\alpha})$. Using computations from the proof of Proposition 6.2 in [1] and the jump intensities given in (1), we find that the covariance matrix of $Y^L(1) - Y^L(0)$ is the identity matrix multiplied by

$$\begin{aligned} &2\left\{\frac{u_s \rho_L}{\pi R_s^2 L^{2\alpha}} \int_{\mathbb{R}^2} (x_1)^2 L_{R_s}(x, 0)\, dx + \frac{u_B}{\pi R_B^2} \int_{\mathbb{R}^2} (x_1)^2 L_{R_B}(x, 0)\, dx\right\} + o(1) \\ &\quad =: 2\sigma_L^2 + o(1), \end{aligned} \tag{7}$$

with σ_L^2 tending to a finite limit $\sigma^2 > 0$ as $L \to \infty$ (by our assumption on $L^{-2\alpha}\rho_L$). The $o(1)$ remainder here is the error we make by considering $\tilde{Y}^L$ as evolving on $\mathbb{R}^2$ instead of $\mathbb{T}(L)$ (see the proof of Proposition 6.2 in [1]). Assumption (ii) is also satisfied, and so Lemma A and its corollaries apply to $(Y^L)_{L\geq 1}$, with the torus side-length L replaced by $L^{1-\alpha}$. Furthermore, X^L_{Aa} and Y^L follow the same evolution outside $B(0, 2R_B)$ for every L. This will be sufficient to prove Proposition 1.2: we shall show that the time the ancestral lineages of A and a need to coalesce once they are within distance $2R_BL^\alpha$ of one another [or, equivalently, once X^L_{Aa} has entered $B(0, 2R_B)$] is negligible compared to the time they need to be gathered at distance $2R_BL^\alpha$. It is therefore the "gathering time" that dictates the coalescence time of two lineages starting at separation $|x_L| \gg L^\alpha$.

REMARK 3.2. It is here that we take advantage of the form of our recombination mechanism (recall Remark 1.1). When $\tilde{X}^L_{Aa}(t) \neq 0$, its future evolution is determined by the homogeneous Poisson point processes of events Π^L_B and Π^L_s, and depends only on the current separation of the two lineages. If $\tilde{X}^L_{Aa}(t) = 0$, the situation depends upon whether the two lineages are in the same individual (i.e., they have coalesced and will require a recombination event to separate again), or in two distinct individuals at the same spatial location. However, because of the form of our recombination mechanism, two lineages can jump onto the same location only if they are descendants of the same parent (in which case they necessarily coalesce). This means that provided we choose our initial condition in such a way that two lineages in the same spatial location are actually in the same individual, with probability one we will never see two lineages in distinct individuals but the same spatial location and so $\tilde{X}^L_{Aa}$ is indeed a Markov process under $\mathbb{P}_{a_L}$.

NOTATION 3.3. As at the beginning of the section, we assume that all Y^L's are defined on the same probability space, and start at x under the probability measure P_x. Since X^L_{Aa} is a function of the genealogical process of A, a, B and b, we retain the notation $\mathbb{P}_{a_L}$ when referring to it, and X^L_{Aa} then starts a.s. at $L^{-\alpha}x_L$ if $x_L \in \mathbb{T}(L)$ is the initial separation between lineages A and a.

The proof of Proposition 1.2 will require two subsidiary results. For each $L \in \mathbb{N}$, let T^L_{Aa} be the first time the two lineages A and a are at separation less than $2R_BL^\alpha$. Equivalently, $\rho_L^{-1}T^L_{Aa}$ is the entrance time of X^L_{Aa} into $B(0, 2R_B)$. By the observation made in the paragraph preceding Remark 3.2, $\rho_L^{-1}T^L_{Aa}$ under $\mathbb{P}_{a_L}$ has the same distribution as $T(2R_B, Y^L)$ under $\mathrm{P}_{L^{-\alpha}x_L}$, which yields the following lemma.

LEMMA 3.4. *Under the assumptions of Proposition 1.2, we have*

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\big[T^L_{Aa} > \rho_L L^{2(t-\alpha)}\big] = \frac{\beta-\alpha}{t-\alpha} \qquad \forall t \in [\beta, 1] \quad \textit{and} \tag{8}$$

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[T^L_{Aa} > \frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)}\log Lt\right] = \frac{\beta-\alpha}{1-\alpha}e^{-t} \qquad \forall t > 0. \tag{9}$$

Furthermore, for any $\beta_0 \in (\alpha, 1)$ *and* $\varepsilon > 0$, *the convergence in the first* (*resp.*, *second*) *expression is uniform in* $\beta, t \in [\beta_0, 1]$ (*resp.*, $\beta \in [\beta_0, 1]$ *and* $t \geq \varepsilon$) *and* a_L *such that* $|x_L| \in [L^\beta/(\log L), L^\beta \log L]$.

PROOF. When $\beta = 1$, the results are a weaker version of Proposition 6.2 in [1], in which the convergence in (9) is uniform over $t \geq 0$ and over the set of sequences $(x_L)_{L\geq 1}$ such that $|x_L| \geq L(\log L)^{-1}$ for every L. Here, we relax the condition on $(x_L)_{L\geq 1}$, but since the arguments in the proof of convergence (without requiring uniformity) only use the asymptotic behavior of $\log|x_L|$, they are still valid.

If $\beta < 1$, the reasoning is the same as in the proofs of Lemma 3.6 in [12] (note that as above we allow more general sequences of initial separations at the expense of the uniformity of the convergence) and Theorem 2 in [5]. This does not come as a surprise, since the same local central limit theorem applies to both Y^L [on $\mathbb{T}(L^{1-\alpha})$] and Zähle, Cox and Durrett's Y [on $\mathbb{T}(L) \cap \mathbb{Z}^2$] up to some constants depending on the geometry of the geographical patches considered. Hence, since X^L_{Aa} starts from $L^{-\alpha}x_L$ and $\log(L^{-\alpha}|x_L|)/(\log L) \to \beta - \alpha$ by assumption, we can write (as in Lemma 3.6 of [12])

$$\begin{aligned}
&\lim_{L\to\infty} \sup_{\beta\leq t\leq\kappa_L} \left| \mathbb{P}_{a_L}\left[T^L_{Aa} > \rho_L L^{2(t-\alpha)}\right] - \frac{\beta-\alpha}{t-\alpha} \right| \\
&\quad = \lim_{L\to\infty} \sup_{\beta\leq t\leq\kappa_L} \left| \mathrm{P}_{L^{-\alpha}x_L}\left[T(2R_B, Y^L) > L^{2(t-\alpha)}\right] - \frac{\beta-\alpha}{t-\alpha} \right| \\
&\quad = 0,
\end{aligned}$$

where $\kappa_L = 1 - (\log\log L)/(2\log L)$ [so that $L^{2(\kappa_L-\alpha)} = L^{2(1-\alpha)}/(\log L)$]. Now, as in Lemma 3.8 of [12], there exists $L_0 \in \mathbb{N}$ and a constant C such that, for every $L \geq L_0$ and $x \in \mathbb{T}(L)$,

$$\mathrm{P}_x\left[Y^L(s) = 0 \text{ for some } s \in \left[\frac{L^{2(1-\alpha)}}{\log L}, L^{2(1-\alpha)} \right] \right] \leq \frac{C\log\log L}{\log L}. \tag{10}$$

Combining these two results, we obtain (8).

Finally, (9) is the analogue of Theorem 2 in [5] and can either be proved using the same technique or in the same way as Proposition 6.2 in [1] (which, in addition, gives the appropriate constant in the time-rescaling). The uniform convergence stated in the second part of Lemma 3.4 follows from a direct application of the techniques of [12] and [1] cited above. □

The next result we need is the time that two lineages starting at separation at most $2R_BL^\alpha$ take to coalesce. Under our assumption that $(\rho_L L^{-2\alpha})_{L\geq 1}$ is bounded, Proposition 6.4(a) in [1] applied with $\psi_L := L^\alpha$ shows that for any sequence $(\phi_L)_{L\geq 1}$ tending to infinity, we have

$$\lim_{L\to\infty} \sup_{a'_L} \mathbb{P}_{a'_L}[\tau^L_{Aa} > \phi_L\rho_L] = 0, \tag{11}$$

where the supremum is taken over all configurations a'_L such that the distance between the blocks containing A and a is at most $2R_BL^\alpha$. Observe that in [1], only one individual reproduces during an event, and so if several lineages are affected by this event, they necessarily coalesce. Here, the distributions λ_s and λ_B of the number of potential parents are more general, but we assumed that their supports were compact. Thus, the probability that several individuals in the area of an event come from the same parent does not vanish as L tends to infinity, which is all that we need to prove (11).

REMARK 3.5. Since (11) shows that coming to within $2R_BL^\alpha$ is almost equivalent to coalescing for two lineages, this is the only point where the distributions λ_s and λ_B appear in our discussion.

PROOF OF PROPOSITION 1.2. Equipped with these results and the corollaries of Lemma A, we can now write for any given $t \in [\beta, 1]$

$$\begin{aligned}&\mathbb{P}_{a_L}\big[\tau^L_{Aa} > \rho_L L^{2(t-\alpha)}\big]\\ &\quad= \mathbb{P}_{a_L}\big[\tau^L_{Aa} > \rho_L L^{2(t-\alpha)}; T^L_{Aa} > \rho_L\big(L^{2(t-\alpha)} - \log L\big)\big]\\ &\qquad+ \mathbb{P}_{a_L}\big[\tau^L_{Aa} > \rho_L L^{2(t-\alpha)}; T^L_{Aa} \le \rho_L\big(L^{2(t-\alpha)} - \log L\big)\big].\end{aligned} \tag{12}$$

The second term on the right-hand side of (12) tends to zero by the strong Markov property applied at time T^L_{Aa} and (11) with $\phi_L = \log L$. Then, we have, for each L,

$$\begin{aligned}&\big|\mathbb{P}_{a_L}\big[\tau^L_{Aa} > \rho_L L^{2(t-\alpha)}; T^L_{Aa} > \rho_L\big(L^{2(t-\alpha)} - \log L\big)\big] - \mathbb{P}_{a_L}\big[T^L_{Aa} > \rho_L L^{2(t-\alpha)}\big]\big|\\ &\quad\le \mathbb{P}_{a_L}\big[\rho_L\big(L^{2(t-\alpha)} - \log L\big) \le T^L_{Aa} \le \rho_L L^{2(t-\alpha)}\big]\\ &\quad= \mathrm{P}_{L^{-\alpha}x_L}\big[L^{2(t-\alpha)} - \log L \le T(2R_B, Y^L) \le L^{2(t-\alpha)}\big],\end{aligned}$$

which tends to zero by Lemma 3.1 applied with L replaced by $L^{1-\alpha}$ (the size of the torus on which Y^L evolves) if $t < 1$, and by (10) if $t = 1$. Lemma 3.4 enables us to deduce (a).

For (b), the same technique applies but with the last argument replaced by the use of Lemma B. □

Proposition 1.2 is, in fact, a particular case of a more general result which we shall use in Section 4.3 (with $k = 4$). Suppose we follow the ancestry at one locus of $k \ge 2$ different individuals. By analogy with above, we label individuals $1, \dots, k$, we write x^L_{ij} for the initial separation of lineages i and j, T^L_{ij} for the time at which their ancestral lineages first come within $2R_BL^\alpha$ and τ^L_{ij} for their coalescence time. We also write T^L_* (resp., τ^L_*) for the minimum over $\{i \ne j\}$ of the T^L_{ij}'s (resp., the τ^L_{ij}'s). Although (in the same way as above) we could state a result for a more general sequence $(a_L)_{L\ge1}$ of initial configurations, for the proof of

Theorem 1.4 we shall need some uniformity in the convergence. For this reason, we consider $\Gamma(L,k,\eta)$, the set of all configurations of k lineages on $\mathbb{T}(L)$ such that all pairwise distances $|x_{ij}^L|$ belong to $[L^\eta/(\log L), L^\eta \log L]$.

