Rescaling limits of the spatial Lambda-Fleming-Viot process with selection

A.M. Etheridge*
Department of Statistics
University of Oxford
24-29 St Giles
Oxford OX1 3LB, UK

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

F. Yu ‡
Department of Mathematics
University of Bristol
University Walk
Bristol BS8 1TW, UK.

July 5, 2018

*AME supported in part by EPSRC Grant EP/I01361X/1

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

‡FY supported in part by EPSRC Grant EP/I028498/1.

Abstract

We consider the spatial Λ-Fleming-Viot process model [BEV10] for frequencies of genetic types in a population living in $\mathbb{R}^d$, with two types of individuals (0 and 1) and natural selection favouring individuals of type 1. We consider two cases, one in which the dynamics of the process are driven by purely 'local' events (that is, reproduction events of bounded radii) and one incorporating large-scale extinction-recolonisation events whose radii have a polynomial tail distribution. In both cases, we consider a sequence of spatial Λ-Fleming-Viot processes indexed by n, and we assume that the fraction of individuals replaced during a reproduction event and the relative frequency of events during which natural selection acts tend to 0 as n tends to infinity. We choose the decay of these parameters in such a way that when reproduction is only local, the measure-valued process describing the local frequencies of the less favoured type converges in distribution to a (measure-valued) solution to the stochastic Fisher-KPP equation in one dimension, and to a (measure-valued) solution to the deterministic Fisher-KPP equation in more than one dimensions. When large-scale extinction-recolonisation events occur, the sequence of processes converges instead to the solution to the analogous equation in which the Laplacian is replaced by a fractional Laplacian (again, noise can be retained in the limit only in one spatial dimension). We also define the process of 'potential ancestors' of a sample of individuals taken from these populations, which takes the form of a system of branching and coalescing symmetric jump processes. We show their convergence in distribution towards a system of Brownian or stable motions which branch at some finite rate. In one dimension, in the limit, pairs of particles also coalesce at a rate proportional to the local time at zero of their separation. In contrast to previous proofs of scaling limits for the spatial Λ-Fleming-Viot process, here the convergence of the more complex *forwards in time* processes is used to prove the convergence of the dual process of potential ancestries.

AMS 2010 subject classifications. *Primary:* 60G57, 60J25, 92D10 ; *Secondary:* 60J75, 60G52.
Key words and phrases: Generalised Fleming-Viot process, natural selection, limit theorems, duality, symmetric stable processes, population genetics.

1 Introduction

The principal aim of mathematical population genetics is to understand the influence of the different forces of evolution that act on a population, and the interactions between them, in shaping the patterns of genetic diversity that we see in the present-day population. One important aspect of this is the interplay between spatial structure of the population and the intrinsic randomness due to reproduction in a finite population (known as genetic drift). This is particularly mathematically challenging in one of the most biologically important situations, when the population is distributed across a two-dimensional spatial continuum. The obstructions to producing a mathematically consistent and analytically tractable model in this setting were highlighted in [Fel75] and dubbed 'the pain in the torus'. The *spatial Λ-Fleming-Viot process* (SLFV), introduced in [Eth08, BEV10], provides one route to overcoming those obstructions, and its relatively simple mathematical structure makes it a powerful tool for investigating genetic diversity in spatially structured populations. In fact, it is not so much a process as a general framework for modelling frequencies of different genetic types in populations which evolve in a spatial continuum. For example, it is readily adapted to include things like the large-scale extinction/recolonisation events which have dominated the demographic history of many species. In this paper, we shall be interested in an extension of this measure-valued process in which some individuals have higher reproductive success than others, modelling the evolution of a spatially structured population subject to *natural selection*.

Variants of the SLFV that incorporate forms of natural selection already appear in a number of studies [EFP17, EFPS17, EFS17, FP17, BEK18], but without a detailed discussion of the construction of the stochastic processes, or whether they are well-defined when the geographic space in which the population evolves is infinite. Our first contribution is to formulate and construct an SLFV with natural selection. The methods that we employ can be readily adapted to capture all of the forms of selection considered to date, and indeed the form of selection considered here contains many of them as special cases.

We shall then turn to using our model to study the interaction between natural selection, spatial structure, and genetic drift. In particular, we are interested in identifying the spatial and temporal scales over which one can expect to see a non-trivial signature of the interaction between these forces. More precisely, we investigate rescaling limits of the model which capture the resultant patterns of genetic diversity over large spatial and temporal scales. In particular, our second contribution is to find suitable scalings of time, space and of the strength of selection for which, in the limit as the scaling parameter n tends to infinity, we recover the Fisher-KPP equation [Fis37, KPP37] and, in one spatial dimension, its stochastic counterpart. In the presence of large-scale demographic events, the appropriate rescalings are different and lead to analogous equations with the Laplacian replaced by the fractional Laplacian, but, intriguingly, no other trace of the large-scale events survives. The limits obtained here assume that the local population densities are high, thus complementing results of [EFS17, EFPS17] which address the interaction of natural selection and genetic drift when local population densities are small.

The Fisher-KPP equation

$$\partial_t p = \frac{\sigma^2}{2}\Delta p + sp(1-p) \tag{1}$$

was introduced independently by Fisher [Fis37], specifically to model the spread of an advantageous gene through a spatially distributed population, and Kolomogorov, Petrovsky & Piskunov [KPP37], who also highlighted the applications to biology. Fisher considered a population living in a one-dimensional space, whereas Kolmogorov et al. worked in two dimensions (although they then assumed that the distribution of types was independent of the second coordinate, thus reducing it to the one-dimensional case). The equation has been extensively studied (and extended in many ways), and is now a standard model of invasion in biology. A major focus of work on has been the travelling wave solutions. When the motion of individuals or genes is not local but has a heavy-tailed distribution, one replaces the Laplacian in (1) by a fractional Laplacian $-(-\Delta)^\alpha$. This, notably, modifies the speed of the travelling wave solutions, which is constant in the diffusive case and increases exponentially in the fractional case; see [CR13] and references therein.

To take into account the stochasticity inherent in reproduction in a finite population, in one dimension one can add a noise term of the form

$$\varepsilon\sqrt{p(1-p)}\dot{\mathcal{W}},$$

to the right hand side of (1), where $\dot{\mathcal{W}}$ is a space-time white noise. This yields the natural continuous space analogue of the classical stepping-stone model of population genetics, introduced without selection in [Kim53], and studied in more generality in, for example, [SS80]. The (continuous space) stochastic Fisher-KPP equation can be obtained from the discrete space counterpart through rescaling (c.f. [BDE02], where the case without selection is treated) and was also obtained as the limit (over appropriate large spatial and temporal scales) of a family of long-range contact processes in [MT95]. It has been the object of intensive study, with the perturbations of solutions due to the noise when ε is very small receiving particular attention, e.g. [MS95, CD05, MMQ11] and a huge body of closely related work inspired by work of Brunet, Derrida and coworkers, e.g. [BD01]. Our results here provide the parameter regimes under which the SLFV with selection can be thought of as a noisy perturbation of the Fisher-KPP equation. Crucially, they apply in two or more spatial dimensions, where the stochastic PDE has no solution.

1.1 The spatial Λ-Fleming-Viot process with selection

The main innovation in the SLFV is that reproduction in the population is based on a Poisson point process of *events*, rather than on individuals. It is this which overcomes the pain in the torus. This is discussed in detail in [BEV10] and so we do not repeat the motivation here. Each event determines the region of space in which reproduction (or extinction/recolonisation) will take place and an *impact* u. As a result of the event, a proportion u of the individuals living in the region are replaced by offspring of a parent chosen from the population immediately before the event. The Poisson structure renders the process particularly amenable to analytic study. In the neutral setting, which has been studied rather extensively (see [BEV13] for a somewhat out of date review), the parent is chosen uniformly at random from the affected region, irrespective of type. There are many possible ways to incorporate natural selection. Here we shall focus on one of the simplest, but also most important, in which in the selection of the parent, individuals are weighted according to their genetic type.

To motivate our definition of the process with (fecundity) selection, suppose that there are two possible types in the population, which we shall denote by 0 and 1. In order to give a slight selective advantage to type 1, we fix a *selection coefficient* $s > 0$ and suppose that, when an event falls, if the proportion of type 0 individuals in the affected region immediately before the event is $\bar{w}$, then the probability of picking a type 0 parent is $p(\bar{w}, s) = \bar{w}/(1 + s(1 - \bar{w}))$. In other words, in the choice of the parent we give a weight 1 to type 0 individuals, and a weight $1 + s > 1$ to type 1 individuals, so that the probability of picking a parent of type 0 is $\bar{w}/(\bar{w} + (1 + s)(1 - \bar{w})) = p(\bar{w}, s)$. Typically one is interested in weak selection, so that $s \ll 1$ and, in this case, we can estimate this probability by $(1 - s)\bar{w} + s\bar{w}^2$. Here again we reap the benefit of the Poisson structure of events: we can think of events as being of one of two types. A proportion $(1 - s)$ of events are 'neutral': the parent is selected exactly as in the neutral setting and has probability $\bar{w}$ of being of type 0. On the other hand, a proportion s of events are 'selective' and then the probability of a type 0 parent is $\bar{w}^2$. One way to achieve this is to dictate that at selective events we choose *two* potential parents, independently, and only if both are type 0 will the offspring be type 0. The Poisson structure allows us to view neutral and selective events as being driven by independent Poisson processes. This approach exactly parallels that usually adopted to incorporate genic selection into the classical Moran model of population genetics (see, e.g., Definition 5.6 in [Eth11]). Of course there are many ways to modify the selection mechanism. For example, as in Definition 1.1 below, we can allow both the distribution of the size of the region affected and of the impact to differ between selective and neutral events, or we can consider *density dependent* selection, in which the fitness of an individual depends on the local distribution of genetic types, e.g. [EFP17].

Let us turn to a precise definition. First we describe the state space of the process, borrowing some results from [VW15] in the special case in which the compact space of possible genetic types is $K = \{0,1\}$. We suppose that the population evolves in $\mathbb{R}^d$ (although the space of geographical locations could equally, for example, be taken to be some subset of $\mathbb{R}^d$, or a d-dimensional torus). At each time t, the population is represented by a measure M_t on $\mathbb{R}^d \times K$ whose first marginal is Lebesgue measure on $\mathbb{R}^d$. As in the neutral setting, this corresponds to assuming that individuals are uniformly distributed over $\mathbb{R}^d$ and for any measurable subset E of $\mathbb{R}^d$ and $\kappa \in \{0,1\}$, $\mathrm{Vol}(E)^{-1} M_t(E \times \{\kappa\})$ gives the proportion of individuals of type κ in E. The space

$$\mathcal{M}_\lambda := \Big\{ M \text{ measure on } \mathbb{R}^d \times \{0,1\} : \ \forall f \in C_c(\mathbb{R}^d), \int_{\mathbb{R}^d} f(x) M(\mathrm{d}x, \mathrm{d}\kappa) = \int_{\mathbb{R}^d} f(x) \mathrm{d}x \Big\} \tag{2}$$

of such measures is equipped with the topology of vague convergence, which makes it a compact set (c.f. Lemma 1.1 in [VW15]). Here $C_c(\mathbb{R}^d)$ denotes the space of all compactly supported continuous functions on $\mathbb{R}^d$. A standard disintegration theorem (see e.g. [Kal02], p.561) gives us the existence of a density $w_t : \mathbb{R}^d \to [0,1]$ such that

$$M_t(\mathrm{d}x, \mathrm{d}\kappa) = \big(w_t(x) \delta_0(\mathrm{d}\kappa) + (1 - w_t(x)) \delta_1(\mathrm{d}\kappa) \big) \, \mathrm{d}x. \tag{3}$$

Morally, $w_t(x)$ represents the local fraction of individuals of type 0 at site $x \in \mathbb{R}^d$ at time t. Note that w_t is defined up to a Lebesgue null set, that is two mappings w_t and $\tilde{w}_t$ will be equivalent if and only if

$$\mathrm{Vol}\big(\{ x \in \mathbb{R}^d \ : \ w_t(x) \neq \tilde{w}_t(x) \} \big) = 0.$$

In what follows, w_t will denote any representative of the equivalence class of densities for M_t. We shall thus equally speak of M_t or w_t, depending on what makes the notation more fluid. However, it should be understood that the object of interest in all our results is the measure-valued evolution $(M_t)_{t \geq 0}$.

Definition 1.1 (SLFV with fecundity selection (SLFVS)). *Let μ, μ' be two σ-finite measures on $(0,\infty)$, and let $\nu = \{\nu_r, \, r > 0\}$, $\nu' = \{\nu'_r, \, r > 0\}$ be two collections of probability measures on $[0,1]$ such that*

$$\int_{(0,\infty)} r^d \int_0^1 u \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) < \infty, \qquad \textit{and} \qquad \int_{(0,\infty)} r^d \int_0^1 u \nu'_r(\mathrm{d}u) \mu'(\mathrm{d}r) < \infty. \tag{4}$$

Further, let Π^N and Π^S be two independent Poisson point processes on $\mathbb{R} \times \mathbb{R}^d \times (0,\infty) \times [0,1]$ with respective intensity measures $\mathrm{d}t \otimes \mathrm{d}x \otimes \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ and $\mathrm{d}t \otimes \mathrm{d}x \otimes \mu'(\mathrm{d}r)\nu'_r(\mathrm{d}u)$.

The spatial Λ-Fleming-Viot process with selection *with driving noises Π^N and Π^S is the $\mathcal{M}_\lambda$-valued process $(M_t)_{t\geq 0}$ with càdlàg paths whose dynamics are given as follows. If $(t,x,r,u) \in \Pi^N$, a neutral event occurs at time t, within the closed ball $B(x,r)$:*

1. *Sample a type κ according to the type distribution within $B(x,r)$ just before the event. That is, $\kappa = 0$ with probability $V_r^{-1} M_{t-}(B(x,r) \times \{0\})$, where V_r is the volume of a d-dimensional ball of radius r; otherwise, $\kappa = 1$.*

2. *Update the value of M_t (only) within $B(x,r)$ by setting*

$$M_t\Big|_{B(x,r)\times\{0,1\}} := (1-u) M_{t-}\Big|_{B(x,r)\times\{0,1\}} + u \, \mathrm{d}x\Big|_{B(x,r)} \otimes \delta_\kappa.$$

In words, at every site $y \in B(x,r)$ we keep a fraction $(1-u)$ of the population as it was just before the event, and we replace the remaining fraction u by descendants of the individual *with type κ chosen during the first step. These offspring all inherit the type κ of their parent. Thus, a representative of the density of M_t can be taken to be $w_t(y) = w_{t-}(y)$ if $y \notin B(x,r)$, and*

$$w_t(y) = (1-u)w_{t-}(y) + u\mathbf{1}_{\{\kappa=0\}} \qquad \text{if } y \in B(x,r).$$

Similarly, if $(t,x,r,u) \in \Pi^S$, a selective event occurs at time t, within the closed ball $B(x,r)$:

1. *Sample two types κ and κ' independently, according to the type distribution within $B(x,r)$ just before the event. We interpret them as the types of two 'potential' parents.*
2. *Update the value of M_t (only) within $B(x,r)$ by setting*

$$M_t\Big|_{B(x,r)\times\{0,1\}} := (1-u)M_{t-}\Big|_{B(x,r)\times\{0,1\}} + u\,\mathrm{d}x\Big|_{B(x,r)} \otimes \delta_{\max\{\kappa,\kappa'\}}.$$

That is, the offspring are of type 0 if and only if both potential parents are of type 0. This time, a representative of the density of M_t can be taken to be $w_t(y) = w_{t-}(y)$ if $y \notin B(x,r)$, and

$$w_t(y) = (1-u)w_{t-}(y) + u\mathbf{1}_{\{\kappa=\kappa'=0\}} \qquad \text{if } y \in B(x,r).$$

Remark 1.2. *In the results expounded later, we shall consider a family $\{w_t,\, t \geq 0\}$ of functions such that at every time $t \geq 0$, w_t is a representative of the density of M_t. In this case, it will be convenient (but not compulsory) to fix a representative w_0 of M_0 and to use the updating procedure described in Definition 1.1 to construct w_t.*

For every function $w : \mathbb{R}^d \to [0,1]$, $x \in \mathbb{R}^d$, $r > 0$ and $u \in [0,1]$, let us define

$$\begin{aligned}\Theta^+_{x,r,u}(w) &:= \mathbf{1}_{B(x,r)^c}w + \mathbf{1}_{B(x,r)}((1-u)w + u), \qquad \text{and}\\ \Theta^-_{x,r,u}(w) &:= \mathbf{1}_{B(x,r)^c}w + \mathbf{1}_{B(x,r)}(1-u)w.\end{aligned} \tag{5}$$

These quantities will correspond to the value of the density immediately after an event (t,x,r,u) if the parent is of type 0 or type 1 respectively. Let us write $C(\mathbb{R}^d)$ (resp., $\mathbb{L}^1(\mathbb{R}^d)$) for the space of all continuous (resp., integrable) functions on $\mathbb{R}^d$. Also, for every $f \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, and every $F \in C(\mathbb{R})$, we set

$$\langle w, f\rangle := \int_{\mathbb{R}^d} w(x)f(x)\mathrm{d}x \tag{6}$$

and define the function $\Psi_{F,f}$ by

$$\Psi_{F,f}(M) := F(\langle w,f\rangle) = F\bigg(\int_{\mathbb{R}^d\times\{0,1\}} f(x)\mathbf{1}_{\{0\}}(\kappa)M(\mathrm{d}x,\mathrm{d}\kappa)\bigg), \tag{7}$$

where w is any representative of the density of M. Assuming that Definition 1.1 gives rise to a well-defined Markov process, we would expect its infinitesimal generator $\mathcal{L}$ to act on functions of the form $\Psi_{F,f}$ as follows:

$$\begin{aligned}\mathcal{L}\Psi_{F,f}(M) = &\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)} \frac{1}{V_r}\Big[w(y)F(\langle\Theta^+_{x,r,u}(w),f\rangle)\\ &\qquad + (1-w(y))F(\langle\Theta^-_{x,r,u}(w),f\rangle) - F(\langle w,f\rangle)\Big]\mathrm{d}y\,\nu_r(\mathrm{d}u)\,\mu(\mathrm{d}r)\,\mathrm{d}x\\ &+ \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\Big[w(y)w(z)F\big(\langle\Theta^+_{x,r,u}(w),f\rangle\big)\\ &\qquad + (1-w(y)w(z))F(\langle\Theta^-_{x,r,u}(w),f\rangle) - F(\langle w,f\rangle)\Big]\mathrm{d}y\,\mathrm{d}z\,\nu'_r(\mathrm{d}u)\,\mu'(\mathrm{d}r)\,\mathrm{d}x.\end{aligned} \tag{8}$$

Our first result is the following.

Theorem 1.3. *Under Condition (4), there exists a unique Markov process* $(M_t)_{t\geq 0}$*, with values in* $\mathcal{M}_\lambda$*, whose infinitesimal generator acting on functions of the form* $\Psi_{F,f}$ *with* $F \in C^1(\mathbb{R})$ *and* $f \in C(\mathbb{R}^d) \cap \mathbb{L}^1(\mathbb{R}^d)$ *is given by (8).*

The result would be an obvious consequence of the Poisson point process formulation if we had chosen a compact set in place of $\mathbb{R}^d$ for the geographical space in which the population evolves, and if the intensities of the Poisson point processes Π^N and Π^S were finite, as then the global rate at which events fall and M_t is updated would be finite. The proof of existence, given in Appendix A, takes a sequence of hypercubes growing to $\mathbb{R}^d$, and two sequences of Poisson point processes whose intensities converge to the (possibly infinite) intensities μ, μ', ν, ν', and thus constructs the process $(M_t)_{t\geq 0}$ of Theorem 1.3 as a limit.

The technical Condition (4) corresponds to Assumption 2.4 in [BEV10], where it was required in the proof of existence and uniqueness of the neutral SLFV. Uniqueness in that case is proved via duality with a system of coalescing random walks that traces the location of the ancestors of individuals in a sample from the population. As we shall see below (Section 1.2), when we consider the analogous process of branching and coalescing random walks that describes the locations of all *potential* ancestors of individuals in a sample from our population with selection, Condition (4) corresponds to requiring that a given (potential) *ancestral lineage* should jump in space at a finite rate.

Observe that the reproduction events encoded by the Poisson point process Π^S favour the subpopulation of individuals of type 1, since during an event determined by Π^S, offspring are of type 0 only if both the potential parents sampled are of type 0. Since we only consider this particular form of selection in this paper, there should be no ambiguity in simply calling this process the SLFV *with selection*, but we emphasise that, although this is certainly one of the most natural, there are many alternative models. For example, one could modify the construction so that one first selects a parental type and then an impact depending on that type, or one could 'kill' with differential weights (c.f. [BP15, Fou13, Mil15] in the non-spatial setting).

We note that [EK18] describes two constructions of the SLFV. The first gives the building blocks for the existence of an SLFV with type-dependent killing, under somewhat weaker conditions than (4). The proof of existence is given (only) in the neutral case, but uniqueness remains open. The second construction, which requires Condition (4), allows for the sort of selection considered here, although, again, the actual proof of existence is only provided in the neutral case.

1.2 A dual process of branching and coalescing jump processes

Uniqueness of our process will follow from duality with a system of branching and coalescing random walks.

Recall that during a neutral event $(t, x, r, u) \in \Pi^N$, a single parental type is chosen according to the type distribution

$$\frac{1}{V_r}\int_{B(x,r)} M_{t-}(\mathrm{d}z, \mathrm{d}\kappa) = \frac{1}{V_r}\int_{B(x,r)} \big(w_{t-}(z)\delta_0(\mathrm{d}\kappa) + (1 - w_{t-}(z))\delta_1(\mathrm{d}\kappa)\big)\mathrm{d}z$$

in $B(x, r)$ at time $t-$. (Recall that V_r is the volume of a ball of radius r in $\mathbb{R}^d$). Although, strictly speaking, the density w_{t-} is only defined up to a Lebesgue null set (and so for a given z the value of $w_{t-}(z)$ may differ between two representatives of the density of M_{t-}), this sampling can be seen as picking a spatial location z uniformly at random within $B(x,r)$, and then choosing a parent from the population at z immediately before the event. Thus the parent if of type 0 with probability $w_{t-}(z)$, or 1 with probability $1-w_{t-}(z)$. Similarly, the independent sampling of two types within $B(x,r)$ during a selective event can be interpreted as choosing two locations z and z' independently and uniformly at random within $B(x,r)$, and then potential parental types according to the type distributions at z and z' just before the event.

Suppose now that we sample $k \in \mathbb{N}$ individuals at some locations $x_1, \dots, x_k \in \mathbb{R}^d$ at time 0, 'the present', assuming that the population has been evolving for an arbitrarily long time. In a neutral SLFV, one can trace back the locations of the ancestors of the individuals in a sample: when an ancestor finds itself in the region affected by a reproduction event, it will be amongst the offspring of the event with probability u, in which case its position jumps to the location of the parent, otherwise it will be unaffected. In particular, we know the distribution of the location of the parent, without any additional information about the distribution of types in the region at the time of the event. In the model with selection this is no longer the case; we are unable to decide which of the 'potential' parents is the true parent of the event without knowing their types. Instead we follow the locations of all 'potential' ancestors. We denote by $(\xi_s^1, \dots, \xi_s^{N_s})$ the locations of the N_s individuals living s units of time in the past from whom the individuals in the sample *might* have inherited their genetic types. That is, we reverse the arrow of time and trace back until the first (neutral or selective) event in which at least one of the sampled individuals belonged to the fraction u of the population replaced during the event. The parent (for a neutral event), or the two potential parents (for a selective event), are potential ancestors from whom our sampled individuals may have inherited their genetic types. Tracing further back in time, we record the locations of all individuals not yet affected by an event plus all potential ancestors. This results in a branching and coalescing system of potential ancestral lineages. Only by knowing the types of all these potential ancestors are we able to extract the true ancestry of the sample. This parallels the construction of the *ancestral selection graph* and its duality relation with the Wright-Fisher diffusion with selection in the case of a panmictic population [KN97, NK97].

Let us define the time-reversed point processes

$$\overleftarrow{\Pi}^i := \big\{(-t,x,r,u) \,:\, (t,x,r,u) \in \Pi^i\big\}, \qquad i \in \{N,S\}. \tag{9}$$

They also form two independent Poisson point processes on $\mathbb{R} \times \mathbb{R}^d \times (0,\infty) \times [0,1]$ with the same intensity measures as the corresponding forwards in time processes. We can now formulate the following definition. Let $\mathcal{M}_p(\mathbb{R}^d)$ denote the set of all finite point measures on $\mathbb{R}^d$, which we endow with the topology of weak convergence.

Definition 1.4 (Branching and coalescing dual)**.** *The branching and coalescing dual process* $(\Xi_t)_{t\geq 0}$ *is the* $\mathcal{M}_p(\mathbb{R}^d)$*-valued Markov process with càdlàg paths whose dynamics are defined as follows. Let*

$$\Xi_0 = \sum_{i=1}^{N_0} \delta_{\xi_0^i}$$

be a finite point measure on $\mathbb{R}^d$*. At any time* $t \geq 0$*, we write*

$$\Xi_t = \sum_{i=1}^{N_t} \delta_{\xi_t^i}, \tag{10}$$

where N_t is the total number of atoms in Ξ_t, and $\xi_t^1, \dots, \xi_t^{N_t}$ are their locations. We shall refer to each atom as a particle*. Then:*
At each event $(t, x, r, u) \in \overleftarrow{\Pi}^N$:

1. *To each particle in $B(x, r)$ at time $t-$ (i.e. such that $\xi_{t-}^i \in B(x, r)$), independently give a mark with probability u, or not with probability $1 - u$;*

2. *If at least one particle is marked, all the marked particles are erased from Ξ_t (i.e. the corresponding Dirac masses are removed) and are replaced by a single particle at a location drawn uniformly at random from within $B(x, r)$.*

At each event $(t, x, r, u) \in \overleftarrow{\Pi}^S$:

1. *To each particle sitting in $B(x, r)$ at time $t-$, independently give a mark with probability u, or not with probability $1 - u$;*

2. *If at least one particle is marked, all the marked particles are erased from Ξ_t and are replaced by two particles whose locations are drawn independently and uniformly from within $B(x, r)$.*

In both cases, if no particles in Ξ_{t-} are marked, then nothing happens.

Note that the point measure Ξ_t always has at least one atom, since any erasure is accompanied by the insertion of at least one new atom. When several lineages belong to the fraction of offspring created during a neutral event, they all coalesce into a single ancestral lineage, initially sitting at the position of the parent chosen during the event (which is uniformly distributed over the area affected by the event). In particular, if only one lineage is affected by a neutral event, we can see the replacement of its location ξ_{t-}^i by the parental location as a spatial jump of the i-th lineage. Because such a jump occurs only when the lineage is in the area affected by a neutral event and belongs to the fraction u of the local population replaced at that time, a lineage currently at location z jumps due to an event of $\overleftarrow{\Pi}^N$ at rate

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbf{1}_{\{|z-x| \le r\}} u \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x = \int_0^\infty \int_0^1 V_r u \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) < \infty, \tag{11}$$

where here again V_r denotes the volume of a d-dimensional ball of radius r, and the finiteness of this integral follows from the first condition in (4). If we now consider two lineages currently at locations y and y', they will merge into a single lineage due to a neutral event when they both (independently) belong to the fraction of the population replaced during the event, which gives us a coalescence rate of

$$\int_{\mathbb{R}^d} \int_0^\infty \int_0^1 \mathbf{1}_{\{|y-x| \le r\}} \mathbf{1}_{\{|y'-x| \le r\}} u^2 \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r) \mathrm{d}x = \int_0^\infty \int_0^1 V_r(y, y') u^2 \, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r), \tag{12}$$

where $V_r(y, y')$ denotes the volume of the intersection $B(y, r) \cap B(y', r)$. Observe that this rate is bounded by the expression on the r.h.s. of (11) and depends only on the distance $|y - y'|$ between the two lineages.