PROPOSITION 3.6. *For any $\beta \in (\alpha, 1]$, $\varepsilon > 0$ and $i \neq j$, we have*

$$\lim_{L\to\infty} \sup_{\beta\le\eta\le t\le 1} \sup_{a_L\in\Gamma(L,k,\eta)} \left| \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{ij}^L \le \rho_L L^{2(t-\alpha)}\right] - \frac{1}{\binom{k}{2}}\left(1 - \left(\frac{\eta-\alpha}{t-\alpha}\right)^{\binom{k}{2}}\right)\right| = 0,$$

$$\lim_{L\to\infty} \sup_{t\ge\varepsilon,\beta\le\eta\le 1} \sup_{a_L\in\Gamma(L,k,\eta)} \left| \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{ij}^L \le \frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)}\log Lt\right] - \frac{1}{\binom{k}{2}}\left(1 - \left(\frac{\eta-\alpha}{1-\alpha}e^{-t}\right)^{\binom{k}{2}}\right)\right| = 0.$$

The same is true with τ^L replaced by T^L.

In essence, Proposition 3.6 tells us that on the timescale $\rho_L L^{2(t-\alpha)}$, $t \in [\eta, 1]$, the time of the first coalescence (or of the first "gathering") is approximately the same as that of the first merger in a Kingman coalescent timechanged by $\log(\frac{t-\alpha}{\eta-\alpha})$, and that the approximation is uniform over η's bounded away from α. Moreover, asymptotically, just as in the Kingman coalescent, each pair of lineages has the same chance to be the first to coalesce. On the other hand, on the timescale $\frac{1-\alpha}{2\pi\sigma^2}\rho_L L^{2(1-\alpha)}\log Lt$, conditional on $T_*^L > \rho_L L^{2(1-\alpha)}$, the asymptotic behavior corresponds to Kingman's coalescent run at speed 1.

SKETCH OF PROOF. The proof of Proposition 3.6 is a straightforward adaptation of those of Lemma 4.2 and of Lemma 5.2 in [12] (see also the comments given in the paragraph following the proof of Lemma 4.2). The interested reader will also find there references to earlier results for the random walks with instantaneous coalescence which are dual to the two-dimensional voter model. □

Let us end this section by recalling a lemma of [1] and by stating an analogous result. For every $L \in \mathbb{N}$, $i \neq j$ and $t \ge 0$, let $\tilde{X}_{ij}^L(t)$ be the separation [on $\mathbb{T}(L)$ at time t] of lineages i and j.

LEMMA C (Lemma 6.9 in [1]). *Suppose $k = 4$ and*

$$\lim_{L\to\infty} \frac{\min_{i\neq j} \log|x_{ij}^L|}{\log L} = \lim_{L\to\infty} \frac{\max_{i\neq j} \log|x_{ij}^L|}{\log L} = 1. \tag{13}$$

Then,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{12}^L; |\tilde{X}_{13}^L(\tau_*^L)| \le \frac{L}{\log L}\right] = 0,$$

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{12}^L; |\tilde{X}_{34}^L(\tau_*^L)| \le \frac{L}{\log L}\right] = 0.$$

These results are also true if τ^L is replaced by T^L.

In words, when two lineages meet and coalesce, with probability tending to one the others are at distance at least $L/\log L$ of each other and of the coalescing pair (in particular, such a merger involves at most two lineages at a time). When the initial distance between the lineages is of the order of L^β with $\beta < 1$, we have instead:

LEMMA 3.7. *Suppose again $k = 4$ and the limit in (13) is equal to $\beta \in (\alpha, 1)$. Then*,

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{12}^L \le \frac{\rho_L L^{2(1-\alpha)}}{\log L}; |\tilde{X}_{13}^L(\tau_*^L)| \notin \left[\frac{L^\alpha}{\log L}\frac{\sqrt{\tau_*^L}}{\sqrt{\rho_L}}, L^\alpha \log L \frac{\sqrt{\tau_*^L}}{\sqrt{\rho_L}}\right]\right] = 0,$$

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[\tau_*^L = \tau_{12}^L \le \frac{\rho_L L^{2(1-\alpha)}}{\log L}; |\tilde{X}_{34}^L(\tau_*^L)| \notin \left[\frac{L^\alpha}{\log L}\frac{\sqrt{\tau_*^L}}{\sqrt{\rho_L}}, L^\alpha \log L \frac{\sqrt{\tau_*^L}}{\sqrt{\rho_L}}\right]\right] = 0.$$

The result is also true if τ^L is replaced by T^L.

Notice the rescalings of time by ρ_L and space by L^α introduced in (6) under which the behavior of the lineages is close to that of finite variance random walks. In fact, although their formulations are rather different, Lemma 3.7 is very similar to Lemma 1 in [6] or Lemma 5.1 in [12] for coalescing random walks.

SKETCH OF PROOF OF LEMMA 3.7. The method of proof is identical to that of Lemma 6.9 in [1], to which we refer for more complete arguments. It is based on two facts. First, by time $\rho_L L^{2(1-\alpha)}/(\log L)$ the separation of the lineages is never on the order of the side of the torus. Second, if $T_*^L = T_{12}^L$, then $L^{-\alpha}\tilde{X}_{13}^L(\rho_L \cdot)$ and $L^{-\alpha}\tilde{X}_{34}^L(\rho_L \cdot)$, considered separately, follow the same law as the difference of two independent lineages (on $\mathbb{R}^2$, by the first fact) conditioned on not entering $B(0, 2R_B)$ before T_*^L/ρ_L. By Lemma 3.4, with high probability $T_*^L/\rho_L \gg L^{2(\beta-\alpha)}$, and so the result for T^L follows from a standard central limit theorem.

The modifications needed for τ^L use the very rapid coalescence of two lineages gathered at distance $2R_BL^\alpha$ to obtain that, with probability tending to 1, if $\tau_*^L = \tau_{12}^L$, then no other pairs of lineages come within $2R_BL^\alpha$ of one another before time τ_*^L. An application of Lemma 3.7 (with T^L) completes the proof. $\square$

4. Genealogies at two loci. From now on, we work with the rescaling of time and space introduced in (6). As we saw in the previous section, these are the appropriate scales on which to understand the behavior of a collection of independent processes following the dynamics driven by (1). Because our lineages move independently as long as they are at distance greater than $2R_B$ (in rescaled units) of each other, it is also the relevant regime in which to understand "gathering" and coalescence of ancestral lineages.

The aim of Sections 4.1 and 4.2 is to understand how two lineages, initially present in the same individual, can "decorrelate" and how much time they need to do so. Once this phenomenon is understood for two lineages, we can consider the more complex situation described in the Section 1.4 and prove Theorems 1.4 and 1.5. This is achieved in Section 4.3.

4.1. *Effective recombination time.* For every L, let X_{AB}^L be the process that records the (rescaled) difference between the locations of the lineages labeled A and B. Recall that under our working assumptions, these lineages start within the same individual (in other words, A and B belong to the same block of the marked partition a_L).

By construction, recombination occurs only during small events. In our rescaled space and time units, a recombination event results in a separation of the lineages of $\mathcal{O}(L^{-\alpha})$, and then small events affect them at rate $\mathcal{O}(\rho_L)$. Hence, it is very likely that (in our rescaled time units) the lineages very rapidly coalesce and have to wait for the next recombination event [i.e., roughly $(\rho_L r_L)^{-1}$ units of rescaled time] to be geographically separated again, and so on. An efficient way for the lineages to escape this "flickering" due to small events is for a large event to send them to a separation of $\mathcal{O}(1)$. This necessarily occurs at a time when $X_{AB}^L \neq 0$. Thus, let us define S^L as the first time t at which at least one of the two lineages is affected by a large event and $X_{AB}^L(t-) \neq 0$ [which does not prohibit $X_{AB}^L(t) = 0$]. We call S^L the *effective recombination time*. Its large-L behavior is given by the following proposition.

PROPOSITION 4.1. *There exist $\theta_1, \theta_2 > 0$ such that for every $\theta > \theta_2$ and every nonvanishing sequence $(\phi_L)_{L\geq 1}$ satisfying $\phi_L \leq L^2/(\rho_L \log L)$ for every L, we have for L large enough*

$$\mathbb{P}_{a_L}\left[S^L \geq \phi_L\left(1 + \frac{\theta \log(\phi_L \rho_L)}{r_L \rho_L}\right)\right] \leq e^{-\theta_1 \phi_L} + e^{-(\theta-\theta_2)\phi_L \log(\phi_L \rho_L)}.$$

The idea of the proof of Proposition 4.1 is to show that, with very high probability, the number of visits to 0 of X^L_{AB} before it has accumulated a time ϕ_L outside 0 is less than $\phi_L \log(\phi_L \rho_L)$. Since each visit lasts a time proportional to $(r_L \rho_L)^{-1}$, the total amount of time it takes for X^L_{AB} to accumulate ϕ_L units of time outside zero is at most of the order of $\phi_L + \phi_L \log(\phi_L \rho_L)/(r_L \rho_L)$. The probability that by this time the two lineages have not been affected by a large event while in distinct locations is bounded by a quantity of the form $e^{-C\phi_L}$.

Let us write $\mathcal{R}_L(x)$ for the rate at which at least one of the lineages is affected by a large event when $X^L_{AB} = x$, and recall that time is rescaled by a factor ρ_L. From the expression for the intensity of Π^L_B, we can find a constant $C_B > 0$ such that $\mathcal{R}_L(x) \geq C_B$ for all $x \in \mathbb{T}(L^{1-\alpha}) \setminus \{0\}$ [in fact, one can even show that the function $x \mapsto \mathcal{R}_L(x)$ is increasing in $|x|$, and so one can take $C_B := \mathcal{R}_L(0) > 0$]. Let $\hat{X}^L$ be a $\mathbb{T}(L^{1-\alpha})$-valued Markov process distributed in the same way as the difference between two lineages subject only to the events of Π^L_s, and $\hat{S}^L$ be an exponential random variable with instantaneous rate $\mathcal{R}_L(\hat{X}^L(t))\mathbf{1}_{\{\hat{X}^L(t)\neq 0\}}$. By the preceding remark, $\hat{S}^L$ is stochastically bounded by an exponential random variable with instantaneous rate $C_B \mathbf{1}_{\{\hat{X}^L(t)\neq 0\}}$. Because large events have no effect when $X^L_{AB} = 0$, the law of the stopped process $\{X^L_{AB}(t), t \in [0, S^L]\}$ is the same as that of $\{\hat{X}^L(t), t \in [0, \hat{S}^L]\}$. Thus, for the proof of Proposition 4.1 we work with $\hat{X}^L$ and $\hat{S}^L$ and use P_x to denote the law of $\hat{X}^L$ under which $\mathbb{P}[\hat{X}^L(0) = x] = 1$.