Let us turn to the effect of the events in $\overleftarrow{\Pi}^S$. When a single lineage is affected by a selective event, it is replaced by two lineages starting at independent locations uniformly distributed over the area of the event. This can be seen as the branching of this lineage into two. Reasoning as before, the rate at which such a branching event affects a lineage sitting at some position y is equal to

$$\int_0^\infty \int_0^1 V_r u\, \nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r), \tag{13}$$

which is also finite by the second condition in (4). When at least two lineages are affected by a selective event, they coalesce and are replaced by two new lineages (instead of one as in the case of neutral events) whose locations are independent and uniformly distributed over the area of the event. Overall, the $\mathcal{M}_p(\mathbb{R}^d)$-valued process of Definition 1.4 can thus be seen as a system of branching and coalescing jump processes describing the trajectories followed by the potential ancestral lineages of the sample. Since we only consider finitely many initial individuals in the sample, the integrability conditions (4) guarantee that the jump rate in this process is finite and so this description gives rise to a well-defined process.

The difficulty that we face in establishing a duality between this system of branching and coalescing lineages and the SLFVS is that the density w_t of the SLFVS at any time is only defined Lebesgue a.e. and so the usual test functions used to establish such dualities in population genetics, which take the form $\prod_{i=1}^k w_t(x_i)$ for fixed points $x_1, \dots, x_k$ when the underlying geographical space is discrete, will not make sense. However, if, instead of taking deterministic points $x_1, \dots, x_k$, we take random points, with a distribution which has a density ψ with respect to Lebesgue measure on $(\mathbb{R}^d)^k$, then such a test function becomes

$$\begin{aligned}&\int_{(\mathbb{R}^d)^k} \psi(x_1, \dots, x_k)\left\{\prod_{j=1}^k w_t(x_j)\right\}\mathrm{d}x_1 \cdots \mathrm{d}x_k \\ &\qquad = \int_{(\mathbb{R}^d\times\{0,1\})^k} \psi(x_1, \dots, x_k)\left\{\prod_{j=1}^k \mathbf{1}_{\{0\}}(\kappa_j)\right\} M_t(\mathrm{d}x_1, \mathrm{d}\kappa_1)\cdots M_t(\mathrm{d}x_k, \mathrm{d}\kappa_k),\end{aligned}$$

which *is* well-defined.

For every vector of k locations $(x_1, \dots, x_k) \in (\mathbb{R}^d)^k$, let us define

$$\Xi[x_1, \dots, x_k] = \sum_{i=1}^k \delta_{x_i} \in \mathcal{M}_p(\mathbb{R}^d). \tag{14}$$

We can now state the following duality property, whose proof is given in Section 2.

Proposition 1.5. *Any $\mathcal{M}_\lambda$-valued Markov process $(M_t)_{t\geq 0}$ with generator $\mathcal{L}$ defined by (8) is dual to the process $(\Xi_t)_{t\geq 0}$, in the sense that for every $k \in \mathbb{N}$, $\psi \in C((\mathbb{R}^d)^k) \cap \mathbb{L}^1((\mathbb{R}^d)^k)$, $M_0 \in \mathcal{M}_\lambda$ and $t \geq 0$, we have for any choice of the representatives w_0 (resp., w_t) of the density of M_0 (resp., M_t):*

$$\begin{aligned}&\mathbb{E}_{M_0}\left[\int_{(\mathbb{R}^d)^k} \psi(x_1, \dots, x_k)\left\{\prod_{j=1}^k w_t(x_j)\right\}\mathrm{d}x_1 \cdots \mathrm{d}x_k\right] \\ &\qquad = \int_{(\mathbb{R}^d)^k} \psi(x_1, \dots, x_k)\mathbb{E}_{\Xi[x_1,\dots,x_k]}\left[\prod_{j=1}^{N_t} w_0\big(\xi_t^j\big)\right]\mathrm{d}x_1 \cdots \mathrm{d}x_k.\end{aligned} \tag{15}$$

In Appendix A, we show the existence of such a Markov process $(M_t)_{t\geq 0}$ by an approximation argument, and then use Proposition 1.5 and the fact that the set of test functions considered is separating to conclude that the operator $\mathcal{L}$ determines a unique $\mathcal{M}_\lambda$-valued process, namely the SLFVS. This motivates the formulation of Proposition 1.5.

Remark 1.6. *Observe that for any* $\psi \in C((\mathbb{R}^d)^k) \cap \mathbb{L}^1((\mathbb{R}^d)^k)$*, the r.h.s. of (15) is well-defined and independent of the representative of the density of* M_0 *chosen. Indeed, suppose first that* $\psi \geq 0$ *and write* $\|\psi\|_1 := \int_{(\mathbb{R}^d)^k} \psi(x_1, \dots, x_k)\mathrm{d}x_1 \cdots \mathrm{d}x_k$*. Then the full integral on the right of (15) can be rewritten as*

$$\|\psi\|_1 \, \mathbb{E}_{\mathcal{X}_0}\bigg[\prod_{j=1}^{N_t} w_0\big(\xi_t^j\big)\bigg] = \|\psi\|_1 \, \mathbb{E}_{\mathcal{X}_0}\bigg[\exp\bigg\{\int_{\mathbb{R}^d} (\ln w_0(y))\Xi_t(\mathrm{d}y)\bigg\}\bigg], \tag{16}$$

where the random k-set of locations of the atoms of $\mathcal{X}_0$ *has density* $\psi/\|\psi\|_1$ *with respect to Lebesgue measure. Now, it is easy to check from Definition 1.4 that when the law of* $(\xi_0^1, \dots, \xi_0^k)$ *is absolutely continuous with respect to Lebesgue measure on* $(\mathbb{R}^d)^k$*, at any time* $t > 0$*, conditional on* N_t*, the law of* $(\xi_t^1, \dots, \xi_t^{N_t})$ *is absolutely continuous with respect to Lebesgue measure on* $(\mathbb{R}^d)^{N_t}$*. Decomposing according to the value of* N_t*, the expression on the l.h.s. of (16) can thus be written as a sum of integrals of the form*

$$\|\psi\|_1 \mathbb{P}_{\mathcal{X}_0}[N_t = m] \int_{(\mathbb{R}^d \times \{0,1\})^m} p_t(z_1, \dots, z_m) \bigg\{\prod_{j=1}^m \mathbf{1}_{\{0\}}(\kappa_j)\bigg\} M_0(\mathrm{d}z_1, \mathrm{d}\kappa_1) \cdots M_0(\mathrm{d}z_m, \mathrm{d}\kappa_m),$$

where $p_t(z_1, \dots, z_m)$ *is the density at* $(z_1, \dots, z_m)$ *of the distribution of the m points in* Ξ_t*, conditional on* $\Xi_0 = \mathcal{X}_0$ *and* $N_t = m$*. This expression is indeed independent of the representative* w_0 *of the density of* M_0 *chosen. The generalisation to any* $\psi \in C((\mathbb{R}^d)^k) \cap \mathbb{L}^1((\mathbb{R}^d)^k)$ *is then straightforward.*

1.3 Statement of the main results

Now that we have introduced the framework in which we shall work, let us state our main results. They concern the patterns of variation that we see under this model if we look over large spatial and temporal scales. In particular, we are interested in the regime of high local population density, corresponding to small impact u. We concentrate on the particular case in which $\mu'(\mathrm{d}r)\nu_r'(\mathrm{d}u) = s\mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ for a small parameter $s > 0$, which corresponds to the weighting of the selection of the parent which motivated our definition of the SLFVS. In fact, we shall choose very special forms for the measures $\mu(\mathrm{d}r)$ and $\nu_r(\mathrm{d}u)$. Our results will certainly hold under much more general conditions, but the proofs become obscured by notation.

More precisely, for each $n \in \mathbb{N}$, we fix a number $u_n \in (0, 1)$ and assume that all events (neutral and selective) have impact u_n, that is,

$$\nu_r(\mathrm{d}u) = \nu_r'(\mathrm{d}u) = \delta_{u_n}(\mathrm{d}u) \qquad \text{for every } r > 0$$

and that

$$\mu' = s_n\mu \qquad \text{for some } s_n > 0.$$

We consider the regime in which u_n and s_n go to 0 as $n \to \infty$; biologically speaking, this corresponds to assuming that the local population densities are very large while selection is weak. This mirrors the usual assumptions in the classical Moran and Wright-Fisher models, in the absence of spatial structure, in which one is interested in the scaling limits that are obtained as population size N tends to infinity while Ns_N remains $\mathcal{O}(1)$ (see, e.g., Chapter 5 in [Eth11]). We shall find scalings of time and space for which the rescaled SLFVS converges to a non-trivial limit as $n \to \infty$. Specifically, we identify a parameter regime in which the SLFVS behaves like the Fisher-KPP equation (with noise in $d = 1$), or its analogue with long-range dispersal. This will tell us how strong selection must be (relative to local population densities) if we are to see its effect over large spatial and temporal scales when local population density is large. With this in mind, let us assume that for some $\gamma, \delta, u, \sigma > 0$,

$$u_n = \frac{u}{n^\gamma}, \qquad \text{and} \qquad s_n = \frac{\sigma}{n^\delta}.$$

We shall consider the following two cases:

- **Fixed radius:** $\mu(\mathrm{d}r) = \delta_R(\mathrm{d}r)$, for some fixed $R > 0$. In this case, we choose $\gamma = 1/3$, $\delta = 2/3$, and set (in any dimension)

  $$\overline{w}_t^n(x) := \frac{1}{V_R} M_{nt}\big(B(n^{1/3}x, R) \times \{0\}\big) = \frac{1}{V_R} \int_{B(n^{1/3}x,R)} w_{nt}(y)\,\mathrm{d}y.$$

  Writing $w_t^n(\cdot) = w_{nt}(n^{1/3}\cdot)$ and $B_n(x) = B(x, n^{-1/3}R)$, we see that

  $$\overline{w}_t^n(x) = \frac{n^{d/3}}{V_R} \int_{B_n(x)} w_t^n(x),$$

  and so this scaling corresponds to scaling down the spatial coordinate by $n^{1/3}$ (so that distance one in the new units corresponds to distance $n^{1/3}$ in the original units), and to considering the timescale $(nt,\, t \geq 0)$. The random variable $\overline{w}_t^n(x)$ gives the local proportion of individuals of the unfavoured type 0 in a small neighbourhood (of radius $n^{-1/3}R$) of the point x and at time t in these new units.

- **Stable radii:** For some $\alpha \in (1, 2)$, we set

  $$\mu(\mathrm{d}r) = \frac{\mathbf{1}_{\{r \geq 1\}}}{r^{d+\alpha+1}}\,\mathrm{d}r,$$

  $$\overline{w}_t^n(x) := \frac{1}{V_1} M_{nt}\big(B(n^\beta x, 1) \times \{0\}\big) = \frac{1}{V_1} \int_{B(n^\beta x,1)} w_{nt}(y)\,\mathrm{d}y,$$

  where

  $$\beta = \frac{1}{2\alpha - 1}, \quad \gamma = \frac{\alpha - 1}{2\alpha - 1} \quad \text{and} \quad \delta = \frac{\alpha}{2\alpha - 1}. \tag{17}$$

In both cases, we write $\overline{M}_t^n$ for the random measure (taking its values in $\mathcal{M}_\lambda$) with density $\overline{w}_t^n$. It is straightforward to check that the integrability conditions (4) are satisfied; in particular, the indicator function $\mathbf{1}_{\{r \geq 1\}}$ in the definition of μ in the stable case prevents microscopic events from accumulating at a rate which would violate these conditions. Consequently, the unscaled $\mathcal{M}_\lambda$-valued process corresponding to each n is well-defined, and so is its scaled and locally averaged version $(\overline{M}_t^n)_{t\geq 0}$. Note however that the process $\overline{M}^n = (\overline{M}_t^n)_{t\geq 0}$ is not Markovian, since the change in the value of $\overline{w}_t^n(y)$ due to an event centered in $B(x, r)$ will depend on the geometry of (and the genetic diversity within a ball centered in) the intersection $B(n^\beta y, R) \cap B(x, r)$.

Remark 1.7. *We recover the parameters for the fixed radius case from those for stable radii on setting $\alpha = 2$, and so there is some sort of continuity between the two regimes. In the fixed radius case, we are able to provide an informal argument which explains why our choice for the parameters β, γ, δ is appropriate (c.f. Section 3). These heuristics also partly explain the choice of the parameter values in the stable case. The missing condition on β, γ, δ in this case is less intuitive and arises from a generator calculation, see also Section 3.*

Let us write $D_{\mathcal{M}_\lambda}[0,\infty)$ for the set of all càdlàg paths with values in $\mathcal{M}_\lambda$. Recall that the space $\mathcal{M}_\lambda$ is equipped with the topology of vague convergence. Let $C_c^\infty(\mathbb{R}^d)$ denote the set of all smooth compactly supported functions on $\mathbb{R}^d$ and recall the notation $\langle w, f\rangle$ from (6). Our main results are as follows.

Theorem 1.8. (Fixed radius) *Suppose that $(\overline{M}_0^n)_{n\geq 1}$ converges in distribution to some $M_0 \in \mathcal{M}_\lambda$. Then, as $n\to\infty$, the process $(\overline{M}_t^n)_{t\geq 0}$ converges weakly in $D_{\mathcal{M}_\lambda}[0,\infty)$ towards a Markov process $(M_t^\infty)_{t\geq 0}$ with initial value $M_0^\infty = M_0$. The limiting process is characterised as follows. Let*

$$\Gamma_R = \frac{1}{V_R}\int_{B(0,R)}\int_{B(x,R)} (z_1)^2 \mathrm{d}z\mathrm{d}x \tag{18}$$

(where z_1 denotes the first coordinate of z).