For each $L \in \mathbb{N}$, let us define the stopping times $(\hat{Q}^L_i)_{i\geq 0}$ and $(\hat{q}^L_i)_{i\geq 0}$ by $\hat{Q}^L_0 = \hat{q}^L_0 = 0$ and for every $i \geq 1$,

$$\hat{Q}^L_i := \inf\{t \geq \hat{q}^L_{i-1} : \hat{X}^L(t) \neq 0\},$$
$$\hat{q}^L_i := \inf\{t \geq \hat{Q}^L_i : \hat{X}^L(t) = 0\}.$$

[Note that $\hat{Q}^L_1 = 0$ if $\hat{X}^L(0) \neq 0$, in which case $\hat{q}^L_1$ is the first hitting time of 0.] By construction, the random variables $(\hat{Q}^L_i - \hat{q}^L_{i-1})_{i\in\mathbb{N}}$ are i.i.d. and distributed according to an exponential random variable with parameter $C_{\mathrm{rec}} r_L \rho_L$, where $C_{\mathrm{rec}} := \pi R_s^2 u_s (1 - \lambda_s(\{1\})) > 0$ (the last factor arises since the number of reproducing individuals needs to be greater than one for recombination to occur). We have the following result for the excursions of $\hat{X}^L$ away from 0.

LEMMA 4.2. *There exist $C_e > 0$ and $u_e > 0$ such that for every $L \geq 1$ and $u_e \leq u \leq L^2/(\log L)$, for every $x \in B(0, 2R_s L^{-\alpha}) \setminus \{0\}$,*

$$\mathrm{P}_x[\hat{q}^L_1 > u\rho_L^{-1}] \geq \frac{C_e}{\log u}.$$

PROOF. Here (and only here) it is easier to work with the initial time and space units and show that the probability of an excursion outside 0 of length greater than u is bounded from below by $C_e/(\log u)$ when u is large. Let us thus define

$\tilde{X}^L$ by $\tilde{X}^L(t) := L^\alpha \hat{X}^L(\rho_L^{-1} t)$ for all $t \geq 0$, with the understanding that $\tilde{X}^L$ starts at $L^\alpha x$ under the probability measure P_x.

The desired result is shown in [11] for standard discrete space random walks whose jumps have finite variance as well as for Brownian motion (with the hitting time of 0 replaced by the entrance time into a ball of fixed radius) in two dimensions. To see why it is true for $\tilde{X}^L$ on $\mathbb{T}(L)$, observe first that by time $L^2/(\log L)$, the process $\tilde{X}^L$ does not see that space is limited, and so it behaves as though it were moving in $\mathbb{R}^2$. More precisely, there exists a constant $C > 0$ such that for all $z \in B(0, 6R_s L^{-\alpha})$,

$$\mathrm{P}_z\Big[\sup_{u \leq L^2/(\log L)} |\tilde{X}^L(u)| > \frac{L}{3}\Big] \leq \frac{C}{\log L}.$$

(Use the L^2-maximal inequality and the fact that $|\tilde{X}^L|$ is bounded by the corresponding quantity for the same process defined on $\mathbb{R}^2$, which is proportional to $L^2/(\log L)$ by equation (22) in [1]). Hence, let us assume that $\tilde{X}^L$ is defined on $\mathbb{R}^2$ instead of $\mathbb{T}(L)$. Since the evolution due to small events depends on L only through the torus sidelength, with our new convention all $\tilde{X}^L$'s have the same distribution and we can drop the exponent L in the notation. For the same reason, we also write $\tilde{q}_1$ for the random times $\rho_L \hat{q}_1^L$, that is, the length of the first excursion outside 0 of $\tilde{X}$.

Let $\tilde{\mathrm{T}}_{(4R_s)}$ denote the first time $\tilde{X}$ leaves $B(0, 4R_s)$ [and so $\tilde{X}(\tilde{\mathrm{T}}_{(4R_s)}) \in B(0, 6R_s) \setminus B(0, 4R_s)$ by our assumption on the jump sizes], and let $\tilde{\mathrm{T}}_{[2R_s]}$ be the first return time of $\tilde{X}$ into $B(0, 2R_s)$ after $\tilde{\mathrm{T}}_{(4R_s)}$. We have for every $x \in B(0, 2R_s L^{-\alpha}) \setminus \{0\}$,

$$\begin{aligned}
&\mathrm{P}_x[\tilde{q}_1 > u] \\
&\quad \geq \mathrm{P}_x\big[\tilde{q}_1 > u; \tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1\big] \\
&\quad \geq \mathrm{P}_x\big[\tilde{q}_1 - \tilde{\mathrm{T}}_{(4R_s)} > u; \tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1\big] \\
&\quad = \mathrm{E}_x\Big[\mathbf{1}_{\{\tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1\}} \mathrm{P}_{\tilde{X}(\tilde{\mathrm{T}}_{(4R_s)})}[\tilde{q}_1 > u]\Big] \\
&\quad \geq \mathrm{E}_x\big[\mathbf{1}_{\{\tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1\}} \mathrm{P}_{\tilde{X}(\tilde{\mathrm{T}}_{(4R_s)})}\big[\tilde{\mathrm{T}}_{[2R_s]} > u\big]\big] \\
&\quad \geq \Big(\inf_{B(0,2R_s)\setminus\{0\}} \mathrm{P}_{L^{-\alpha}y}\big[\tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1\big]\Big)\Big(\inf_{B(0,6R_s)\setminus B(0,4R_s)} \mathrm{P}_{L^{-\alpha}z}\big[\tilde{\mathrm{T}}_{[2R_s]} > u\big]\Big).
\end{aligned} \tag{14}$$

The first infimum is strictly positive. To see this, note that $\mathrm{P}_{L^{-\alpha}y}[\tilde{\mathrm{T}}_{(4R_s)} < \tilde{q}_1]$ is bounded from below by the probability that the first four small events affecting the lineages send them to a distance at least $4R_s$ of each other before they coalesce, and the infimum over $B(0, 2R_s) \setminus \{0\}$ of the latter probability is positive since $u_s < 1$ (if $u_s = 1$, only one of the lineages can be in the geographical range of such separating events, and so their probability of occurrence shrinks to 0 as $|y| \to 0$).

For the second infimum in (14), we use the same construction as in the proof of Skorokhod embedding (see, e.g., [3]) to write the path of $\tilde{X}$ as that of a standard Brownian motion W considered at particular times. More precisely, if $(\tilde{\sigma}_i)_{i\in\mathbb{N}}$ is the sequence of jump times of $\tilde{X}$, we can find a sequence of Brownian stopping times $(\sigma_i)_{i\in\mathbb{N}}$ such that $(W(\sigma_i))_{i\geq 0}$ has the same joint distributions as $(\tilde{X}(\tilde{\sigma}_i))_{i\geq 0}$. For every $i \in \mathbb{N}$, conditional on $W(\sigma_{i-1})$, σ_i is the first time greater than σ_{i-1} at which W leaves $B(W(\sigma_{i-1}), l_i)$, where the random variable l_i is independent of W and of $\{\sigma_j, j < i\}$ and has the same distribution as the length of the first jump of $\tilde{X}$. As a consequence, if $\tilde{n}(u) := \max\{i : \tilde{\sigma}_i \leq u\}$, by comparing the paths of $\tilde{X}$ and of W we obtain

$$\mathrm{P}_{L^{-\alpha}z}\big[\tilde{\mathrm{T}}_{[2R_s]} > u\big] \geq \mathrm{P}_z\big[W(t) \notin B(0, 2R_s), \forall t \leq \sigma_{\tilde{n}(u)}\big].$$

Now, each $\tilde{\sigma}_i - \tilde{\sigma}_{i-1}$ is stochastically bounded from below by an exponential random variable with positive parameter $k_1 > 0$, and so by standard large deviation results we can find $k_2 > 0$ large enough and $k_3 > 0$ such that for all $u > 1$ and $y \in \mathbb{R}^2$,

$$\mathrm{P}_y[\tilde{n}(u) > k_2 u] \leq e^{-k_3 u}.$$

By construction, each $\sigma_i - \sigma_{i-1}$ is stochastically bounded from above by the first time Brownian motion started at 0 leaves $B(0, 2R_s)$, which also has an exponential moment. Hence, there exist $k_4, k_5 > 0$ such that for all $u > 1$ and $y \in \mathbb{R}^2$,

$$\mathrm{P}_y\big[\sigma_{\lfloor k_2 u\rfloor+1} > k_4 u\big] \leq e^{-k_5 u}.$$

Using these bounds and the result already established in [11] for Brownian motion at time $k_4 u$, Lemma 4.2 is proved. □

We now have all the ingredients we require to prove Proposition 4.1.

PROOF OF PROPOSITION 4.1. Set

$$\psi_L := \phi_L\left(1 + \frac{\theta \log(\phi_L \rho_L)}{r_L \rho_L}\right), \tag{15}$$

and call $\mathbf{t}(\psi_L)$ the time $\hat{X}^L$ spends away from 0 before time ψ_L. We have, for every L,

$$\begin{aligned}\mathrm{P}_0[\hat{S}^L \geq \psi_L] &= \mathrm{P}_0[\hat{S}^L \geq \psi_L; \mathbf{t}(\psi_L) \leq \phi_L] + \mathrm{P}_0[\hat{S}^L \geq \psi_L; \mathbf{t}(\psi_L) > \phi_L] \\ &\leq \mathrm{P}_0[\mathbf{t}(\psi_L) \leq \phi_L] + e^{-C_B \phi_L},\end{aligned}$$

where C_B is the lower bound on the rate of effective large events introduced just below the statement of the proposition. Next, if we set $\hat{k}_L := \sup\{i : \hat{Q}_i^L \leq \psi_L\}$, that is, $\hat{k}_L$ is the number of excursions of $\hat{X}^L$ away from 0 which start before time ψ_L, we can write

$$\begin{aligned}\mathrm{P}_0[\mathbf{t}(\psi_L) \leq \phi_L] = \mathrm{P}_0[&\mathbf{t}(\psi_L) \leq \phi_L; \hat{k}_L \leq \phi_L \log(\phi_L \rho_L)] \\ &+ \mathrm{P}_0[\mathbf{t}(\psi_L) \leq \phi_L; \hat{k}_L > \phi_L \log(\phi_L \rho_L)].\end{aligned}$$

On the one hand,

$$\begin{aligned}
\mathrm{P}_0[\mathbf{t}(\psi_L) \le \phi_L; \hat{k}_L > \phi_L \log(\phi_L \rho_L)] &\le \mathrm{P}_0\left[\sum_{i=1}^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor} (\hat{q}_i^L - \hat{Q}_i^L) \le \phi_L\right] \\
&\le \mathrm{P}_0[\hat{q}_1^L - \hat{Q}_1^L \le \phi_L]^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor} \\
&\le \left(1 - \frac{C_e}{\log(\phi_L \rho_L)}\right)^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor} \\
&\le e^{-C'_e \phi_L},
\end{aligned}$$

for a constant $C'_e > 0$ and L large enough. The second line is obtained by an obvious recursion using the strong Markov property at the successive times $\hat{q}_i^L$ in decreasing order, and the third line uses Lemma 4.2 [recall that by assumption on ϕ_L, we have $\phi_L \rho_L \to \infty$ and $\phi_L \rho_L \le L^2/(\log L)$]. Hence, we can set $\theta_1 := C_B \wedge C'_e$. On the other hand,

$$\begin{aligned}
&\mathrm{P}_0[\mathbf{t}(\psi_L) \le \phi_L; \hat{k}_L \le \phi_L \log(\phi_L \rho_L)] \\
&\quad \le \mathrm{P}_0\left[\sum_{i=1}^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor + 1} (\hat{Q}_i^L - \hat{q}_{i-1}^L) \ge \psi_L - \phi_L\right] \\
&\quad = \mathrm{P}_0\left[\exp\left\{r_L \rho_L \sum_{i=1}^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor + 1} (\hat{Q}_i^L - \hat{q}_{i-1}^L)\right\} \ge \exp\{r_L \rho_L (\psi_L - \phi_L)\}\right] \\
&\quad \le e^{-\theta \phi_L \log(\phi_L \rho_L)} \mathrm{E}_0\left[\exp\left\{r_L \rho_L \sum_{i=1}^{\lfloor \phi_L \log(\phi_L \rho_L)\rfloor + 1} (\hat{Q}_i^L - \hat{q}_{i-1}^L)\right\}\right],
\end{aligned}$$

where the last line uses the Markov inequality. As we pointed out above, the random variables $r_L \rho_L (\hat{Q}_i^L - \hat{q}_{i-1}^L)$ are i.i.d. with law $\mathrm{Exp}(C_{\mathrm{rec}})$. Therefore, we can write for a constant $\theta_2 > 0$

$$\mathrm{P}_0[\mathbf{t}(\psi_L) \le \phi_L; \hat{k}_L \le \phi_L \log(\phi_L \rho_L)] \le e^{-(\theta - \theta_2)\phi_L \log(\phi_L \rho_L)}.$$

Combining these results, the proof of Proposition 4.1 is complete. □

Finally, let us use Proposition 4.1 to obtain some estimates on the time two lineages starting in the same individual need to reach a separation at which they start to evolve independently. The following lemma will be a key result for the proof of Proposition 4.4 in the next section. For every $L \in \mathbb{N}$, let $\mathrm{T}^L_{(3R_B)}$ denote the exit time of X^L_{AB} from $B(0, 3R_B)$.

LEMMA 4.3. *There exists a constant $\theta_3 > 0$ such that if $(\phi_L)_{L \ge 1}$ is as in Proposition 4.1, $\phi_L \to \infty$ as $L \to \infty$ and $\theta > \theta_2$, there exists $L_0 = L_0(\theta,$*

$(\phi_L)_{L\in\mathbb{N}}$) *such that for every* $L \geq L_0$,

$$\mathbb{P}_{a_L}\left[\mathrm{T}^L_{(3R_B)} \geq \phi_L\left(1+\frac{\theta\log(\phi_L\rho_L)}{r_L\rho_L}\right)\right] \leq \sqrt{\phi_L}e^{-\theta_3\sqrt{\phi_L}}.$$

PROOF. For conciseness, we again use the notation ψ_L introduced in (15). This time we define $Q^L_0 = q^L_0 = 0$ and

$$\begin{aligned}
Q^L_i &:= \inf\{t > q^L_{i-1} : t \text{ is the epoch of an effective recombination}\},\\
q^L_i &:= \inf\{t \geq Q^L_i : X^L_{AB}(t) = 0 \text{ or } X^L_{AB}(t) \notin B(0,3R_B)\},\\
k_L &:= \max\{i : Q^L_i \leq \mathrm{T}^L_{(3R_B)}\}.
\end{aligned}$$

First, we claim that there exists a constant $\wp > 0$ independent of L such that, for L large enough, $k_L + 1$ is stochastically bounded by a geometric random variable with success probability $\wp$. In other words, the probability that X^L_{AB} starting at $x \in B(0,3R_B)\setminus\{0\}$ leaves $B(0,3R_B)$ before hitting 0 is bounded from below by $\wp$, independently of x. The proof of this claim is given in the first paragraph of the proof of Lemma 6.6 in [1]. (The quantity $\wp$ is taken to be the probability that a sequence of large events sends the lineages to a distance of at least $3R_B$ without meanwhile being counteracted by small events bringing them too close together.) As a consequence, for any large L,

$$\mathbb{P}_{a_L}[\mathrm{T}^L_{(3R_B)} \geq \psi_L] \leq \mathbb{P}_{a_L}[\mathrm{T}^L_{(3R_B)} \geq \psi_L; k_L < \sqrt{\phi_L}] + (1-\wp)^{\sqrt{\phi_L}}.$$

Next, let us write

$$\begin{aligned}
&\mathbb{P}_{a_L}[\mathrm{T}^L_{(3R_B)} \geq \psi_L; k_L < \sqrt{\phi_L}] \\
&\quad= \mathbb{P}_{a_L}\left[\mathrm{T}^L_{(3R_B)} \geq \psi_L; k_L < \sqrt{\phi_L}; \sum_{i=1}^{k_L}(Q^L_i - q^L_{i-1}) \geq \frac{\psi_L}{2}\right] \qquad (16)\\
&\qquad+ \mathbb{P}_{a_L}\left[\mathrm{T}^L_{(3R_B)} \geq \psi_L; k_L < \sqrt{\phi_L}; \sum_{i=1}^{k_L}(Q^L_i - q^L_{i-1}) < \frac{\psi_L}{2}\right]. \qquad (17)
\end{aligned}$$

The quantity in (16) is bounded by

$$\begin{aligned}
\mathbb{P}_{a_L}\left[\sum_{i=1}^{\lfloor\sqrt{\phi_L}\rfloor}(Q^L_i - q^L_{i-1}) \geq \frac{\psi_L}{2}\right] &= 1 - \mathbb{P}_{a_L}\left[\sum_{i=1}^{\lfloor\sqrt{\phi_L}\rfloor}(Q^L_i - q^L_{i-1}) < \frac{\psi_L}{2}\right]\\
&\leq 1 - \mathbb{P}_{a_L}\left[\forall i \leq \lfloor\sqrt{\phi_L}\rfloor, Q^L_i - q^L_{i-1} < \frac{\psi_L}{2\sqrt{\phi_L}}\right] \qquad (18)\\
&\leq 1 - \left(1 - \sup_{a'_L}\mathbb{P}_{a'_L}\left[S^L \geq \frac{\psi_L}{2\sqrt{\phi_L}}\right]\right)^{\lfloor\sqrt{\phi_L}\rfloor},
\end{aligned}$$

where the last line is obtained by recursion (notice that, conditionally on q_{i-1}^L, $Q_i^L - q_{i-1}^L$ has the same law as the effective recombination time S_L) and the supremum is taken over all initial configurations a_L' in which lineages A and B are either at distance 0 or at distance greater than $3R_B$. We can in fact restrict our attention to the set of configurations in which A and B belong to the same block. Indeed, if $|X_{AB}^L(0)| > 3R_B$, we can decompose the probability that $S^L \geq \psi_L/(2\sqrt{\phi_L})$ into the sum of:

- the probability that $S^L \geq \psi_L/(2\sqrt{\phi_L})$ and X_{AB}^L does not hit 0 before time $\psi_L/(4\sqrt{\phi_L})$, which decreases like $e^{-C\psi_L/\sqrt{\phi_L}}$ since the rate at which large events affect the lineages when $X_{AB}^L \neq 0$ is bounded from below by a positive constant;
- the probability that $S^L \geq \psi_L/(2\sqrt{\phi_L})$ and X_{AB}^L hits 0 before time $\psi_L/(4\sqrt{\phi_L})$, which boils down to the case $X_{AB}^L(0) = 0$ by the strong Markov property applied at the first time $X_{AB}^L = 0$.

Now, by Proposition 4.1 applied with ϕ_L replaced by $\sqrt{\phi_L}/2$, we have

$$\begin{aligned}\mathbb{P}_{a_L'}\left[S^L \geq \frac{\psi_L}{2\sqrt{\phi_L}}\right] &\leq \mathbb{P}_{a_L'}\left[S^L \geq \frac{\sqrt{\phi_L}}{2}\left(1 + \frac{\theta \log(\sqrt{\phi_L}\rho_L/2)}{r_L\rho_L}\right)\right] \\ &\leq e^{-(\theta_1/2)\sqrt{\phi_L}} + e^{-((\theta-\theta_2)/2)\sqrt{\phi_L}\log(\sqrt{\phi_L}\rho_L/2)}.\end{aligned}$$

Substituting in (18) and using the asymptotic relation $1 - (1 - e^{-t})^t \sim te^{-t}$ as $t \to \infty$, we obtain that for L large enough, the quantity in (16) is bounded by $\sqrt{\phi_L}e^{-(\theta_1/4)\sqrt{\phi_L}}$.

As concerns (17), observe that there exists $\theta_4 > 0$ such that for every $L \geq 1$, each of the $q_i^L - Q_i^L$ is stochastically bounded by an exponential random variable with parameter θ_4. Indeed, when X_{AB}^L lies within $B(0, (3/2)R_B)$, the rate at which a coalescence occurs due to a large event is bounded from below by a positive constant. On the other hand, it is not difficult to check that when X_{AB}^L lies within $B(0, (3/2)R_B)^c$, the rate at which the two lineages are sent at a distance greater than $3R_B$ by a large event is also bounded from below by a positive constant. The quantity in (17) is therefore bounded by

$$\mathbb{P}_{a_L}\left[\sum_{i=1}^{\lfloor\sqrt{\phi_L}\rfloor}(q_i^L - Q_i^L) \geq \frac{\psi_L}{2}\right] \leq \mathbb{P}\left[\sum_{i=1}^{\lfloor\sqrt{\phi_L}\rfloor}\mathcal{E}_i \geq \frac{\psi_L}{2}\right] \leq \exp\left\{-\frac{\psi_L}{2} + c\sqrt{\phi_L}\right\},$$

where $(\mathcal{E}_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. exponential random variables with parameter θ_4 and c is a positive constant expressed in terms of the exponential moment of $\mathcal{E}_1$. The result follows. □

4.2. *Decorrelation time of two lineages starting in the same individual.* In the previous section we obtained some information on the time required for two

lineages starting in the same individual to become separated by a distance greater than $3R_B$. We know that the lineages behave independently whenever they are at distance greater than $2R_B$. However, nothing guarantees that after the random time $\mathrm{T}^L_{(3R_B)}$ of Lemma 4.3, the ancestral lineages of A and B will evolve independently. Indeed, it is very likely that after some time they will once again be within distance $2R_B$ of one another and coalescence events will keep them close together for a potentially long period of time. Hence, in order to prove Theorem 1.4, we would like to know how much time our lineages need before they start "looking" as if they were independent. That is, we are interested in the time until their separation is of the same order as if they had evolved according to independent copies of ℓ^L started from 0. Recall from Lemma A that for (large) times less than $L^{2(1-\alpha)}/\sqrt{\log L}$, the difference of two independent lineages behaves like Brownian motion on $\mathbb{R}^2$. The following proposition thus tells us that the decorrelation time we are looking for is asymptotically bounded from above by $(\log L)^5(1+\frac{\log \rho_L}{r_L\rho_L})$.

PROPOSITION 4.4. *Let $(T_L)_{L\geq 1}$ be a sequence of times such that $(\log L)^5(1+\frac{\log \rho_L}{r_L\rho_L}) \leq T_L \leq \frac{L^{2(1-\alpha)}}{\log L}$ for every L. Then,*

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\left[|X^L_{AB}(T_L)| \notin \left[\frac{\sqrt{T_L}}{\log L}, \sqrt{T_L}\log L\right]\right] = 0.$$

The scheme of the proof of Proposition 4.4 will again be to decompose the path of X^L_{AB} into appropriate excursions and incursions. We shall show that the proportion of the time before T_L that X^L_{AB} spends in the region of space where it does *not* evolve like the difference of two independent lineages is asymptotically negligible.

To this end, for every $L \in \mathbb{N}$, let us define the stopping times $(Q^L_i)_{i\geq 0}$ and $(q^L_i)_{i\geq 0}$ by $q^L_0 = Q^L_0 = 0$, and for every $i \geq 1$,

$$Q^L_i := \inf\{t > q^L_{i-1} : X^L_{AB}(t) \notin B(0, 3R_B)\},$$
$$q^L_i := \inf\{t > Q^L_i : X^L_{AB}(t) \in B(0, 2R_B)\},$$

with the convention that $\inf \varnothing = +\infty$. We also write k_L for the number of "excursions" that start before time T_L, that is,

$$k_L := \max\{i : Q^L_i \leq T_L\}.$$

The first step in proving Proposition 4.4 is to show that

LEMMA 4.5. *For every $\delta \in (0, 1/2)$, there exist $K(\delta) > 0$ such that for all L large enough,*

$$\mathbb{P}_{a_L}[k_L > K(\delta)\log T_L] \leq \delta.$$

We postpone the proof of Lemma 4.5 until the end of the section and instead exploit it to prove Proposition 4.4.