(i) When $d=1$, $(M_t^\infty)_{t\geq 0}$ is the unique process for which, for every choice of the representative w_s^∞ of the density of M_s^∞ at every time s, and for every $f,g\in C_c^\infty(\mathbb{R})$,

$$\mathcal{Z}^f := \left(\langle w_t^\infty, f\rangle - \langle w_0^\infty, f\rangle - \int_0^t \left\{\frac{u\Gamma_R}{2}\langle w_s^\infty, \Delta f\rangle - 2Ru\sigma\langle w_s^\infty(1-w_s^\infty), f\rangle\right\}\mathrm{d}s\right)_{t\geq 0}$$

is a continuous zero-mean martingale with quadratic variation at time t equal to

$$4R^2u^2\int_0^t \langle w_s^\infty(1-w_s^\infty), f^2\rangle\,\mathrm{d}s.$$

Furthermore, the bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ is given by

$$\left[\mathcal{Z}^f, \mathcal{Z}^g\right]_t = 4R^2u^2\int_0^t \langle w_s^\infty(1-w_s^\infty), fg\rangle\,\mathrm{d}s.$$

(ii) When $d\geq 2$, $(M_t^\infty)_{t\geq 0}$ is the unique (deterministic) process for which, for every choice of the representative w_s^∞ of the density of M_s^∞ at every time s, and for every $f\in C_c^\infty(\mathbb{R}^d)$ and $t\geq 0$,

$$\langle w_t^\infty, f\rangle = \langle w_0^\infty, f\rangle + \int_0^t \left\{\frac{u\Gamma_R}{2}\langle w_s^\infty, \Delta f\rangle - u\sigma V_R\langle w_s^\infty(1-w_s^\infty), f\rangle\right\}\mathrm{d}s.$$

Informally, in one space dimension, one can see the time-indexed family of densities of the limiting process $(M_t^\infty)_{t\geq 0}$ as a weak solution to the stochastic partial differential equation

$$\frac{\partial w}{\partial t} = \frac{u\Gamma_R}{2}\Delta w - 2Ru\sigma w(1-w) + 2Ru\sqrt{w(1-w)}\,\dot{\mathcal{W}}$$

(independently of the representative chosen at every time t), where $\mathcal{W}$ a space-time white noise. In dimension $d\geq 2$, on the other hand, the noise term disappears in the limit and the time-indexed family of densities of $(M_t^\infty)_{t\geq 0}$ can be seen as a weak solution to the deterministic Fisher-KPP equation

$$\frac{\partial w}{\partial t} = \frac{u\Gamma_R}{2}\Delta w - u\sigma V_R\, w(1-w).$$

To state the corresponding result for stable radii, we need some more notation. Again, we write $V_r(x,y)$ for the volume of $B(x,r)\cap B(y,r)$ and define

$$\Phi(|z-y|) := \int_{\frac{|z-y|}{2}}^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\,\mathrm{d}r.$$

Now set

$$\mathcal{D}^\alpha f(y) = u \int_{\mathbb{R}^d} \Phi(|z-y|)(f(z)-f(y))\mathrm{d}z. \tag{19}$$

We shall check in Lemma 5.1 that this defines the generator of a symmetric stable process (that is, it is a constant multiple of the fractional Laplacian). Our result for stable radii is then as follows.

Theorem 1.9. (Stable radii) *Suppose that $\overline{M}_0^n$ converges in distribution to some $M_0 \in \mathcal{M}_\lambda$. Then, as $n \to \infty$, the process $(\overline{M}_t^n)_{t\geq 0}$ converges weakly in $D_{\mathcal{M}_\lambda}[0,\infty)$ towards a Markov process $(M_t^\infty)_{t\geq 0}$ with initial value M_0. Furthermore, if $\mathcal{D}^\alpha$ denotes the generator of the symmetric α-stable process defined in (19), then*

(i) When $d=1$, $(M_t^\infty)_{t\geq 0}$ is the unique process for which, for every choice of the representative w_s^∞ of the density of M_s^∞ at every time s, and for every $f,g \in C_c^\infty(\mathbb{R})$,

$$\mathcal{Z}^f := \langle w_t^\infty, f\rangle - \langle w_0^\infty, f\rangle - \int_0^t \left\{ \langle w_s^\infty, \mathcal{D}^\alpha f\rangle - \frac{2u\sigma}{\alpha}\langle w_s^\infty(1-w_s^\infty), f\rangle \right\} \mathrm{d}s \Bigg)_{t\geq 0}$$

is a continuous zero-mean martingale with quadratic variation at time t equal to

$$\frac{4u^2}{\alpha-1}\int_0^t \langle w_s^\infty(1-w_s^\infty), f^2\rangle \,\mathrm{d}s.$$

Furthermore, the bracket process between $\mathcal{Z}^f$ and $\mathcal{Z}^g$ is given by

$$\left[\mathcal{Z}^f, \mathcal{Z}^g\right]_t = \frac{4u^2}{\alpha-1}\int_0^t \langle w_s^\infty(1-w_s^\infty), fg\rangle \,\mathrm{d}s.$$

(ii) When $d \geq 2$, $(M_t^\infty)_{t\geq 0}$ is the unique (deterministic) process for which, for every choice of the representative w_s^∞ of the density of M_s^∞ at every time s, and for every $f \in C_c^\infty(\mathbb{R}^d)$ and $t \geq 0$,

$$\langle w_t^\infty, f\rangle = \langle w_0^\infty, f\rangle + \int_0^t \left\{ \langle w_s^\infty, \mathcal{D}^\alpha f\rangle - \frac{u\sigma V_1}{\alpha}\langle w_s^\infty(1-w_s^\infty), f\rangle \right\} \mathrm{d}s.$$

Observe from the expression of $\mathcal{D}^\alpha$ given in (19) that, as in the fixed radius case, the drift component of the limiting process is proportional to u and the quadratic variation is proportional to u^2, so that u can be thought of as scaling time (we elaborate on this in Remark 3.1). Moreover, the limiting process that we obtain in the stable radius case can be seen as a weak solution to a (stochastic) PDE which only differs from that obtained in the fixed radius case in that the Laplacian has been replaced by the generator of a symmetric stable process. This is, perhaps, at first sight rather surprising. The only effect of the large scale events is on the spatial motion of individuals in the population, and we see no trace of the correlations in their movement, or of the selection or genetic drift acting over large scales, that we have in the prelimiting model. Notice also that the scaling of s_n (relative to u_n) that leads to a nontrivial limit is independent of spatial dimension. In contrast, in [FP17], the authors consider a different scaling for the parameters and prove a similar convergence result and a central limit theorem, in which the order of magnitude and the limit of the fluctuations around the deterministic limiting process are dimension-dependent.

Theorems 1.8, 1.9 have counterparts for the corresponding rescaled dual processes.

Theorem 1.10. (Fixed radius) *For every $n \in \mathbb{N}$, let $(\Xi_t)_{t\geq 0}$ be the process of branching and coalescing jump processes which is dual to the unscaled process $(M_t)_{t\geq 0}$ with parameters $\mu = \delta_R$, $\mu' = s_n\delta_R$, $\nu_R = \nu'_R = \delta_{u_n}$, where $s_n = \sigma n^{-2/3}$ and $u_n = un^{-1/3}$. Define the rescaled process $(\Xi^n_t)_{t\geq 0}$ so that for every $t \geq 0$,*

$$\Xi^n_t = \sum_{i=1}^{N^n_t} \delta_{\xi^{n,i}_t} := \sum_{i=1}^{N_{nt}} \delta_{n^{-1/3}\xi^i_{nt}}.$$

Finally, let $k \in \mathbb{N}$, ψ be a continuous probability density on $(\mathbb{R}^d)^k$ and suppose that for any $n \geq 1$ $\Xi^n_0 = \Xi[X]$, where the random variable X has density ψ with respect to Lebesgue measure (recall the notation $\Xi[\mathbf{x}]$ from (14)). Then, as $n \to \infty$, $(\Xi^n_t)_{t\geq 0}$ converges in distribution in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$ to a branching Brownian motion $(\Xi^\infty_t)_{t\geq 0}$, in which particles follow independent Brownian motions with variance parameter $u\Gamma_R$, and branch at rate $u\sigma V_R$ into two new particles, started at the location of the parent. When $d = 1$, in addition to branching and diffusing, each pair of particles, independently, also coalesces at rate $4R^2u^2$ times the local time at 0 of their separation.

The result for stable radii has the same flavour:

Theorem 1.11. (Stable radii) *For every $n \in \mathbb{N}$, let $(\Xi_t)_{t\geq 0}$ be the system of branching and coalescing jump processes which is dual to the unscaled process $(M_t)_{t\geq 0}$ corresponding to the case of stable radii. Define the rescaled process $(\Xi^n_t)_{t\geq 0}$ with parameters $u_n = u/n^{-\gamma}$ and $s_n = \sigma/n^{-\delta}$ in such a way that for every $t \geq 0$,*

$$\Xi^n_t = \sum_{i=1}^{N^n_t} \delta_{\xi^{n,i}_t} := \sum_{i=1}^{N_{nt}} \delta_{n^{-\beta}\xi^i_{nt}}.$$

Finally, let $k \in \mathbb{N}$, ψ be a continuous probability density on $(\mathbb{R}^d)^k$ and suppose that for any $n \geq 1$ we have $\Xi^n_0 = \Xi[X]$, where the random variable X has density ψ with respect to Lebesgue measure. Then $(\Xi^n_t)_{t\geq 0}$ converges in distribution in $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$ to a system $(\Xi^\infty_t)_{t\geq 0}$ of independent symmetric α-stable processes, which branch at rate $u\sigma V_1/\alpha$ into two particles starting at the location of their parent. The motion of a single particle is described by the generator $\mathcal{D}^\alpha$ defined in (19). In addition, when $d = 1$, each pair of particles, independently, coalesces at rate $4u^2/(\alpha - 1)$ times the local time at zero of their separation.

In fact, we shall use knowledge of the limiting forwards in time model to recover the corresponding limiting results for our rescaled branching and coalescing duals. The difficulty with proving these results directly stems from problems with identifying the limiting coalescence mechanism in one dimension. This contrasts with the situation of uniformly bounded local population densities (i.e., the impact u not tending to zero) considered in [BEV12] in the neutral case and in [EFS17] in the selective case, where it is the ability to identify the limiting behaviour of the (analytically tractable) coalescent dual that allows us to prove results about the forwards in time model.

As remarked above, we would obtain the same results under much more general conditions. For example, in selecting the regions to be affected by events, not only could one take more general measures μ (it is the tail behaviour of $\mu(\mathrm{d}r)$ that we see in our limits), but also reproduction events do not need to be based on balls. We anticipate that this robustness will also be maintained if one replaces our selection mechanism with any other in which one type is favoured over the other (with appropriate modifications if the strength of selection is *density dependent*, that is the parameter s_n depends on the local frequencies of the different types in the population), and it should be clear how to modify our proofs in such cases.

1.4 Structure of the paper

The rest of the paper is laid out as follows. In Section 2, we provide an expression for the generator of the SLFVS applied to a different set of test functions than those considered in (7), before proving the duality relation stated in Proposition 1.5. In Section 3, we provide heuristic arguments to explain our rescalings. In Section 4, we turn to proving Theorem 1.8, the scaling limit in the case of fixed radii, and Theorem 1.10 which provides the corresponding result for the rescaled duals. In Section 5, we prove Theorems 1.9 and 1.11, the analogous results for stable radii. In Appendix A, we prove Theorem 1.3. In Appendices B and C, we obtain continuity estimates for the rescaled SLFVS of Sections 4 and 5. In particular, these rather technical estimates are key ingredients in (and nice complements to) the proofs of Theorems 1.8 and 1.9.

2 Duality between the SLFVS and its potential ancestry

In this section, we prove Proposition 1.5, namely that if $(M_t)_{t\geq 0}$ is an $\mathcal{M}_\lambda$-valued Markov process whose generator applied to functions of the form

$$\Psi_{F,f}(M) := F\big(\langle w, f\rangle\big), \qquad f \in C(\mathbb{R}^d)\cap \mathbb{L}^1(\mathbb{R}^d),\ F\in C(\mathbb{R}),$$

is given by (8), then the duality formula (15) holds true.

To this end, we first need to extend the expression for the generator $\mathcal{L}$ to an appropriate class of test functions. This is the aim of the following lemma.

Lemma 2.1. *Let $k\in\mathbb{N}$ and $\psi\in C((\mathbb{R}^d)^k)$ be integrable. Define the function $\Phi_\psi : \mathcal{M}_\lambda\to\mathbb{R}$ by*

$$\begin{aligned}\Phi_\psi(M) &:= \int_{(\mathbb{R}^d\times\{0,1\})^k} \psi(x_1,\ldots,x_k)\bigg\{\prod_{j=1}^k \mathbf{1}_{\{0\}}(\kappa_j)\bigg\} M(\mathrm{d}x_1,\mathrm{d}\kappa_1)\cdots M(\mathrm{d}x_k,\mathrm{d}\kappa_k)\\ &= \int_{(\mathbb{R}^d)^k}\psi(x_1,\ldots,x_k)\bigg\{\prod_{j=1}^k w(x_j)\bigg\}\,\mathrm{d}x_1\cdots\mathrm{d}x_k. \end{aligned} \tag{20}$$

(Again, the above expression is independent of the representative w of the density of M chosen.) Then, writing I for the set of indices of the locations x_i that lie in the region affected by an

event, the generator $\mathcal{L}$ (defined in (8)) applied to Φ_ψ is equal to

$$\begin{aligned}
&\mathcal{L}\Phi_\psi(M) \\
&= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\int_{(\mathbb{R}^d)^k}\psi(x_1,\dots,x_k)\Big\{\prod_{j\in I^c}w(x_j)\Big\}\Big[w(y)\prod_{j\in I}\big((1-u)w(x_j)+u\big) \\
&\qquad + (1-w(y))\prod_{j\in I}\big((1-u)w(x_j)\big)-\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\dots\mathrm{d}x_k\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\int_{(\mathbb{R}^d)^k}\psi(x_1,\dots,x_k)\Big\{\prod_{j\in I^c}w(x_j)\Big\}\Big[w(y)w(z)\prod_{j\in I}\big((1-u)w(x_j)+u\big) \\
&\qquad + (1-w(y)w(z))\prod_{j\in I}\big((1-u)w(x_j)\big)-\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\dots\mathrm{d}x_k\mathrm{d}y\mathrm{d}z\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x. \qquad (21)
\end{aligned}$$

Proof of Lemma 2.1.

Any continuous integrable function $\psi:(\mathbb{R}^d)^k\to\mathbb{R}$ can be uniformly approximated by linear combinations of functions of the product form $\psi_1(x_1)\cdots\psi_k(x_k)$ with $\psi_i\in C(\mathbb{R}^d)\cap\mathbb{L}^1(\mathbb{R}^d)$ for every i. Furthermore, by polarisation, the test function

$$\prod_{i=1}^k\left(\int_{\mathbb{R}^d\times\{0,1\}}\psi_i(x_i)\mathbf{1}_{\{0\}}(\kappa_i)M(\mathrm{d}x_i,\mathrm{d}\kappa_i)\right)=\prod_{i=1}^k\langle w,\psi_i\rangle$$

can in turn be written as a linear combination of functions of the form $\langle w,f\rangle^m$, with $m\in\mathbb{N}$ and $f\in C(\mathbb{R}^d)\cap\mathbb{L}^1(\mathbb{R}^d)$, for which we can use (8) to obtain:

$$\begin{aligned}
\mathcal{L}\Psi_{(\cdot)^m,f}(M) &= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\Big[w(y)\big\langle\mathbf{1}_{B(x,r)^c}w+\mathbf{1}_{B(x,r)}((1-u)w+u),f\big\rangle^m \\
&\qquad + (1-w(y))\big\langle\mathbf{1}_{B(x,r)^c}w+\mathbf{1}_{B(x,r)}(1-u)w,f\big\rangle^m-\langle w,f\rangle^m\Big]\mathrm{d}y\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\Big[w(y)w(z)\big\langle\mathbf{1}_{B(x,r)^c}w+\mathbf{1}_{B(x,r)}((1-u)w+u),f\big\rangle^m \\
&\qquad + (1-w(y)w(z))\big\langle\mathbf{1}_{B(x,r)^c}w+\mathbf{1}_{B(x,r)}(1-u)w,f\big\rangle^m-\big\langle w,f\big\rangle^m\Big] \\
&\qquad\qquad \mathrm{d}y\,\mathrm{d}z\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x. \qquad (22)
\end{aligned}$$

Now, taking $k=m$, and $\psi(x_1,\dots,x_m)=\prod_{i=1}^m f(x_i)$ in (20), we obtain that

$$\Phi_\psi(M)=\langle w,f\rangle^m$$

(where w is any representative of the density of M). Let us thus show that, in this case, the expression on the r.h.s. of (21) coincides with (22). We focus on the first term in the r.h.s. of (21), since the computations are the same for the other terms. For fixed x,r,u,y, and writing

B for $B(x,r)$ to simplify the notation, we have

$$\int_{(\mathbb{R}^d)^m} f(x_1)\cdots f(x_m)\Big[\prod_{j:x_j\notin B} w(x_i)\Big]\Big[\prod_{j:x_j\in B}\big((1-u)w(x_j)+u\big)\Big]\mathrm{d}x_1\ldots\mathrm{d}x_m$$
$$=\sum_{J\subseteq\{1,\ldots,m\}}\int_{(\mathbb{R}^d)^m} f(x_1)\cdots f(x_m)\Big[\prod_{j\in J^c}\mathbf{1}_{\{x_j\notin B\}}w(x_j)\Big]\Big[\prod_{j\in J}\mathbf{1}_{\{x_j\in B\}}((1-u)w(x_j)+u)\Big]\mathrm{d}x_1\ldots\mathrm{d}x_m$$
$$=\sum_{J\subseteq\{1,\ldots,m\}}\big\langle\mathbf{1}_{B^c}w,f\big\rangle^{m-|J|}\big\langle\mathbf{1}_B((1-u)w+u),f\big\rangle^{|J|}$$
$$=\sum_{j=0}^{m}\binom{m}{j}\big\langle\mathbf{1}_{B^c}w,f\big\rangle^{m-j}\big\langle\mathbf{1}_B((1-u)w+u),f\big\rangle^{j}=\big\langle\mathbf{1}_{B^c}w+\mathbf{1}_B\big((1-u)w+u\big),f\big\rangle^{m},$$

which coincides with the integrand in the first part of (22). Checking that the same holds for the three other parts of (22), we can conclude that the two expressions for the action of $\mathcal{L}$ on functions of the form $\langle w,f\rangle^m$ coincide, and so do they on functions of the form $\prod_{i=1}^k\langle w,\psi_i\rangle$.

Now, for every integrable $\psi\in C((\mathbb{R}^d)^k)$ and every $M\in\mathcal{M}_\lambda$, the expression for $\mathcal{L}\Phi_\psi(M)$ given in (21) can be rewritten as

$$\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\int_{(\mathbb{R}^d)^k}\psi(x_1,\ldots,x_k)\Big\{\prod_{j\in I^c}w(x_j)\Big\}\Big[w(z)\Big(\sum_{J\subset I,J\neq I}(1-u)^{|J|}u^{|I\setminus J|}\prod_{j\in J}w(x_j)\Big)$$
$$+(1-w(z))\big((1-u)^{|I|}-1\big)\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\ldots\mathrm{d}x_k\mathrm{d}z\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$
$$+\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\int_{(\mathbb{R}^d)^k}\psi(x_1,\ldots,x_k)\Big\{\prod_{j\in I^c}w(x_j)\Big\}$$
$$\times\Big[w(z)w(z')\Big(\sum_{J\subset I,J\neq I}(1-u)^{|J|}u^{|I\setminus J|}\prod_{j\in J}w(x_j)\Big)+(1-w(z)w(z'))\big((1-u)^{|I|}-1\big)\prod_{j\in I}w(x_j)\Big]$$
$$\mathrm{d}x_1\ldots\mathrm{d}x_k\mathrm{d}z\mathrm{d}z'\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x.$$

Bounding w and $1-w$ by 1, and using the fact that

$$\sum_{J\subsetneq I}(1-u)^{|J|}u^{|I\setminus J|}=1-(1-u)^{|I|}\leq ku$$

for every $I\subseteq\{1,\ldots,k\}$, we obtain that

$$|\mathcal{L}\Phi_\psi(M)|\leq k\|\psi\|_1\Big(\int_0^\infty\int_0^1 uV_r\,\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)+\int_0^\infty\int_0^1 uV_r\,\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\Big),$$

which is finite by Condition (4). Thus if a sequence ψ_n converges in $\mathbb{L}^1$ to $\psi\in C((\mathbb{R}^d)^k)\cap\mathbb{L}^1((\mathbb{R}^d)^k)$, the sequence $\mathcal{L}\Phi_{\psi_n}(M)$ converges too, and so we can use the density argument mentioned at the beginning of the proof to extend (21) to any $\psi\in C((\mathbb{R}^d)^k)\cap\mathbb{L}^1((\mathbb{R}^d)^k)$. □

Either of these two sets of test functions, (7) or (20), characterises the law of the SLFVS (see Lemma 2.1 in [VW15] for the second set), and so we can use them interchangeably. In particular, the family (7) will be more convenient in proving the convergence of our rescaled

$\mathcal{M}_\lambda$-valued processes, whereas the duality relation that will give us the uniqueness of the limit is based on the family (20).

Armed with Lemma 2.1, we can now prove Proposition 1.5.

Proof of Proposition 1.5.

By linearity, there is no loss of generality in supposing that ψ is a probability density on $(\mathbb{R}^d)^k$, which we shall think of as the distribution of the locations of the atoms of Ξ_0.

To complete the proof, it suffices to evaluate the generator, $\mathcal{G}$, of the $\mathcal{M}_p(\mathbb{R}^d)$-valued process of jumping, branching and coalescing particles $(\Xi_t)_{t\geq 0}$ on an appropriate class of test functions. Here again, we write I for the set of indices of the atoms (or lineages) in Ξ_s that lie in the region affected by an event. Fix $M \in \mathcal{M}_\lambda$, and a given representative w of the density of M. We consider the function

$$\Phi^w : \Xi = \sum_{i=1}^{l} \delta_{x_i} \in \mathcal{M}_p(\mathbb{R}^d) \mapsto \prod_{j=1}^{l} w(x_j) = \exp\bigg(\int_{\mathbb{R}^d} (\ln w(x))\Xi(\mathrm{d}x)\bigg).$$

Then for any $l \in \mathbb{N}$ and any $x_1, \ldots, x_l \in \mathbb{R}^d$, the generator of $(\Xi_t)_{t\geq 0}$ applied to Φ^w at Ξ is given by

$$\begin{aligned}
&\mathcal{G}\Phi^w(\Xi) \\
&= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)} \frac{1}{V_r}\prod_{j\in I^c} w(x_j) \\
&\qquad \times \Bigg[\sum_{D\subseteq I; |D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\bigg(w(z)\prod_{i\in D^c} w(x_i) - \prod_{i\in I} w(x_i)\bigg)\Bigg]\mathrm{d}z\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2} \frac{1}{V_r^2}\prod_{j\in I^c} w(x_j) \\
&\qquad \times \Bigg[\sum_{D\subseteq I; |D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\bigg(w(z)w(z')\prod_{i\in D^c} w(x_i) - \prod_{i\in I} w(x_i)\bigg)\Bigg]\mathrm{d}z\mathrm{d}z'\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)} \frac{1}{V_r}\prod_{j\in I^c} w(x_j)\Bigg[w(z)\bigg(\sum_{D\subseteq I; |D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\prod_{i\in D^c} w(x_i)\bigg) \\
&\qquad - \big(1-(1-u)^{|I|}\big)\prod_{i\in I} w(x_i)\Bigg]\mathrm{d}z\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2} \frac{1}{V_r^2}\prod_{j\in I^c} w(x_j)\Bigg[w(z)w(z')\bigg(\sum_{D\subseteq I; |D|\geq 1} u^{|D|}(1-u)^{|I\setminus D|}\prod_{i\in D^c} w(x_i)\bigg) \\
&\qquad - \big(1-(1-u)^{|I|}\big)\prod_{i\in I} w(x_i)\Bigg]\mathrm{d}z\mathrm{d}z'\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x.
\end{aligned}$$

(Note that a priori this expression depends on the representative w of the density of M, but it will then be integrated with respect to ψ times Lebesgue measure on $(\mathbb{R}^d)^l$, which will make this dependence disappear.)

On the other hand, by Lemma 2.1, for any integrable function $\varphi : (\mathbb{R}^d)^l \to \mathbb{R}$, the generator of the SLFVS applied to the function Φ_φ defined as in (20) is equal to

$$\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\int_{(\mathbb{R}^d)^l}\varphi(x_1,\dots,x_l)\prod_{j\in I^c}w(x_j)\Big[w(z)\prod_{j\in I}\big((1-u)w(x_j)+u\big) \tag{23}$$
$$+(1-w(z))\prod_{j\in I}\big((1-u)w(x_j)\big)-\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\dots\mathrm{d}x_l\mathrm{d}z\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x$$
$$+\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_r^2}\int_{(\mathbb{R}^d)^l}\varphi(x_1,\dots,x_l)\prod_{j\in I^c}w(x_j)\Big[w(z)w(z')\prod_{j\in I}\big((1-u)w(x_j)+u\big)$$
$$+(1-w(z)w(z'))\prod_{j\in I}\big((1-u)w(x_j)\big)-\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\dots\mathrm{d}x_l\mathrm{d}z\mathrm{d}z'\nu'_r(\mathrm{d}u)\mu'(\mathrm{d}r)\mathrm{d}x$$

Now, the first integral in the above is equal to

$$\int_{\mathbb{R}^d}\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_r}\int_{(\mathbb{R}^d)^l}\varphi(x_1,\dots,x_l)\prod_{j\in I^c}w(x_j)\Big[w(z)\sum_{D\subseteq I;|D|\geq 1}u^{|D|}(1-u)^{|I\setminus D|}\prod_{j\in D^c}w(x_j)$$
$$-\big(1-(1-u)^{|I|}\big)\prod_{j\in I}w(x_j)\Big]\mathrm{d}x_1\dots\mathrm{d}x_l\mathrm{d}z\nu_r(\mathrm{d}u)\mu(\mathrm{d}r)\mathrm{d}x,$$

which, by Fubini's Theorem, is equal to the first term of $\mathcal{G}\Phi^w$ integrated with respect to the (signed) measure $\varphi(x_1,\dots,x_l)\mathrm{d}x_1\dots\mathrm{d}x_l$. In the same way, expanding the product, we see that the 'selection' term in $\mathcal{L}\Phi_\varphi$ is equal to the integral of the second term of $\mathcal{G}\Phi^w$ with respect to the same measure.

Recall from Remark 1.6 that if the distribution of the locations of the $N_0 = k$ atoms of Ξ_0 is absolutely continuous with respect to Lebesgue measure on $(\mathbb{R}^d)^k$, then for any $s > 0$ and conditionally on N_s, the distribution of the locations of the atoms of Ξ_s also has a density with respect to Lebesgue measure on $(\mathbb{R}^d)^{N_s}$. Now for each $0 \leq s \leq t$, partitioning on the events $\{N_s = l\}$ and applying the calculation above with φ equal to ψ times the density function of $\{\xi_s^1,\dots,\xi_s^l\}$ conditional on $N_s = l$, we deduce that for every $t \geq 0$, provided that the distribution the atoms of Ξ_0 has density ψ with respect to Lebesgue measure,

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}_{(M_0,\Xi_0)}\left[\prod_{j=1}^{N_{t-s}}w_s(\xi_{t-s}^j)\right]=0,$$

(where the expectation is with respect to the joint distribution of M_s and Ξ_{t-s}) for every $s \in [0,t]$. Integrating s over $[0,t]$, we thus have

$$\mathbb{E}_{(M_0,\Xi_0)}\left[\prod_{j=1}^{N_0}w_t(\xi_0^j)\right]=\mathbb{E}_{(M_0,\Xi_0)}\left[\prod_{j=1}^{N_t}w_0(\xi_t^j)\right],$$

which is precisely (15). □

3 Heuristics

In this section, we provide an informal justification of our choices for the parameters β, γ and δ in our scalings. Recall from the introduction that we should like to establish scalings of the selection and impact parameters, s_n and u_n, for which selection will leave a macroscopic, but not overwhelming, trace in the evolution of the population on large time- and space-scales. We wish to complement the work of [EFS17, EFPS17], in which the case of very low local population densities is considered, by identifying the order of magnitude of the strength of selection (s_n) relative to local population density ($1/u_n$), and the spatial scale over which to look, to see the local genetic diversities evolve in a non-trivial way when population densities are very large. To understand which scalings will yield non trivial limiting processes in Theorems 1.8 and 1.9, it is convenient to think about the corresponding scaled branching and coalescing dual processes.

First consider the case of fixed (or more generally bounded variance) radius events. Ignoring for a moment the selective (branching) events and assuming that space is scaled down by n^β, a single ancestral lineage in the scaled dual makes mean zero, finite variance, jumps of size of order $1/n^\beta$ at rate proportional to $nu_n = n^{1-\gamma}$. Thus, provided that $1 - \gamma = 2\beta$, its spatial motion will converge to Brownian motion (with a given variance parameter) as $n \to \infty$.