PROOF OF PROPOSITION 4.4. We construct a coupling between X^L_{AB} and a compound Poisson process Y^L which evolves as the difference between two *independent* copies of ℓ^L. Define Y^L as follows: during an excursion of X^L_{AB}, Y^L makes the same jumps as X^L_{AB} at the same times, that is,

$$\forall i \geq 1, \forall t \in (Q^L_i, q^L_i], \qquad Y^L(t) - Y^L(t-) = X^L_{AB}(t) - X^L_{AB}(t-).$$

During the remaining time, Y^L jumps independently of X^L_{AB} with a jump intensity equal to twice that given in (1) rescaled in an appropriate manner. It is easy to check that the law of Y^L is indeed as claimed, since outside $B(0, 2R_B)$, X^L_{AB} evolves like the difference of two independent lineages and so the jump intensity corresponding to the process Y^L is equal to twice that in the rescaled version of (1) at any time. Furthermore, by construction, the difference between X^L_{AB} and Y^L changes only during the time intervals $[q^L_{i-1}, Q^L_i]$. For convenience, we retain the notation $\mathbb{P}$ for the probability measures on the (larger) space of definition of the pair (X^L_{AB}, Y^L), and set $Y^L(0) = 0$, $\mathbb{P}_{a_L}$-a.s.

Let us call I_L the amount of time before T_L during which X^L_{AB} and Y^L behave independently, that is,

$$I_L := \sum_{i=1}^{k_L} (Q^L_i - q^L_{i-1}) + (T_L - q^L_{k_L})_+.$$

If θ_2 is as in Proposition 4.1, we have

$$\begin{aligned}
&\mathbb{P}_{a_L}\left[|X^L_{AB}(T_L)| \notin \left[\frac{\sqrt{T_L}}{\log L}, \sqrt{T_L}\log L\right]\right] \\
&\quad \leq \mathbb{P}_{a_L}\left[I_L < (\log L)^2\left(1 + \frac{2\theta_2 \log(\rho_L \log L)}{r_L \rho_L}\right);\right. \\
&\qquad\qquad \left. |X^L_{AB}(T_L)| \notin \left[\frac{\sqrt{T_L}}{\log L}, \sqrt{T_L}\log L\right]\right] \\
&\quad + \mathbb{P}_{a_L}\left[I_L \geq (\log L)^2\left(1 + \frac{2\theta_2 \log(\rho_L \log L)}{r_L \rho_L}\right)\right].
\end{aligned} \tag{19}$$

First, let us show that the second term in the right-hand side of (19) converges to 0 as $L \to \infty$. Let $\delta \in (0, 1/2)$. By Lemma 4.5, there exists $K > 1$ such that for L large enough, $\mathbb{P}_{a_L}[k_L > K \log T_L] \leq \delta$. Hence, we can write

$$\begin{aligned}
&\mathbb{P}_{a_L}\left[I_L \geq (\log L)^2\left(1 + \frac{2\theta_2 \log(\rho_L \log L)}{r_L \rho_L}\right)\right] \\
&\quad \leq \mathbb{P}_{a_L}\left[I_L \geq (\log L)^2\left(1 + \frac{2\theta_2 \log(\rho_L \log L)}{r_L \rho_L}\right); k_L \leq K \log T_L\right] + \delta.
\end{aligned}$$

Now, by the same reasoning as in (18), we have

$$
\begin{aligned}
&\mathbb{P}_{a_L}\Big[I_L \geq (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big); k_L \leq K\log T_L\Big]\\
&\quad\leq \mathbb{P}_{a_L}\Big[\sum_{i=1}^{\lfloor K\log T_L\rfloor+1} Q_i^L - q_{i-1}^L \geq (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big)\Big]\\
&\quad\leq 1-\Big(1-\sup_{a_L'}\mathbb{P}_{a_L'}\Big[Q_1^L \geq \frac{(\log L)^2}{K\log T_L+1}\\
&\qquad\qquad\times\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big)\Big]\Big)^{\lfloor K\log T_L\rfloor+1},
\end{aligned}
\tag{20}
$$

where the supremum is taken over all initial configurations a_L' in which the distance between the blocks containing A and B is at most $2R_B$. Again, as in (18), we can restrict our attention to initial configurations in which A and B belong to the same block (recall from the proof of Lemma 4.3 that the rate at which a sequence of "separating" events occurs is bounded from below by a positive constant whenever $X_{AB}^L \neq 0$). By assumption, $\log T_L \leq 2\log L$ and $K > 1$, and so using Lemma 4.3 with $\phi_L = (\log L)/(2K)$ for the last inequality we obtain that for all large L, uniformly in a_L' as above,

$$
\begin{aligned}
&\mathbb{P}_{a_L'}\Big[Q_1^L \geq \frac{(\log L)^2}{K\log T_L+1}\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big)\Big]\\
&\quad\leq \mathbb{P}_{a_L'}\Big[Q_1^L \geq \frac{\log L}{2K}\Big(1+\frac{2\theta_2\log((2K)^{-1}\rho_L\log L)}{r_L\rho_L}\Big)\Big]\\
&\quad\leq \sqrt{(2K)^{-1}\log L}\,e^{-\theta_3\sqrt{(2K)^{-1}\log L}}.
\end{aligned}
$$

Consequently, we obtain from the asymptotic relation $1-(1-te^{-t})^{t^2} \sim t^3e^{-t}$ that the quantity in the right-hand side of (20) tends to zero as $L\to\infty$ and

$$
\limsup_{L\to\infty}\mathbb{P}_{a_L}\Big[I_L \geq (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big)\Big] \leq \delta.
$$

Since δ was arbitrary, this limit is actually zero.

Let us now show that the first term in the right-hand side of (19) tends to zero as $L\to\infty$. To this end, observe that it is bounded by

$$
\begin{aligned}
&\mathbb{P}_{a_L}\Big[I_L < (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big);\\
&\quad |X_{AB}^L(T_L) - Y^L(T_L)| > (\log\log L)(\log L)\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2}\Big]
\end{aligned}
\tag{21}
$$

$$+\mathbb{P}_{a_L}\Bigg[I_L < (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big);$$
$$|X^L_{AB}(T_L)| \notin \Big[\frac{\sqrt{T_L}}{\log L}, \sqrt{T_L}\log L\Big];$$
$$|X^L_{AB}(T_L) - Y^L(T_L)| \le (\log\log L)(\log L)$$
$$\times\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2}\Bigg].$$

Because the difference $X^L_{AB} - Y^L$ changes only during the periods $[q^L_{i-1}, Q^L_i]$, during which $|X^L_{AB}| \le 3R_B$ and Y^L jumps around according to twice the jump intensity given by the appropriate rescaling of (1), the first term in (21) is bounded by

$$\mathbb{P}_{a_L}\Bigg[I_L < (\log L)^2\Big(1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big);$$
$$|\hat{Y}^L(I_L)| + 3R_B > (\log\log L)(\log L)\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2}\Bigg],$$

where $\hat{Y}^L$ is an independent copy of Y^L starting from 0. Hence, we also have as an upper bound

$$\mathbb{P}_{a_L}[|\hat{Y}^L(I_L)| > (\log\log L)\sqrt{I_L} - 3R_B],$$

which tends to zero by a standard use of Markov's inequality and equation (22) of [1].

As concerns the second term in (21), it is bounded by

$$\mathbb{P}_{a_L}\Bigg[|Y^L(T_L)| \notin \Big[\frac{\sqrt{T_L}}{\log L} + (\log\log L)(\log L)\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2},$$
$$\sqrt{T_L}\log L - (\log\log L)(\log L)\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2}\Big]\Bigg]$$
$$= \mathbb{P}_{a_L}\Bigg[|Y^L(T_L)| \notin \Big[\frac{\sqrt{T_L}}{\log L}\big(1+\varepsilon^{(1)}_L\big), \sqrt{T_L}\log L\big(1-\varepsilon^{(2)}_L\big)\Big]\Bigg],$$

where by assumption on T_L and the fact that $\rho_L \ge \log L$,

$$\varepsilon^{(1)}_L := \frac{(\log L)^2\log\log L}{\sqrt{T_L}}\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2} \le C\frac{\log\log L}{\sqrt{\log L}}$$

and

$$\varepsilon^{(2)}_L := \frac{\log\log L}{\sqrt{T_L}}\Big\{1+\frac{2\theta_2\log(\rho_L\log L)}{r_L\rho_L}\Big\}^{1/2} \le C'\frac{\log\log L}{(\log L)^{5/2}}.$$

An application of the central limit theorem then gives the result. □

The proof of Lemma 4.5 rests upon the following lemma.

LEMMA 4.6. *There exists $C_q, v_q > 0$ such that for every L large enough, $v_q \le v \le L^{2(1-\alpha)}/(\log L)$ and every initial condition a'_L in which the separation between A and B belongs to $B(0, 5R_B) \setminus B(0, 3R_B)$,*

$$\mathbb{P}_{a'_L}[q_1^L > v] \ge \frac{C_q}{\log v}.$$

The proof of Lemma 4.6 uses the same arguments as the second half of the proof of Lemma 4.2 (based on Skorokhod embedding) and so we omit it.

PROOF OF LEMMA 4.5. Our strategy is to show that if we choose K large enough, the probability that none of the first $K \log T_L$ excursions outside $B(0, 3R_B)$ has duration of $\mathcal{O}(T_L)$ is smaller than δ. To achieve this, let $K > 0$. We have

$$\begin{aligned}
\mathbb{P}_{a_L}[k_L > K \log T_L] &= \mathbb{P}_{a_L}\big[Q^L_{\lfloor K \log T_L \rfloor + 1} \le T_L\big] \\
&= \mathbb{P}_{a_L}\left[\sum_{i=1}^{\lfloor K \log T_L \rfloor} (q_i^L - Q_i^L) + \sum_{i=1}^{\lfloor K \log T_L \rfloor + 1} (Q_i^L - q_{i-1}^L) \le T_L\right] \\
&\le \mathbb{P}_{a_L}\left[\sum_{i=1}^{\lfloor K \log T_L \rfloor} (q_i^L - Q_i^L) \le T_L\right] \\
&\le \mathbb{P}_{a_L}[\forall i \in \{1, \dots, \lfloor K \log T_L \rfloor\}, q_i^L - Q_i^L \le T_L].
\end{aligned}$$

Using a recursion and Lemma 4.6 together with the fact that $|X^L_{AB}(Q_i^L)| \in [3R_B, 5R_B]$ (recall the jump lengths are bounded by $2R_B$), we arrive at

$$\begin{aligned}
\mathbb{P}_{a_L}[\forall i \in \{1, \dots, \lfloor K \log T_L \rfloor\}, q_i^L - Q_i^L \le T_L] &\le \left(1 - \frac{C_q}{\log T_L}\right)^{\lfloor K \log T_L \rfloor} \\
&\to e^{-KC_q} \qquad \text{as } L \to \infty.
\end{aligned}$$

Now choose $K(\delta)$ large enough that $e^{-K(\delta)C_q} \le \delta/2$, and Lemma 4.5 is proved. □

4.3. *Proof of the main results.* Now that we understand decorrelation better, we can prove Theorems 1.4 and 1.5. Recall the rescalings of time by a factor ρ_L and of space by $L^{-\alpha}$ that have been in force since the beginning of Section 4 and the notation τ_{ij}^L for the coalescence time of lineages i and j in *original*

units. In order to work in the rescaled setting, we define $\mathrm{t}_{ij}^L := \tau_{ij}^L/\rho_L$ for every $i, j \in \{A, a, B, b\}$, and $\mathrm{t}^L := \mathrm{t}_{Aa}^L \wedge \mathrm{t}_{Bb}^L$. We denote the genealogical process (on the original space and time scales) of the four loci corresponding to step L by $\mathcal{A}^L$. As explained in Section 1.4, this Markov process takes its values in the set of all marked partitions of $\{A, a, B, b\}$. For any $t \geq 0$, each block of $\mathcal{A}^L(t)$ contains the labels of the lineages present in the same individual at (genealogical) time t, and its mark gives the current location on $\mathbb{T}(L)$ of this common ancestor.