Now consider what happens at a selective event. The two new potential ancestral lineages are born at a separation of order $1/n^\beta$. If we are to 'see' the event, the two lineages must move apart to a separation of order one before (perhaps) coalescing. The number of excursions they must make away from the region in which they can both be affected by an event (and thus coalesce) before we can expect to see such a 'long' excursion is order 1 in $d \geq 3$, order $\log n$ in $d = 2$ and order n^β in $d = 1$. On the other hand, when they are sufficiently close together that they can be hit by the same event, given that one of them jumps, there is a probability of order $u_n = u/n^\gamma$ that the other one is affected by the same event and so they coalesce. So the number of times they come close to one another before they coalesce is order n^γ. Thus, in the limit as $n \to \infty$, for each branching event in the dual, in dimensions at least 2, the probability that there is a long excursion before coalescence (and so we 'see' the event) tends to one. Moreover, the same argument tells us that we will never see coalescence of any other lineages in our system, since each time they come sufficiently close to each other to have a chance to merge, with probability tending to one they separate again without coalescing. In one dimension, we can expect to see both branching and coalescence provided that the number of excursions we expect to wait before seeing a coalescence and the number we expect to wait before the lineages escape to a distance of order one, that is $1/u_n$ and n^β respectively, are comparable. This gives $\beta = \gamma$ and, combining with the condition $1 - \gamma = 2\beta$ above, we find $\beta = \gamma = 1/3$. Finally, selection events affect a given lineage at a rate proportional to ns_nu_n in the rescaled process (recall that time is accelerated by a factor n and a lineage is affected by an event only if it belongs to the fraction u_n of the local population replaced), and so we choose $\delta = 2/3$ to make this order one.

We now turn to the stable case. As before, we first consider the motion of a single rescaled lineage. This lineage jumps only when it is in the fraction u_n of individuals which are created during an event that overlaps it, and its new position is chosen uniformly over a ball whose radius is given by the intensity measure μ with polynomial decay described in (17). Consequently, if we choose $nu_n \propto n^{\alpha\beta}$, i.e. $1 - \gamma = \alpha\beta$, then in the limit as $n \to \infty$ its motion will converge to a symmetric α-stable process with index α. Second, in order to see any branching of lineages through selection events at all, we need again ns_nu_n to be order one, that is $1 - \gamma - \delta = 0$. Finally, let us consider coalescence. Since $u_n \to 0$ as $n \to \infty$, although it is now the case that two lineages can always be affected by the same event (the radii of the events are not bounded), 'most of the time' they will not and the motions are almost independent. Consequently, the difference of their positions is also approximately described by an α-stable process. Now, because events of radius $\mathcal{O}(1)$ are much more frequent than events of large radii $\mathcal{O}(n^a)$ for any $a > 0$, and the probability that both lineages belong to the fraction of the local population replaced during an event is tiny (u_n^2), if coalescence is to happen in the limit, then we expect it to be driven by small events. In more than one dimension, the rotation-invariant α-stable processes with $\alpha \in (1,2)$ are transient (see Example 37.19(ii) in [Sat99]), and so as in the fixed radius case, this tells us that the two lineages do not spend enough time close together for coalescence to occur, whatever our choice of β, γ, α consistent with the previous conditions. In one dimension, we have not found a simple heuristic explanation for the last condition on the parameters (which one would expect to be analogous to the comparison between the number and lengths of visits in a neighbourhood of zero for the difference process, and the coalescence rate of the two lineages, carried out in the fixed radius case). Instead, the condition $\gamma = (\alpha - 1)\beta$ will emerge when we control the second term on the r.h.s. of (56), which corresponds to the variance term in the limiting process (and thus to the coalescence term in the dual process). See also Equations (59) and (60) and the surrounding paragraphs. In the end, we have three equations in three unknowns (in one dimension) and solving gives the values in Equation (17).

Remark 3.1. *At first sight, these scalings do not perhaps look altogether natural. The reason for this is that the* timescale *of the SLFV process is not one of generations. Suppose that one thinks of a generation as being the time that it takes for an 'individual' in the SLFV to be affected by a reproduction event. Then a generation is proportional to $1/u$ units of SLFV time. In the 'generation timescale', we are speeding up time by a factor of nu_n and then we recognise the scaling in the fixed radius case as exactly the diffusive rescaling and the scaling in the stable case as its natural analogue when we have long-range dispersal.*

Note also that by choosing $1 - \gamma = \alpha\beta$ and $1 - \gamma - \delta = 0$ but with $\gamma > (\alpha - 1)\beta$, we could eliminate the coalescence term in one dimension, corresponding to removing the noise term in the forwards in time description of allele frequencies.

It turns out to be highly non-trivial to turn these heuristics into a rigorous proof and so, instead, we work with the forwards in time model and deduce convergence of the dual processes as a corollary.

4 Convergence of the rescaled SLFVS and its dual - the fixed radius case

In this section, we prove Theorem 1.8 and, from it, deduce Theorem 1.10. Recall our notation $V_r(x,y)$ for the volume of the intersection $B(x,r) \cap B(y,r)$ and that in the fixed radius case, all reproduction events have the same radius $R > 0$. Let $C_c^\infty(\mathbb{R}^d)$ denote the set of all functions of class C^∞ with compact support on $\mathbb{R}^d$.

4.1 Proof of Theorem 1.8.

The proof proceeds in the usual way. First we show that the sequence of 'nearly-Markovian' rescaled and locally averaged processes is tight, then we identify the possible limit points and finally uniqueness of the limit point guarantees that the whole sequence in fact converges.

1) Tightness.

Since the state space of the processes is compact (in the topology of vague convergence), we know that any possible limit will take its values in $\mathcal{M}_\lambda$ and, furthermore, we can use the Aldous-Rebolledo criterion [Ald78, Reb80] to reduce the problem to tightness of the sequences of the finite variation parts and of the quadratic variation of the martingale parts of $(\Psi_{F,f}(\overline{M}_t^n))_{t\geq 0}$

for every $F \in C^3(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R}^d)$, where $\Psi_{F,f}$ was defined in (7). Let us therefore establish an expression for these quantities.

Although we are interested in the sequence $\overline{M}^n$, we begin by considering the scaling acting on M. A judicious choice of test functions will then allow us to deduce expressions for the finite and quadratic variation parts of $\mathcal{L}\Psi_{F,\varphi}(\overline{M})$.

Recall the notation Θ^+ and Θ^- from (5), and fix $F \in C^3(\mathbb{R})$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$. By construction (see Theorem 1.3), we know that before scaling space and time, the generator of the SLFVS with reproduction events of fixed radius R, and parameters u_n, s_n, acting on the function $\Psi_{F,\varphi}$ is given by

$$\mathcal{L}\Psi_{F,\varphi}(M) = \int_{\mathbb{R}^d}\int_{B(x,R)^2} \frac{1}{V_R^2}\Big\{w(y)(1+s_n w(z))\big[F(\langle\Theta^+_{x,R,u_n}(w),\varphi\rangle) - F(\langle w,\varphi\rangle)\big] \\ + (1-w(y)+s_n(1-w(y)w(z)))\big[F(\langle\Theta^-_{x,R,u_n}(w),\varphi\rangle) - F(\langle w,\varphi\rangle)\big]\Big\}\mathrm{d}y\mathrm{d}z\mathrm{d}x,$$

where w is a representative of the density of M. From the proof of Theorem 1.3 given in Appendix A, the finite variation part of SLFVS $(M_t)_{t\geq 0}$ is

$$\mathcal{A}_t = \int_0^t \mathcal{L}\Psi_{F,\varphi}(M_s)\mathrm{d}s,$$

and its quadratic variation is given by

$$\mathcal{Q}_t = \int_0^t\int_{\mathbb{R}^d}\int_{B(x,R)^2} \frac{1}{V_R^2}\Big\{w_s(y)(1+s_n w_s(z))\big[F(\langle\Theta^+_{x,R,u_n}(w_s),\varphi\rangle) - F(\langle w_s,\varphi\rangle)\big]^2 \\ + (1-w_s(y)+s_n(1-w_s(y)w_s(z)))\big[F(\langle\Theta^-_{x,R,u_n}(w_s),\varphi\rangle) - F(\langle w_s,\varphi\rangle)\big]^2\Big\}\mathrm{d}y\mathrm{d}z\mathrm{d}x\mathrm{d}s.$$

(Notice that we have suppressed the dependence of $(M_t)_{t\geq 0}$ on n in the notation for simplicity. Also, w_s is a representative of M_s for every s, which can be chosen in a consistent way as explained in Remark 1.2.) Let us now consider the Markov process $(M_t^n)_{t\geq 0}$ whose density at time t is $w_t^n(\cdot) := w_{nt}(n^{1/3}\,\cdot)$. We set

$$B_n(x) = B(x, n^{-1/3}R) \tag{24}$$

and write $\overline{w}(x) = n^{d/3}V_R^{-1}\int_{B_n(x)} w(z)\mathrm{d}z$. In particular, in the notation of Section 1.3 we have for every $t \geq 0$

$$\frac{n^{d/3}}{V_R}\int_{B_n(x)} w_t^n(z)\mathrm{d}z = \frac{1}{V_R}\int_{B(n^{1/3}x,R)} w_{nt}(y)\mathrm{d}y = \overline{w}_t^n(x).$$

From our expression for $\mathcal{L}$, accelerating time by a factor n and performing several changes of the spatial variables, we obtain that the generator of M^n is given by

$$\begin{aligned}
&\mathcal{L}^n\Psi_{F,\varphi}(M)\\
&= n\int_{\mathbb{R}^d}\int_{B(x,R)^2}\frac{1}{V_R^2}\Big\{w(n^{-1/3}y)\big(1+s_n w(n^{-1/3}z)\big)\big[F(\langle\Theta^+_{n^{-1/3}x,n^{-1/3}R,u_n}(w),\varphi\rangle)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - F(\langle w,\varphi\rangle)\big]\\
&\quad + (1-w(n^{-1/3}y)+s_n(1-w(n^{-1/3}y)w(n^{-1/3}z)))\big[F(\langle\Theta^-_{n^{-1/3}x,n^{-1/3}R,u_n}(w),\varphi\rangle)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - F(\langle w,\varphi\rangle)\big]\Big\}\mathrm{d}y\mathrm{d}z\mathrm{d}x\\
&= n^{1+\frac{d}{3}}\int_{\mathbb{R}^d}\Big\{\overline{w}(x)(1+s_n\overline{w}(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w),\varphi\rangle)-F(\langle w,\varphi\rangle)\big]\\
&\qquad + \big(1-\overline{w}(x)+s_n(1-\overline{w}(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w),\varphi\rangle)-F(\langle w,\varphi\rangle)\big]\Big\}\mathrm{d}x. \qquad (25)
\end{aligned}$$

The finite variation part of $(\Psi_{F,\varphi}(M^n_t))_{t\geq 0}$ is then obtained by integrating $\mathcal{L}^n\Psi_{F,\varphi}(M^n_s)$ with respect to time. Likewise, its quadratic variation at time t is equal to

$$\begin{aligned}
&n^{1+\frac{d}{3}}\int_0^t\int_{\mathbb{R}^d}\Big\{\overline{w}^n_s(x)(1+s_n\overline{w}^n_s(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w^n_s),\varphi\rangle)-F(\langle w^n_s,\varphi\rangle)\big]^2\\
&\quad + \big(1-\overline{w}^n_s(x)+s_n(1-\overline{w}^n_s(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w^n_s),\varphi\rangle)-F(\langle w^n_s,\varphi\rangle)\big]^2\Big\}\mathrm{d}x\mathrm{d}s.
\end{aligned}$$

Finally, it remains to evaluate the above expressions with φ of the form

$$\varphi_f(x) = \frac{n^{d/3}}{V_R}\int_{B_n(x)} f(y)\mathrm{d}y \qquad (26)$$

for some $f\in C_c^\infty(\mathbb{R}^d)$ and to use the fact that, by Fubini's Theorem,

$$\langle w^n,\varphi_f\rangle = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} w^n(y)\frac{n^{d/3}}{V_R}f(z)\mathbf{1}_{\{|z-y|\leq n^{-1/3}R\}}\mathrm{d}y\mathrm{d}z = \langle\overline{w}^n,f\rangle,$$

to obtain that the finite variation part of $(\Psi_{F,f}(\overline{M}^n_t))_{t\geq 0}$ is given by

$$\mathcal{A}^n_t = \int_0^t \mathcal{L}^n\Psi_{F,\varphi_f}\big(M^n_s\big)\mathrm{d}s, \qquad (27)$$

with $\mathcal{L}^n\Psi_{F,\varphi_f}$ as in (25), and its quadratic variation is given by

$$\begin{aligned}
\mathcal{Q}^n_t &= n^{1+\frac{d}{3}}\int_0^t\int_{\mathbb{R}^d}\Big\{\overline{w}^n_s(x)(1+s_n\overline{w}^n_s(x))\big[F(\langle\Theta^+_{x,n^{-1/3}R,u_n}(w^n_s),\varphi_f\rangle)-F(\langle w^n_s,\varphi_f\rangle)\big]^2\\
&\quad + \big(1-\overline{w}^n_s(x)+s_n(1-\overline{w}^n_s(x)^2)\big)\big[F(\langle\Theta^-_{x,n^{-1/3}R,u_n}(w^n_s),\varphi_f\rangle)-F(\langle w^n_s,\varphi\rangle)\big]^2\Big\}\mathrm{d}x\mathrm{d}s. \qquad (28)
\end{aligned}$$

Note that

$$\begin{aligned}
&\langle\Theta^+_{x,n^{-1/3}R,u_n}(w),\varphi_f\rangle - \langle w,\varphi_f\rangle = u_n\,\langle\mathbf{1}_{B_n(x)}(1-w),\varphi_f\rangle\\
&\langle\Theta^-_{x,n^{-1/3}R,u_n}(w),\varphi_f\rangle - \langle w,\varphi_f\rangle = -u_n\,\langle\mathbf{1}_{B_n(x)}w,\varphi_f\rangle,
\end{aligned}$$

so that both increments are of the order of $u_n n^{-d/3}$. Moreover, f has compact support S_f in $\mathbb{R}^d$ and thus so has φ_f. This will enable us to control the integrals over space of these increments.