REMARK 4.7. Several times during the course of the proofs below we shall apply Proposition 4.4 with $T_L = L^{2(\beta-\alpha)}$. Strictly speaking, we can only do this if $L^{2(\beta-\alpha)} \geq (\log L)^5(1 + \frac{\log \rho_L}{r_L \rho_L})$, at least for L large enough, which is not guaranteed by (2). However, if it is not the case, we can still find a sequence $(\phi_L)_{L\in\mathbb{N}}$ tending to infinity and such that

$$\phi_L L^{2(\beta-\alpha)} \geq (\log L)^5\left(1 + \frac{\log \rho_L}{r_L \rho_L}\right) \qquad \forall L \in \mathbb{N} \quad \text{and} \quad \lim_{L\to\infty} \frac{\log \phi_L}{\log L} = 0.$$

Now, for the sake of clarity we presented the results of Lemma 3.4 at times of the form $\rho_L L^{2(t-\alpha)}$ but its proof shows that, because $\log(\phi_L L^{2(\beta-\alpha)}) \sim \log(L^{2(\beta-\alpha)})$ as $L \to \infty$, we also have

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\big[T_{Aa}^L > \rho_L \phi_L L^{2(\beta-\alpha)}\big] = 1.$$

(Another way to see this is to use the inequality $\mathbb{P}_{a_L}[T_{Aa}^L > \rho_L \phi_L L^{2(t-\alpha)}] \leq \mathbb{P}_{a_L}[T_{Aa}^L > \rho_L \phi_L L^{2(\beta-\alpha)}]$ for any fixed $t > \beta$ and L large enough, and then let t tend to β.) Hence, all the above arguments carry over with $L^{2(\beta-\alpha)}$ replaced by $\phi_L L^{2(\beta-\alpha)}$. Since the modifications are minor, we work with $L^{2(\beta-\alpha)}$ in all cases.

PROOF OF THEOREM 1.4. The main difficulty is that we are interested in the first coalescence times of the pairs (A, a) and (B, b), regardless of that of any other pair. As a consequence, several coalescence and subsequent recombination events may occur before t^L, creating some correlation between lineages originally far from each other (A and b, e.g.). The point is to show that on the timescale of interest, decorrelation occurs fast enough for the system of ancestral lineages to behave like two independent genealogical processes, one for each locus.

Let us start by showing (a). Note that we can assume $\beta < 1$, since otherwise the result follows from Proposition 1.2 and the bound

$$\mathbb{P}_{a_L}\big[\mathrm{t}^L \leq L^{2(1-\alpha)}\big] \leq \mathbb{P}_{a_L}\big[\mathrm{t}_{Aa}^L \leq L^{2(1-\alpha)}\big] + \mathbb{P}_{a_L}\big[\mathrm{t}_{Bb}^L \leq L^{2(1-\alpha)}\big] \to 0$$

as $L \to \infty$.

Hence, suppose $\beta < 1$, fix $t \in (\beta, 1]$ (the case $t = \beta$ is treated as above) and let $L \in \mathbb{N}$. By the Markov property applied to $\mathcal{A}^L$ at time $\rho_L L^{2(\beta-\alpha)}$, we have

$$\begin{aligned}
&\mathbb{P}_{a_L}[\mathrm{t}^L > L^{2(t-\alpha)}] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbf{1}_{\{\mathrm{t}^L > L^{2(\beta-\alpha)}\}} \mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\qquad - \mathbb{E}_{a_L}\big[\mathbf{1}_{\{\mathrm{t}^L \le L^{2(\beta-\alpha)}\}} \mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big].
\end{aligned} \tag{22}$$

Again, the second term in (22) is bounded by

$$\mathbb{P}_{a_L}\big[\mathrm{t}^L_{Aa} \le L^{2(\beta-\alpha)}\big] + \mathbb{P}_{a_L}\big[\mathrm{t}^L_{Bb} \le L^{2(\beta-\alpha)}\big],$$

which tends to 0 as $L \to \infty$ by Proposition 1.2. Since Lemma 3.7 shows that, with probability tending to 1, at most two lineages at a time can meet at distance less than $2R_B$, we can define T^L_1 as the first time two of the four lineages come within distance $2R_B$ of each other and write

$$\begin{aligned}
&\mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}^L_1 > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\qquad + \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}^L_1 \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big].
\end{aligned} \tag{23}$$

Setting aside the first term in the right-hand side of (23) for a moment, we further decompose the event corresponding to the second term:

$$\begin{aligned}
&\mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}^L_1 \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}^L_1 \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m^L_1 \notin \{Aa, Bb\}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\qquad + \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}^L_1 \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m^L_1 \in \{Aa, Bb\}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big],
\end{aligned} \tag{24}$$

where m^L_1 denotes the pair of labels of the lineages which "meet" at time T^L_1. Let us show that the second term in (24) tends to 0 as $L \to \infty$. Using Lemma 3.4, we know that, with probability tending to one, no pairs of lineages starting at (rescaled) separation $L^{-\alpha} x_L$ have met at distance less than $2R_B$ by time $L^{2(\beta-\alpha)}$. Hence, until this time any of these pairs taken separately evolves like two independent compound Poisson processes, and their mutual distance at time $L^{2(\beta-\alpha)}$ lies within $[L^{\beta-\alpha}/(\log L), L^{\beta-\alpha} \log L]$ with probability tending to one (by a standard application of the Central Limit Theorem). On the other hand, by condition (2) we can use Proposition 4.4 with $T_L = L^{2(\beta-\alpha)}$ (see Remark 4.7) and conclude that

with probability tending to 1, the distance at time T_L between each pair of lineages starting within the same individual also lies in $[L^{\beta-\alpha}/(\log L), L^{\beta-\alpha}\log L]$. The situation has thus become rather symmetric by time $L^{2(\beta-\alpha)}$. Suppose, for instance, that $m_1^L = Aa$. Then, either $\mathrm{T}_1^L < L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - \log L$ and $\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}$ or $\mathrm{T}_1^L \in [L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - \log L, L^{2(t-\alpha)} - L^{2(\beta-\alpha)}]$. The probability of the first event tends to 0 by (11), which shows that once A and a are gathered at distance smaller than $2R_B$, they coalesce in a time smaller than $\log L$. Lemma 3.1 (if $t < 1$) or (10) (if $t = 1$) shows that the probability of the second event also tends to 0 as $L \to \infty$. Hence, the second term in (24) does indeed vanish as $L \to \infty$.

So far, we have obtained

$$\begin{aligned}
&\mathbb{P}_{a_L}\bigl[\mathrm{t}^L > L^{2(t-\alpha)}\bigr] \\
&\quad = \mathbb{E}_{a_L}\bigl[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\bigl[\mathrm{T}_1^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\bigr]\bigr] \\
&\qquad + \mathbb{E}_{a_L}\bigl[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\bigl[\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m_1^L \notin \{Aa, Bb\}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\bigr]\bigr] + \delta_L^1,
\end{aligned} \tag{25}$$

where $\delta_L^1 \to 0$ as $L \to \infty$. Next, by the strong Markov property applied to $\mathcal{A}^L$ at time $\rho_L \mathrm{T}_1^L$ and the fact that $\mathrm{T}_1^L < \mathrm{t}^L$ a.s., we have

$$\begin{aligned}
&\mathbb{E}_{a_L}\bigl[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\bigl[\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m_1^L \notin \{Aa, Bb\}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\bigr]\bigr] \\
&\quad = \mathbb{E}_{a_L}\bigl[\mathbb{E}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\bigl[\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m_1^L \notin \{Aa, Bb\}\}} \\
&\qquad\qquad \times \mathbb{P}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\bigl[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - \mathrm{T}_1^L\bigr]\bigr]\bigr].
\end{aligned}$$

If $t < 1$, Lemma 3.7 tells us that with probability tending to 1, the mutual distance between each of the 5 pairs of lineages different from m_1^L at time T_1^L belongs to the interval $[(\mathrm{T}_1^L)^{1/2}/(\log L), (\mathrm{T}_1^L)^{1/2}\log L]$. If $t = 1$, equation (10) shows that we can replace $\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(1-\alpha)} - L^{2(\beta-\alpha)}\}}$ by $\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(1-\alpha)}/(\log L)\}}$, up to an asymptotically vanishing error term, and so Lemma 3.7 still applies. Hence, by the uniform convergence stated in Lemma 3.4, the probability that one of these pairs meet at distance less than $2R_B$ before $2\mathrm{T}_1^L$ tends to zero. Furthermore, Proposition 4.4 guarantees that with very high probability, the pair that meet at time T_1^L is also at a distance belonging to $[(\mathrm{T}_1^L)^{1/2}/(\log L), (\mathrm{T}_1^L)^{1/2}\log L]$ after another T_1^L units of time. (This statement uses a conditioning on T_1^L, which turns $2\mathrm{T}_1^L$ into a deterministic time and enables us to use Proposition 4.4.) Defining T_2^L and m_2^L in the same manner as above (we number the different quantities which appear here to make the recursion clearer) and using exactly the same arguments as those leading to (25), we can thus write that with probability tending

to 1,

$$\begin{aligned}
&\mathbb{P}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\big[\mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - \mathrm{T}_1^L\big] \\
&\quad = \mathbb{E}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\big[\mathrm{T}_2^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L\big]\big] \\
&\qquad + \mathbb{E}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L \mathrm{T}_1^L)}\big[\mathrm{T}_2^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L; \\
&\qquad\qquad m_2^L \notin \{Aa, Bb\}; \\
&\qquad\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L\big]\big] + \delta_L^2,
\end{aligned}$$

with $\delta_L^2 \to 0$ as $L \to \infty$. It is easy to check that the above equality is also valid if $L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L \le 0$. By induction, we obtain for any $k \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{P}_{a_L}\big[\mathrm{t}^L > L^{2(t-\alpha)}\big] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbb{P}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathrm{T}_1^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\big]\big] \\
&\qquad + \mathbb{E}_{a_L}\big[\mathbb{E}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m_1^L \notin \{Aa, Bb\}\}} \\
&\qquad \times \mathbb{P}_{\mathcal{A}^L(2\rho_L \mathrm{T}_1^L)}\big[\mathrm{T}_2^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L\big]\big]\big] + \cdots \\
&\qquad + \mathbb{E}_{a_L}\big[\mathbb{E}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}; m_1^L \notin \{Aa, Bb\}\}} \\
&\qquad \times \mathbb{E}_{\mathcal{A}^L(2\rho_L \mathrm{T}_1^L)}\big[\mathbf{1}_{\{\mathrm{T}_2^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L\}} \mathbf{1}_{\{m_2^L \notin \{Aa, Bb\}\}} \\
&\qquad \times \mathbb{E}_{\mathcal{A}^L(2\rho_L \mathrm{T}_2^L)}\big[\cdots \mathbb{E}_{\mathcal{A}^L(2\rho_L \mathrm{T}_{k-2}^L)}\big[\mathbf{1}_{\{\mathrm{T}_{k-1}^L < L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L - \cdots - 2\mathrm{T}_{k-2}^L\}} \\
&\qquad \times \mathbf{1}_{\{m_{k-1}^L \notin \{Aa, Bb\}\}} \mathbb{P}_{\mathcal{A}^L(2\rho_L \mathrm{T}_{k-1}^L)}\big[\mathrm{T}_k^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} \\
&\qquad\qquad - \cdots - 2\mathrm{T}_{k-1}^L\big]\big]\cdots\big]\big]\big] \\
&\qquad + \mathbb{E}_{a_L}\big[\cdots \mathbb{P}_{\mathcal{A}^L(2\rho_L \mathrm{T}_{k-1}^L)}\big[\mathrm{T}_k^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L - \cdots - 2\mathrm{T}_{k-1}^L; \\
&\qquad \mathrm{t}^L > L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L - \cdots - 2\mathrm{T}_{k-1}^L\big]\cdots\big] + \sum_{i=1}^{k} \delta_L^i,
\end{aligned} \tag{26}$$

in which all occurrences of $L^{2(1-\alpha)}$ are replaced by $L^{2(1-\alpha)}/(\log L)$ if we are considering the case $t = 1$. In order to stop the recursion, let us show that for any $\varepsilon > 0$, there exists $k \in \mathbb{N}$ such that the last but one term in (26) is bounded by ε for all L large enough. To this end, define the sequence of random times $(\gamma_i^L)_{i \ge 1}$ by

$$\gamma_1^L := \inf\big\{t \ge L^{2(\beta-\alpha)} : 2 \text{ rescaled lineages meet at distance less than } 2R_B\big\},$$

and for any $i \ge 2$,

$$\gamma_i^L := \inf\{t \ge 2\gamma_{i-1}^L : 2 \text{ rescaled lineages meet at distance less than } 2R_B\}.$$

A simple recursion shows that for all $i \in \mathbb{N}$, γ_i^L and $2\gamma_i^L$ are stopping times. We can thus apply the strong Markov property at time $\rho_L \gamma_1^L$, then $\rho_L \gamma_2^L$, and so on, and obtain that

$$\begin{aligned}&\mathbb{E}_{a_L}\big[\mathbb{E}_{\mathcal{A}^L(\rho_L L^{2(\beta-\alpha)})}\big[\mathbf{1}_{\{\mathrm{T}_1^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)}\}}\mathbb{E}_{\mathcal{A}^L(2\rho_L \mathrm{T}_1^L)}\big[\mathbf{1}_{\{\mathrm{T}_2^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L\}}\\ &\qquad \cdots \times \mathbb{P}_{\mathcal{A}^L(2\rho_L \mathrm{T}_{k-1}^L)}\big[\mathrm{T}_k^L \le L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\mathrm{T}_1^L - \cdots - 2\mathrm{T}_{k-1}^L\big]\cdots\big]\big]\big]\\ &\quad = \mathbb{P}_{a_L}\big[\gamma_k^L \le L^{2(t-\alpha)}\big].\end{aligned} \tag{27}$$

Since with probability tending to 1 at each time $2\gamma_i^L$ the four lineages are at distance of the order of $(\gamma_i^L)^{1/2}$ of each other, Proposition 3.6 guarantees that, up to an asymptotically vanishing error term, the conditional probability that γ_{i+1}^L is less than $L^{2(t-\alpha)} - L^{2(\beta-\alpha)} - 2\gamma_1^L - \cdots - 2\gamma_i^L$ is bounded from above by $\mathcal{C} := (1+c)(1 - (\frac{\beta-\alpha}{t-\alpha})^6)$, where $c > 0$ can be chosen arbitrarily close to 0. It remains to choose $k \in \mathbb{N}$ such that $\mathcal{C}^k \le \varepsilon$ and to notice that the left-hand side of (27) is an upper bound for the last but one term in (26) to conclude.

Finally, let us show that the other terms in (26) are close to those corresponding to a system of four independent lineages. Using the integer $k = k(\varepsilon)$ obtained in the last paragraph, we rewrite the decomposition (26) in terms of $(\gamma_i^L)_{i\in\mathbb{N}}$ as follows (we retain the notation m_i^L for the labels of the two lineages meeting at time γ_i^L and we set $\gamma_0^L := 0$):

$$\begin{aligned}\mathbb{P}_{a_L}\big[\mathrm{t}^L > L^{2(t-\alpha)}\big] = \eta_L(\varepsilon) + \sum_{j=1}^{k} \mathbb{P}_{a_L}\big[&\gamma_{j-1}^L \le L^{2(t-\alpha)};\\ &m_l^L \notin \{Aa, Bb\}\ \forall l \in \{1,\dots,j-1\};\\ &\gamma_j^L > L^{2(t-\alpha)}\big],\end{aligned} \tag{28}$$

where $\eta_L(\varepsilon)$ is the sum of the last but one term in (26) and of the error terms δ_L^i, and is smaller than 2ε for L large enough by definition of $k(\varepsilon)$. Now, let us denote by $\hat{\mathcal{A}}^L$ a system of four independent lineages moving around on $\mathbb{T}(L)$ according to the law of the motion of a single (unrescaled) lineage, and let us define $(\hat{\gamma}_i^L)_{i\ge1}$ in the same way as $(\gamma_i^L)_{L\ge1}$ but with $\mathcal{A}^L$ replaced by $\hat{\mathcal{A}}^L$. Let us also write $\hat{\mathrm{t}}_{Aa}^L$ (resp., $\hat{\mathrm{t}}_{Bb}^L$) for the smallest time t such that the lineages A and a (resp., B and b) meet at distance less than $2R_B L^\alpha$ at time $\rho_L t$, and $\hat{m}_i^L$ for the indices of the pair meeting at time $\hat{\gamma}_i^L$. Exactly the same chain of arguments as above leads to a decomposition of $\mathbb{P}_{a_L}[\hat{\mathrm{t}}_{Aa}^L \wedge \hat{\mathrm{t}}_{Bb}^L > L^{2(t-\alpha)}]$ of the form (28), with another sequence $(\hat{\eta}_L(\varepsilon))_{L\ge1}$ whose terms are bounded by 2ε whenever L is large enough. Now, let us emphasize that Proposition 3.6 also applies to the meeting times at distance less than $2R_B L^\alpha$, before which the evolutions of $\mathcal{A}^L$ and $\hat{\mathcal{A}}^L$ have the same distribution. As a consequence, morally, we should have that the distributions of the pairs

of indices m_i^L and $\hat{m}_i^L$ both converge to a uniform draw from the set of distinct pairs of labels (in other words, each pair has asymptotically the same chance to be that meeting), and, furthermore, if γ_i^L and $\hat{\gamma}_i^L$ are of the same logarithmic order, so should γ_{i+1}^L and $\hat{\gamma}_{i+1}^L$ be.

More formally, let us define, for every $L \in \mathbb{N}$ and $j \geq 1$,

$$\mathcal{L}_j^L := \frac{\log \gamma_j^L}{2 \log L} \mathbf{1}_{\{\gamma_j^L \leq L^{2(1-\alpha)}/(\log L)\}} + \infty \mathbf{1}_{\{\gamma_j^L > L^{2(1-\alpha)}/(\log L)\}},$$

and $\hat{\mathcal{L}}_j^L$ in a similar manner. Our goal is to show that for each j, the vectors $V_j^L := (\mathcal{L}_1^L, m_1^L, \ldots, \mathcal{L}_j^L, m_j^L)$ and $\hat{V}_j^L := (\hat{\mathcal{L}}_1^L, \hat{m}_1^L, \ldots, \hat{\mathcal{L}}_j^L, \hat{m}_j^L)$ converge in distribution as $L \to \infty$ to the same random vector, whose law is obtained by successive uses of Proposition 3.6. Thus, let us prove by recursion that the distribution functions of the two vectors converge to the same limit. The case $j = 1$ is a direct consequence of Proposition 3.6, which shows that for any $s \in [\beta, 1]$ and $i_1 \neq i_2$,

$$\lim_{L \to \infty} \mathbb{P}_{a_L}[\mathcal{L}_1^L \leq s - \alpha; m_1^L = i_1 i_2] = \frac{1}{6}\left(1 - \left(\frac{\beta - \alpha}{s - \alpha}\right)^6\right) \quad \text{and}$$

$$\lim_{L \to \infty} \mathbb{P}_{a_L}[\mathcal{L}_1^L = \infty; m_1^L = i_1 i_2] = \frac{1}{6}\left(\frac{\beta - \alpha}{1 - \alpha}\right)^6.$$

[Recall the analysis made at the beginning of the proof, according to which the lineages meet before time $L^{2(\beta-\alpha)}$ with probability tending to zero, and at that time they are all at pairwise distance $\mathcal{O}(L^{\beta-\alpha})$.]

Suppose the distribution functions of V_j^L and $\hat{V}_j^L$ converge to the same (nondegenerate) limit as L tends to infinity. Let then $s \in [\beta, 1]$, $i_1 \neq i_2$ and $\mathcal{B}$ be an event of the form $\{\mathcal{L}_1^L \leq s_1 - \alpha; m_1^L = i_1^{(1)} i_2^{(1)}; \ldots; \mathcal{L}_j^L \leq s_j - \alpha; m_j^L = i_1^{(j)} i_2^{(j)}\}$ for some given $\beta \leq s_1 \leq \cdots \leq s_j \leq s$. Using the strong Markov property with $\mathcal{A}^L$ at time $2\rho_L \gamma_j^L = 2\rho_L L^{2\mathcal{L}_j^L}$ and recalling the definition of T_1^L as the first time two rescaled lineages come at distance less than $2R_B$ of each other, we obtain

$$\begin{aligned}
&\mathbb{P}_{a_L}[V_j^L \in \mathcal{B}; \mathcal{L}_{j+1}^L \leq s - \alpha; m_{j+1}^L = i_1 i_2] \\
&\quad = \mathbb{E}_{a_L}\Big[\mathbf{1}_{\{V_j^L \in \mathcal{B}\}} \mathbb{P}_{\mathcal{A}^L(2\rho_L L^{2\mathcal{L}_j^L})}\big[T_1^L \leq L^{2(s-\alpha)} - 2L^{2\mathcal{L}_j^L}; m_1^L = i_1 i_2\big]\Big] \\
&\quad = \mathbb{E}_{a_L}\left[\mathbf{1}_{\{V_j^L \in \mathcal{B}\}} \times \frac{1}{6}\left(1 - \left(\frac{\mathcal{L}_j^L}{s - \alpha}\right)^6\right)\right] \qquad (29) \\
&\qquad + \mathbb{E}_{a_L}\left[\mathbf{1}_{\{V_j^L \in \mathcal{B}\}}\left\{\mathbb{P}_{\mathcal{A}^L(2\rho_L L^{2\mathcal{L}_j^L})}[T_1^L \leq L^{2(s-\alpha)} - 2L^{2\mathcal{L}_j^L}; m_1^L = i_1 i_2]\right.\right. \\
&\qquad\qquad \left.\left. - \frac{1}{6}\left(1 - \left(\frac{\mathcal{L}_j^L}{s - \alpha}\right)^6\right)\right\}\right].
\end{aligned}$$