Using this observation, we first show that $|\mathcal{A}_t^n|$ is bounded by a constant independent of n. To this end, we write it as the sum of a neutral term and a selective term and perform a Taylor expansion of F (truncating at second order in the neutral term and at first order in the selective term). This yields, for any $t \geq 0$,

$$\mathcal{A}_t^n = \int_0^t \big(A_n(s) + B_n(s) + C_n(s) + D_n(s) + E_n(s)\big)\,\mathrm{d}s,$$

where

$$\begin{aligned}
A_n(s) &= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \Big[\overline{w}_s^n(x)\langle \mathbf{1}_{B_n(x)}(1-w_s^n), \varphi_f\rangle \\
&\qquad\qquad\qquad\qquad - (1-\overline{w}_s^n(x))\langle \mathbf{1}_{B_n(x)} w_s^n, \varphi_f\rangle\Big]\mathrm{d}x, \\
B_n(s) &= u_n^2 n^{1+\frac{d}{3}} \frac{F''(\langle \overline{w}_s^n, f\rangle)}{2} \int_{\mathbb{R}^d} \Big[\overline{w}_s^n(x)\langle \mathbf{1}_{B_n(x)}(1-w_s^n), \varphi_f\rangle^2 \\
&\qquad\qquad\qquad\qquad + (1-\overline{w}_s^n(x))\langle \mathbf{1}_{B_n(x)} w_s^n, \varphi_f\rangle^2\Big]\mathrm{d}x, \\
C_n(s) &\leq \mathcal{C} n^{1+\frac{d}{3}} \int_{\mathbb{R}^d} \big(u_n \mathrm{Vol}(B_n(x))\big)^3 \mathbf{1}_{\{B_n(x)\cap S_f \neq \emptyset\}}\mathrm{d}x, \\
D_n(s) &= u_n s_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \Big[\overline{w}_s^n(x)^2\langle \mathbf{1}_{B_n(x)}(1-w_s^n), \varphi_f\rangle \\
&\qquad\qquad\qquad\qquad - (1-\overline{w}_s^n(x)^2)\langle \mathbf{1}_{B_n(x)} w_s^n, \varphi_f\rangle\Big]\mathrm{d}x, \\
E_n(s) &\leq \mathcal{C}' n^{1+\frac{d}{3}} s_n u_n^2 \int_{\mathbb{R}^d} \mathrm{Vol}(B_n(x))^2 \mathbf{1}_{\{B_n(x)\cap S_f \neq \emptyset\}}\mathrm{d}x,
\end{aligned}$$

for some constant $\mathcal{C}, \mathcal{C}'$ independent of n and s. To control these expressions, we take a Taylor expansion of φ_f. We illustrate with the term $A_n(s)$. In fact, in identifying the limiting process we shall need a precise expression for the limit of $A_n(s)$ and so we perform the expansion slightly more carefully than would be required to simply conclude boundedness.

Let us write $D\varphi_f$ for the vector of first derivatives of φ_f and $H\varphi_f$ for the corresponding Hessian ($H\varphi_f = DD\varphi_f$). We also denote the compact support of f by S_f. Then

$$\begin{aligned}
A_n(s) &= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \Big[\overline{w}_s^n(x)\langle \mathbf{1}_{B_n(x)}, \varphi_f\rangle - \langle \mathbf{1}_{B_n(x)} w_s^n, \varphi_f\rangle\Big]\mathrm{d}x \\
&= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int\int \mathbf{1}_{\{|y-x|\leq n^{-1/3}R\}} \mathbf{1}_{\{|z-x|\leq n^{-1/3}R\}} w_s^n(y)(\varphi_f(z) - \varphi_f(y))\mathrm{d}z\mathrm{d}y\mathrm{d}x \\
&= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}_{\{|y-x|\leq n^{-1/3}R\}} \mathbf{1}_{\{|z-x|\leq n^{-1/3}R\}} w_s^n(y) \\
&\qquad\qquad \times \big[D\varphi_f(y)(z-y) + \frac{1}{2}(z-y)H\varphi_f(y)(z-y) + \mathcal{O}(|z-y|^3)\mathbf{1}_{\{y\in S_f\}}\big]\mathrm{d}z\mathrm{d}y\mathrm{d}x.
\end{aligned}$$

Consider the first term on the right. Integrating first with respect to x (using Fubini's Theorem) this term is

$$\frac{u_n n^{1+\frac{2d}{3}}}{V_R} F'(\langle \overline{w}_s^n, f\rangle) \int_{\mathbb{R}^d} w_s^n(y) \int_{\mathbb{R}^d} \mathrm{Vol}(B_n(y)\cap B_n(z)) D\varphi_f(y)(z-y)\mathrm{d}z\mathrm{d}y,$$

and since $\mathrm{Vol}(B_n(y)\cap B_n(z))$ is a function of $|z-y|$ alone, the integrand is antisymmetric as a function of $z-y$ and so the integral with respect to z vanishes.

Similary, the integrals corresponding to the off-diagonal terms in the Hessian will vanish, leaving

$$u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f\rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}_{\{|y-x|\le n^{-1/3}R\}}\mathbf{1}_{\{|z-x|\le n^{-1/3}R\}} w^n_s(y) \times \frac{1}{2}\sum_{i=1}^d (z_i-y_i)^2 \frac{\partial^2}{\partial y_i^2}\varphi_f(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x$$

plus a lower order term. Now observe that since $f\in C_c^\infty(\mathbb{R}^d)$, another Taylor expansion argument enables us to write that

$$\frac{\partial^2}{\partial y_i^2}\varphi_f(y) = \varphi_{\frac{\partial^2 f}{\partial y_i^2}}(y) = \frac{\partial^2 f}{\partial y_i^2}(y) + \mathcal{O}(n^{-2/3})\mathbf{1}_{\{B_n(y)\cap S_f\neq\emptyset\}}$$

(where the term $\mathcal{O}(n^{-2/3})$ is independent of y). This yields

$$\begin{aligned}
A_n(s) &= u_n n^{1+\frac{d}{3}} F'(\langle \overline{w}^n_s, f\rangle) \int_{\mathbb{R}^d} \frac{n^{d/3}}{V_R} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}_{\{|y-x|\le n^{-1/3}R\}}\mathbf{1}_{\{|z-x|\le n^{-1/3}R\}} w^n_s(y) \\
&\qquad \times \frac{1}{2}\sum_{i=1}^d (z_i-y_i)^2 \frac{\partial^2 f}{\partial y_i^2}(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x \\
&\quad + \mathcal{O}(n^{-2/3}) n^{\frac{2}{3}(1+d)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}_{\{|y-x|\le n^{-1/3}R\}}\mathbf{1}_{\{|z-x|\le n^{-1/3}R\}} |z-y|^2 \mathbf{1}_{\{B_n(y)\cap S_f\neq\emptyset\}} \\
&= \frac{u n^{\frac{2}{3}(1+d)}}{2V_R} F'(\langle \overline{w}^n_s, f\rangle) \int_{\mathbb{R}^d}\int_{B_n(x)^2} w^n_s(y)\sum_{i=1}^d (z_i-y_i)^2\frac{\partial^2 f}{\partial y_i^2}(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x + \mathcal{O}(n^{-2/3}) \\
&= \frac{u\Gamma_R}{2} F'(\langle \overline{w}^n_s, f\rangle) \int_{\mathbb{R}^d} w^n_s(y)\Delta f(y)\mathrm{d}y + \mathcal{O}(n^{-2/3}) \\
&= \frac{u\Gamma_R}{2} F'(\langle \overline{w}^n_s, f\rangle)\,\langle \overline{w}^n_s, \Delta f\rangle + \mathcal{O}(n^{-2/3}),
\end{aligned} \tag{29}$$

where

$$\Gamma_R = \frac{n^{\frac{2}{3}(1+d)}}{V_R}\int_{B_n(y)}\int_{B_n(x)} (z_1-y_1)^2\mathrm{d}z\mathrm{d}x = \frac{1}{V_R}\int_{B(0,R)}\int_{B(x,R)} (z_1)^2\mathrm{d}z\mathrm{d}x$$

was defined in (18), and the last inequality uses another Taylor expansion to show that for any s,

$$\langle w^n_s, \Delta f\rangle = \langle \overline{w}^n_s, \Delta f\rangle + \mathcal{O}(n^{-2/3}) \tag{30}$$

with an error term uniformly bounded in s. In particular, since $|\langle \overline{w}^n_s, f\rangle| \le \|f\|\mathrm{Vol}(S_f)$, we can conclude that $|A_n(s)|\le \mathcal{C}_A$ uniformly in s and n.

Very similar arguments allow us to control the other terms:

$$|B_n(s)| \le \frac{u_n^2 n^{1+\frac{d}{3}}}{2}\,|F''(\langle \overline{w}^n_s, f\rangle)| \int 2\mathrm{Vol}(B_n(x))^2 \mathbf{1}_{\{x\in S_f\}}\|f\|^2\mathrm{d}x \le \mathcal{C}_B\, n^{\frac{1-d}{3}},$$

and, again by the same arguments,

$$|C_n(s)| \le \mathcal{C}_C\, n^{-\frac{2d}{3}}, \quad |D_n(s)|\le \mathcal{C}_D \quad \text{and} \quad |E_n(s)| \le \mathcal{C}_E\, n^{-\frac{1+d}{3}}.$$

Consequently, for every $s < t$ we have

$$|\mathcal{A}_t^n - \mathcal{A}_s^n| \leq \big(\mathcal{C}_A + \mathcal{C}_B\, n^{\frac{1-d}{3}} + \mathcal{C}_C\, n^{-\frac{2d}{3}} + \mathcal{C}_D + \mathcal{C}_E\, n^{-\frac{1+d}{3}}\big)(t-s),$$

which shows that the sequence of finite variation parts of $(\Psi_{F,f}(\overline{M}_t^n))_{t\geq 0}$ is tight. (In fact, its modulus of continuity is uniformly bounded and so we actually have tightness in the topology of uniform convergence over compact time intervals).

Similarly, we obtain that

$$\big[F(\langle \Theta^{\pm}_{x,n^{-1/3}R,u_n}(w_s^n), \varphi_f\rangle) - F(\langle w_s^n, \varphi_f\rangle)\big]^2 \leq \mathcal{C}_F''\|f\|^2 u_n^2 \mathrm{Vol}(B_n(x))^2 \mathbf{1}_{\{B_n(x)\cap S_f \neq \emptyset\}}.$$

Notice that this bound is independent of the value of w_s^n. Substituting into the definition of $\mathcal{Q}_t^n$ given in (28), we obtain that for every $s < t$,

$$|\mathcal{Q}_t^n - \mathcal{Q}_s^n| \leq \mathcal{C}_F\, n^{\frac{1-d}{3}}(t-s), \tag{31}$$

and the sequence of quadratic variations of the martingale part of $(\Psi_{F,f}(\overline{M}_t^n))_{t\geq 0}$ is not only tight, but also when $d \geq 2$ it tends to 0 uniformly over compact time intervals.

Combining these results with the Aldous-Rebolledo criterion, we conclude that $(\overline{M}^n)_{n\geq 1}$ is tight in $D_{\mathcal{M}_\lambda}[0,\infty)$, as required.

2) Limiting process.

We now identify the limiting process. In what follows, we suppose that $(M_t^\infty)_{t\geq 0} \in D_{\mathcal{M}_\lambda}[0,\infty)$ is the weak limit of a subsequence of $(\overline{M}^n)_{n\geq 1}$ and for any $t \geq 0$, we write w_t^∞ for (some representative of) the density of M_t^∞. To characterise the law of M^∞, it is sufficient to consider test functions of the forms

$$\int_{\mathbb{R}^d\times\{0,1\}} f(x)\mathbf{1}_{\{0\}}(\kappa) M(\mathrm{d}x, \mathrm{d}\kappa) = \langle w, f\rangle \qquad \text{and} \qquad \langle w, f\rangle^2,$$

with $f \in C_c^\infty(\mathbb{R}^d)$. Indeed, by polarisation (and Itô's formula) we can then obtain the semimartingale decomposition of $\langle w_t^\infty, f\rangle\langle w_t^\infty, g\rangle$ for every $f, g \in C_c^\infty(\mathbb{R}^d)$. Using Itô's formula, we can extend this decomposition to any process of the form $(\prod_{i=1}^k \langle w_t^\infty, f_i\rangle)_{t\geq 0}$. Since functions of the product form $F(x_1,\ldots,x_k) = f_1(x_1)\cdots f_k(x_k)$ are dense in the set of all continuous integrable functions $\psi : (\mathbb{R}^d)^k \to \mathbb{R}$, and since the family of test functions (20) is sufficient to characterise the law of an $\mathcal{M}_\lambda$-valued process by Lemma 2.1 in [VW15], the law of the limit will indeed be uniquely specified.