Since V_j^L converges in distribution to V_j^∞ as $L \to \infty$, and since the law of V_j^∞ does not charge the boundary of $\mathcal{B}$, the first term in the right-hand side of (29) converges to

$$\mathbb{E}\left[\mathbf{1}_{\{V_j^\infty \in \mathcal{B}\}} \times \frac{1}{6}\left(1 - \left(\frac{\mathcal{L}_j^\infty}{s-\alpha}\right)^6\right)\right] =: \mathbb{P}[V_j^\infty \in \mathcal{B}; \mathcal{L}_{j+1}^\infty \le s - \alpha; m_{j+1}^\infty = i_1 i_2].$$

For the second term in (29), we already saw that, up to an asymptotically vanishing error term, we can insert the indicator function of the set $\{\mathcal{A}^L(2\rho_L L^{2\mathcal{L}_j^L}) \in \Gamma(L, 4, \mathcal{L}_j^L + \alpha)\}$ within the expectation, where $\Gamma(L, 4, \eta)$ is defined at the end of Section 3 as the set of all configurations of four lineages in which all pairwise distances between the locations of the lineages belong to $[L^\eta/(\log L), L^\eta \log L]$. Now, we can also replace the first probability within the curly brackets by the probability that $T_1^L \le L^{2(s-\alpha)}$ and $m_1^L = i_1 i_2$ by Lemma 3.1. Then, the uniform convergence stated in Proposition 3.6 easily gives us that the second term in the right-hand of (29) tends to 0 as $L \to \infty$. Likewise, as L tends to infinity,

$$\begin{aligned}\mathbb{P}_{a_L}[V_j^L \in \mathcal{B}; \mathcal{L}_{j+1}^L = \infty; m_{j+1}^L = i_1 i_2] &\to \mathbb{E}\left[\mathbf{1}_{\{V_j^\infty \in \mathcal{B}\}} \frac{1}{6}\left(\frac{\mathcal{L}_j^\infty}{1-\alpha}\right)^6\right] \\ &=: \mathbb{P}[V_j^\infty \in \mathcal{B}; \mathcal{L}_{j+1}^\infty = \infty; m_{j+1}^\infty = i_1 i_2],\end{aligned}$$

and an analogous result can be established when we allow some of the $\mathcal{L}_i^L$, $i \le j$ (and so the subsequent ones) to be infinite. Since this convergence holds for all s and $i_1 i_2$ as above, we obtain the convergence in law of V_{j+1}^L toward V_{j+1}^∞, whose distribution is determined by the above limits. By the induction principle, for every $j \in \mathbb{N}$ the sequence $(V_j^L)_{L \ge 1}$ converges in distribution to a random vector V_j^∞. Since the same arguments apply to $(\hat{V}_j^L)_{L \ge 1}$, the distribution function of $\hat{V}_j^L$ also converges to that of V_j^∞ and convergence in distribution also holds. As a consequence, coming back to (28), we obtain that for each term of the sum,

$$\begin{aligned}|\mathbb{P}_{a_L}[\gamma_{j-1}^L \le L^{2(t-\alpha)}; m_l^L \notin \{Aa, Bb\}\ \forall l \in \{1, \dots, j-1\}; \gamma_j^L > L^{2(t-\alpha)}] \\ - \mathbb{P}_{a_L}[\hat{\gamma}_{j-1}^L \le L^{2(t-\alpha)}; \hat{m}_l^L \notin \{Aa, Bb\}\ \forall l \in \{1, \dots, j-1\}; \\ \hat{\gamma}_j^L > L^{2(t-\alpha)}]| \to 0\end{aligned}$$

as $L \to \infty$, and so

$$\limsup_{L \to \infty} |\mathbb{P}_{a_L}[\mathrm{t}^L > L^{2(t-\alpha)}] - \mathbb{P}_{a_L}[\hat{\mathrm{t}}^L > L^{2(t-\alpha)}]| \le 4\varepsilon.$$

Since ε was arbitrary, this limit is actually zero. But $\hat{\mathcal{A}}^L$ is a system of four independent lineages, and so

$$\mathbb{P}_{a_L}[\hat{\mathrm{t}}^L > L^{2(t-\alpha)}] = \mathbb{P}_{a_L}[\hat{\mathrm{t}}_{Aa}^L > L^{2(t-\alpha)}] \times \mathbb{P}_{a_L}[\hat{\mathrm{t}}_{Bb}^L > L^{2(t-\alpha)}] \to \left(\frac{\beta - \alpha}{t - \alpha}\right)^2$$

by Proposition 1.2. This concludes the proof of Theorem 1.4(a).

The arguments for the case (b) are very similar, using this time Lemma B for a bound on the probability that some lineages meet during a small interval of time, Lemma C for the distance separating the other lineages when two of them meet and merge and setting $\mathcal{L}_j^L := \gamma_j^L/(\frac{1-\alpha}{2\pi\sigma^2}L^{2(1-\alpha)}\log L)$. □

The proof of Theorem 1.5 uses essentially the same arguments, except that now, before time $\rho_L L^{2(\gamma-\alpha)}$, we cannot use Proposition 4.4 and the lineages starting within the same individual are still highly correlated. In fact, because recombination acts on a linear timescale whereas ancestral relations evolve on an exponential timescale, the proof will show that a phase transition occurs: during a first phase, recombination does not act and so the ancestral lines of the two loci of the same individual are not yet separated, and at time $\rho_L L^{2(\gamma-\alpha)}$ recombination appears in the picture and is quick enough to fully decorrelate the genealogies at the two loci.

PROOF OF THEOREM 1.5. The case (a) is a consequence of the result for two lineages. Indeed, if condition (3) is fulfilled, then necessarily $(\log\rho_L)/(r_L\rho_L)$ tends to infinity and for any $\varepsilon > 0$ there exists $L_0(\varepsilon)$ such that for every $L \geq L_0(\varepsilon)$,

$$\frac{\log\rho_L}{r_L\rho_L} \geq L^{2(\gamma-\alpha)-\varepsilon}.$$

Hence, since we assumed $\rho_L \leq CL^{2\alpha}$, we have for $t \in [\beta,\gamma)$, $\varepsilon := \gamma - t$ and $L \geq L_0(\varepsilon)$,

$$r_L\rho_L L^{2(t-\alpha)} \leq \log\rho_L L^{2(t-\alpha-\gamma+\alpha)+(\gamma-t)} \leq C'L^{-(\gamma-t)}\log L \to 0 \qquad \text{as } L\to\infty.$$

Therefore, with probability tending to one, no recombinations occur by time $\rho_L L^{2(t-\alpha)}$ and $\mathcal{A}^L$ boils down to a system of two lineages, one ancestral to each of the two individuals sampled. Proposition 1.2 enables us to conclude.

If $t = \gamma$ and $r_L\rho_L L^{2(\gamma-\alpha)}$ does not tend to zero (otherwise recombination is too slow and the same argument as above applies), then the probability that there is no coalescence by time $r_L^{-1}/(\log L)$ tends to $(\beta-\alpha)/(\gamma-\alpha)$. Indeed, the recombination rate on the modified timescale is of the order of $r_L\rho_L$, and so with high probability no recombinations separate the two loci in any of our two sampled individuals before time $(r_L\rho_L)^{-1}/(\log L)$. Moreover,

$$\frac{\log((r_L\rho_L)^{-1}/(\log L))}{\log L} = \frac{\log(\log\rho_L/(r_L\rho_L)) - \log\log\rho_L - \log\log L}{\log L} \to 2(\gamma-\alpha) \qquad \text{as } L\to\infty,$$

hence, by Proposition 1.2 (see also Remark 4.7), the probability that no coalescence occurs before $r_L^{-1}/(\log L)$ tends to $(\beta-\alpha)/(\gamma-\alpha)$. The last step is to observe that, again by Proposition 1.2 and Remark 4.7, the probability that any of

the pairs of lineages Aa and Bb (considered separately) coalesces during the time interval $[r_L^{-1}/(\log L), \rho_L L^{2(\gamma-\alpha)}]$ tends to 0 as L tends to infinity.

For (b), apply the Markov property at time $\psi_L := \rho_L(L^{2(\gamma-\alpha)} \vee (\log L)^5(1 + \frac{\log \rho_L}{r_L \rho_L}))$:

$$\begin{aligned}
&\mathbb{P}_{a_L}\big[\tau_{Aa}^L \wedge \tau_{Bb}^L > \rho_L L^{2(t-\alpha)}\big] \\
&\quad = \mathbb{E}_{a_L}\big[\mathbf{1}_{\{\tau_{Aa}^L \wedge \tau_{Bb}^L > \psi_L\}} \mathbb{P}_{\mathcal{A}^L(\psi_L)}\big[\tau_{Aa}^L \wedge \tau_{Bb}^L > \rho_L L^{2(t-\alpha)} - \psi_L\big]\big] \\
&\quad = \frac{(\gamma-\alpha)^2}{(t-\alpha)^2} \mathbb{P}_{a_L}[\tau_{Aa}^L \wedge \tau_{Bb}^L > \psi_L] + o(1),
\end{aligned}$$

where the second equality comes from Proposition 4.4, Theorem 1.4(a) and dominated convergence. Now, by the case (a) and Remark 4.7,

$$\mathbb{P}_{a_L}[\tau_{Aa}^L \wedge \tau_{Bb}^L > \psi_L] \to \frac{\beta-\alpha}{\gamma-\alpha} \qquad \text{as } L \to \infty,$$

which yields the desired result.

Case (c) is identical to (b). □

Acknowledgments. We thank the referees for their very careful reading and their useful remarks.

REFERENCES

[1] BARTON, N. H., ETHERIDGE, A. M. and VÉBER, A. (2010). A new model for evolution in a spatial continuum. *Electron. J. Probab.* **15** 162–216. MR2594876

[2] BARTON, N. H., KELLEHER, J. and ETHERIDGE, A. M. (2010). A new model for extinction and recolonization in two dimensions: Quantifying phylogeography. *Evolution* **64** 2701–2715.

[3] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York. MR1324786

[4] COX, J. T. (1989). Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17** 1333–1366. MR1048930

[5] COX, J. T. and DURRETT, R. (2002). The stepping stone model: New formulas expose old myths. *Ann. Appl. Probab.* **12** 1348–1377. MR1936596

[6] COX, J. T. and GRIFFEATH, D. (1986). Diffusive clustering in the two-dimensional voter model. *Ann. Probab.* **14** 347–370. MR0832014

[7] COX, J. T. and GRIFFEATH, D. (1990). Mean field asymptotics for the planar stepping stone model. *Proc. London Math. Soc.* (3) **61** 189–208. MR1051103

[8] ETHERIDGE, A. M. (2008). Drift, draft and structure: Some mathematical models of evolution. In *Stochastic Models in Biological Sciences. Banach Center Publ.* **80** 121–144. Polish Acad. Sci. Inst. Math., Warsaw. MR2433141

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

[10] LIMIC, V. and STURM, A. (2006). The spatial Λ-coalescent. *Electron. J. Probab.* **11** 363–393 (electronic). MR2223040

[11] RIDLER-ROWE, C. J. (1966). On first hitting times of some recurrent two-dimensional random walks. *Z. Wahrsch. Verw. Gebiete* **5** 187–201. MR0199901

[12] ZÄHLE, I., COX, J. T. and DURRETT, R. (2005). The stepping stone model. II. Genealogies and the infinite sites model. *Ann. Appl. Probab.* **15** 671–699. MR2114986

DEPARTMENT OF STATISTICS
UNIVERSITY OF OXFORD
1 SOUTH PARKS ROAD
OXFORD OX1 3TG
UNITED KINGDOM

CMAP—ÉCOLE POLYTECHNIQUE
ROUTE DE SACLAY
91128 PALAISEAU CEDEX
FRANCE