We begin with the case $d \geq 2$. Having (31), any limit of $(\overline{M}^n)_{n\geq 1}$ will be deterministic. It remains to identify the limit in $\mathbb{L}^1$ of the finite variation part. Specialising the computation of $\mathcal{A}^n$ above to the case $F = \mathrm{Id}$, we have

$$\begin{aligned} A_n(s) &= \frac{u\Gamma_R}{2}\langle \overline{w}_s^n, \Delta f\rangle + \mathcal{O}(n^{-2/3}) = \frac{u\Gamma_R}{2}\int_{\mathbb{R}^d\times\{0,1\}} \Delta f(x)\mathbf{1}_{\{0\}}(\kappa)\overline{M}_s^n(\mathrm{d}x,\mathrm{d}\kappa) + \mathcal{O}(n^{-2/3}) \\ &\to \frac{u\Gamma_R}{2}\langle w_s^\infty, \Delta f\rangle \qquad \text{as } n\to\infty, \end{aligned} \tag{32}$$

where the convergence is in distribution along the subsequence considered. These quantities being bounded by $(u\Gamma_R/2)\|\Delta f\|\mathrm{Vol}(S_f) + \mathcal{O}(n^{-2/3})$, independently of s, the convergence also happens in $\mathbb{L}^1$ norm. Next,

$$\begin{aligned}
D_n(s) &= \sigma u n^{d/3} \int_{\mathbb{R}^d} \Big\{ \overline{w}_s^n(x)^2 \langle \mathbf{1}_{B_n(x)}(1-w_s^n), (f(x) + \mathcal{O}(|y-x|)) \rangle \\
&\qquad\qquad - (1-\overline{w}_s^n(x)^2) \langle \mathbf{1}_{B_n(x)} w_s^n, (f(x)+\mathcal{O}(|y-x|))\rangle \Big\} \mathrm{d}x \\
&= \sigma u V_R \int_{\mathbb{R}^d} \left\{ \overline{w}_s^n(x)^2 (1-\overline{w}_s^n(x)) - (1-\overline{w}_s^n(x)^2)\overline{w}_s^n(x) \right\} f(x) \mathrm{d}x + \mathcal{O}(n^{-1/3}) \\
&= -\sigma u V_R \, \langle \overline{w}_s^n (1-\overline{w}_s^n), f \rangle + \mathcal{O}(n^{-1/3}). \qquad (33)
\end{aligned}$$

As above, the part of $D_n(s)$ which is linear in $\overline{w}_s^n$ converges (weakly and in $\mathbb{L}^1$) along the subsequence considered towards

$$-\sigma u V_R \, \langle w_s^\infty, f \rangle. \qquad (34)$$

We now would like to show that the 'quadratic' part of $D_n(s)$ converges to

$$\sigma u V_R \, \langle (w_s^\infty)^2, f \rangle.$$

Note that this is not a simple consequence of the weak convergence of $\overline{M}^n$ to $(M_t^\infty)_{t\geq 0}$, as $\langle (\overline{w}_s^n)^2, f \rangle$ cannot be written as an integral with respect to $\overline{M}_s^n$ or $(\overline{M}_s^n)^{\otimes 2}$. Instead, we shall approximate this expression by an integral with respect to $(\overline{M}_s^n)^{\otimes 2}$ and use the continuity estimates obtained in Proposition B.1 to bound the remaining terms. (The statement and proof of this proposition are postponed until Appendix B to ease the reading).

Let $\varepsilon \in (0, 1/2)$, and let p_ε be a continuous probability density function on $\mathbb{R}^d$ supported in $B(0,\varepsilon)$. For every $n \geq 1$ and $s \geq 0$, we have

$$\begin{aligned}
&\left| \langle (\overline{w}_s^n)^2, f \rangle - \langle (w_s^\infty)^2, f \rangle \right| \\
&\leq \left| \int_{\mathbb{R}^d} f(x) \overline{w}_s^n(x)^2 \mathrm{d}x - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{w}_s^n(x) \overline{w}_s^n(y) p_\varepsilon(y-x) \mathrm{d}y \mathrm{d}x \right| \\
&\quad + \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{w}_s^n(x) \overline{w}_s^n(y) p_\varepsilon(y-x) \mathrm{d}y \mathrm{d}x - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y-x) \mathrm{d}y \mathrm{d}x \right| \\
&\quad + \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y-x) \mathrm{d}y \mathrm{d}x - \int_{\mathbb{R}^d} f(x) w_s^\infty(x)^2 \mathrm{d}x \right|. \qquad (35)
\end{aligned}$$

The second term in the r.h.s. can be rewritten as

$$\begin{aligned}
&\int_{(\mathbb{R}^d \times \{0,1\})^2} f(x) p_\varepsilon(y-x) \mathbf{1}_{\{0\}}(\kappa) \mathbf{1}_{\{0\}}(\kappa') \overline{M}_s^n(\mathrm{d}y, \mathrm{d}\kappa') \overline{M}_s^n(\mathrm{d}x, \mathrm{d}\kappa) \\
&\to \int_{(\mathbb{R}^d \times \{0,1\})^2} f(x) p_\varepsilon(y-x) \mathbf{1}_{\{0\}}(\kappa) \mathbf{1}_{\{0\}}(\kappa') M_s^\infty(\mathrm{d}y, \mathrm{d}\kappa') M_s^\infty(\mathrm{d}x, \mathrm{d}\kappa) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) w_s^\infty(x) w_s^\infty(y) p_\varepsilon(y-x) \mathrm{d}y \mathrm{d}x
\end{aligned}$$

as n tends to infinity (since the mapping $(x,y) \mapsto f(x)p_\varepsilon(y-x)$ belongs to $C_c((\mathbb{R}^d)^2)$), and since these terms are bounded uniformly in n (and ε, s), this convergence also happens in $\mathbb{L}^1$ norm. That is, the second term in (35) tends to 0 in $\mathbb{L}^1$ norm.

Concerning the first term in the r.h.s. of (35), because $\overline{w}_s^n$ takes its values in $[0,1]$, we have, by Fubini's Theorem,

$$\mathbb{E}\Big[\Big|\int_{\mathbb{R}^d} f(x)\overline{w}_s^n(x)^2\mathrm{d}x - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)\overline{w}_s^n(x)\overline{w}_s^n(y)p_\varepsilon(y-x)\mathrm{d}y\mathrm{d}x\Big|\Big]$$
$$\leq \|f\| \int_{S_f}\int_{B(x,\varepsilon)} \mathbb{E}\big[\big|\overline{w}_s^n(x) - \overline{w}_s^n(y)\big|\big] p_\varepsilon(y-x)\mathrm{d}y\mathrm{d}x$$

By Proposition B.1, there exists $a, v, \lambda, C > 0$ independent of n such that for every $x, y \in \mathbb{R}^d$ satisfying $|x-y| < 1$ and every $s \in [0,T]$, we have

$$\mathbb{E}\big[\big|\overline{w}_s^n(x) - \overline{w}_s^n(y)\big|\big] \leq C\Big\{n^{-a} + \tau_n(x,y) + \Big(|x-y|^{1/4} + \tau_n(x,y)^{1/2}\Big)e^{\lambda(|x|+Rn^{-1/3})}$$
$$+ n^{(1-d)/6}\tau_n(x,y)^{(2-d)/4}\Big\},$$

where

$$\tau_n(x,y) = n^{-v} \vee |x-y|^{2/(d+1)}.$$

Thus, using the facts that the support S_f of f is compact, that p_ε is a probability density supported in $B(0,\varepsilon)$, and that $\tau_n(x,y) \leq \varepsilon^{2/(d+1)}$ for n large enough whenever $|x-y| \leq \varepsilon$, we can write that the first term in the r.h.s. of (35) is bounded by

$$C'\Big(n^{-a} + \varepsilon^{2/(d+1)} + \varepsilon^{1/4} + \varepsilon^{1/(d+1)} + n^{(1-d)/6}\varepsilon^{1/(d+1)}\Big).$$

Likewise, by taking $n \to \infty$ in Proposition B.1 (along the converging subsequence), we obtain that the last term in the r.h.s. of (35) is bounded by

$$C'\Big(\varepsilon^{1/4} + \varepsilon^{2/(d+1)} + \varepsilon^{1/(d+1)} + \varepsilon^{1/(d+1)}\mathbf{1}_{\{d=1\}}\Big).$$

Combining the above, we have that for every $\varepsilon \in (0,1/2)$,

$$\limsup_{n\to\infty} \mathbb{E}\big[\big|\langle(\overline{w}_s^n)^2, f\rangle - \langle(w_s^\infty)^2, f\rangle\big|\big] \leq C\big(\varepsilon^{1/4} + \varepsilon^{1/(d+1)}\big),$$

and letting ε tend to 0 we can conclude that the part of the expression (33) for $D_n(s)$ which is quadratic in $\overline{w}_s^n$ indeed converges in $\mathbb{L}^1$ towards

$$\sigma u V_R \,\langle(w_s^\infty)^2, f\rangle. \tag{36}$$

Combining (32), (34) and (36), and using the fact that the terms $B_n(s)$, $C_n(s)$ and $E_n(s)$ tend to zero uniformly in all possible values of $\overline{M}^n$, we conclude that any limit point of $(\overline{M}^n)_{n\geq 1}$ satisfies, for every $f \in C_c^\infty(\mathbb{R}^d)$, every family $\{w_t^\infty,\, t \geq 0\}$ of representatives of the density of M_t^∞ at each time t, and every $t \geq 0$,

$$\langle w_t^\infty, f\rangle = \langle w_0, f\rangle + \int_0^t \left\{\frac{u\Gamma_R}{2}\,\langle w_s^\infty, \Delta f\rangle - \sigma u V_R\,\langle w_s^\infty(1-w_s^\infty), f\rangle\right\}\mathrm{d}s. \tag{37}$$

Because the limit is a deterministic process, this property is sufficient to characterise its evolution.

Let us finally show that the system of equations (37) has at most one solution. As explained earlier in this proof, any test function of the form

$$M \mapsto \int_{(\mathbb{R}^d)^k} \psi(x_1,\dots,x_k)\bigg\{\prod_{i=1}^k w(x_i)\bigg\}\mathrm{d}x_1\cdots\mathrm{d}x_k,$$

considered in (15) (where as before w is any representative of the density of M), can be uniformly approximated by linear combinations of functions of the form $\prod_{i=1}^k\langle\cdot,f_i\rangle$ with $f_i\in C_c^\infty(\mathbb{R}^d)$ for every i. Thus, we can extend (37) to this more general class of functions. Then in Chapter 7 of [Lia09], it is proved that, when $\sigma=0$, any solution to (37) is dual, through the set of functional relations (15), to a system of independent Brownian motions with variance parameter $u\Gamma_R$, in which individuals never coalesce. This is easily modified to $\sigma>0$, in which case individuals branch into two at rate $u\sigma V_R$, independently of each other. Since the set of all test functions of the form considered in (15) is separating, this is enough to conclude that the system of equations (37) has at most one solution. Hence, this solution exists and the full sequence $(\overline{M}^n)_{n\geq 0}$ converges to it in distribution, as stated in Theorem 1.8(ii).

We now turn to the case $d=1$. As in the case $d\geq 2$, we know that for every $f\in C_c^\infty(\mathbb{R})$ (taking again $F=\mathrm{Id}$ so that $F''=0$),

$$\begin{aligned}W_t^n(f) &:= \langle\overline{w}_t^n,f\rangle - \langle\overline{w}_0^n,f\rangle - \int_0^t \mathcal{L}^n\Psi_{\mathrm{Id},\varphi_f}\big(M_s^n\big)\mathrm{d}s\\ &= \langle\overline{w}_t^n,f\rangle - \langle\overline{w}_0^n,f\rangle - \int_0^t\left\{\frac{u\Gamma_R}{2}\langle\overline{w}_s^n,\Delta f\rangle - \sigma u V_R\langle\overline{w}_s^n(1-\overline{w}_s^n),f\rangle\right\}\mathrm{d}s + \mathcal{O}(n^{-2/3})\end{aligned}$$

is a zero-mean martingale with quadratic variation

$$\begin{aligned}&u_n^2 n^{4/3}\int_0^t\int_{\mathbb{R}^d}\Big\{\overline{w}_s^n(x)(1+s_n\overline{w}_s^n(x))\langle\mathbf{1}_{B_n(x)}(1-w_s^n),f\rangle^2\\ &\qquad\qquad + (1-\overline{w}_s^n(x)+s_n(1-\overline{w}_s^n(x)^2))\langle\mathbf{1}_{B_n(x)}w_s^n,f\rangle^2\Big\}\mathrm{d}x\mathrm{d}s\\ &= u^2V_R^2\int_0^t\langle\overline{w}_s^n(1-\overline{w}_s^n),f^2\rangle\mathrm{d}s + \mathcal{O}(n^{-1/3}).\end{aligned}$$

In addition, (32), (34) and (36) still hold. Consequently, using the canonical decomposition of $W^n(f)$ into the sum of a time-changed Brownian motion and a pure-jump martingale whose jump sizes are uniformly bounded by $Cn^{-1/3}$ (and hence tends to zero), Theorem 6.3.4 in [EK86] implies that the sequence $(W^n(f))_{n\geq 1}$ converges weakly along the subsequence considered to the (time-changed Brownian motion) solution to the stochastic differential equation

$$dW_t = uV_R\sqrt{\langle w_t^\infty(1-w_t^\infty),f^2\rangle}\,dB_t^f,$$

where B^f denotes standard Brownian motion (the convergence of the quadratic variation is obtained using the same type of arguments as in (34) and (36)). As a consequence, any limit point $(M_t^\infty)_{t\geq 0}$ of $(\overline{M}^n)_{n\geq 1}$ satisfies the following system of stochastic differential equations: for every $f\in C_c^\infty(\mathbb{R}^d)$,

$$\mathrm{d}\langle w_t^\infty,f\rangle = \left\{\frac{u\Gamma_R}{2}\langle w_t^\infty,\Delta f\rangle - \sigma uV_R\langle w_t^\infty(1-w_t^\infty),f\rangle\right\}\mathrm{d}t + uV_R\sqrt{\langle w_t^\infty(1-w_t^\infty),f^2\rangle}\mathrm{d}B_t^f, \quad (38)$$

with initial value $\langle w_0, f\rangle$.

It remains to find the semi-martingale characterisation for $((\langle w_t^\infty, f\rangle^2)_{t\geq 0}$. Taking $F(x) = x^2$ in our previous calculations, we obtain that the finite variation part of $\langle \overline{w}_t^n, f\rangle^2$ can be written for any large $n \geq 1$ and $t \geq 0$:

$$\int_0^t \langle \overline{w}_s^n, f\rangle \Big\{ u\Gamma_R \langle \overline{w}_s^n, \Delta f\rangle - 2u\sigma V_R \langle \overline{w}_s^n(1-\overline{w}_s^n), f\rangle \Big\} \mathrm{d}s + u^2 V_R^2 \int_0^t \langle \overline{w}_s^n(1-\overline{w}_s^n), f^2\rangle \mathrm{d}s + \mathcal{O}\big(n^{-2/3}\big).$$

In addition, its quadratic variation part is equal to (taking advantage of the facts $s_n \to 0$, that the increment of $\langle w^n, \varphi_f\rangle$ during a jump is proportional to $u_n \to 0$, and finally that $B_n(x)$ is the ball of radius $Rn^{-1/3}$ around x to discard the negligible terms)

$$\begin{aligned}
&n^{4/3} \int_0^t \int_{\mathbb{R}^d} \Big\{ \overline{w}_s^n(x)(1+s_n \overline{w}_s^n(x)) \big[2\langle \overline{w}_s^n, f\rangle u_n \langle \mathbf{1}_{B_n(x)}(1-w), \varphi_f\rangle\big]^2 \\
&\qquad\qquad + \big(1-\overline{w}_s^n(x) + s_n(1-\overline{w}_s^n(x)^2)\big) \big[2\langle \overline{w}_s^n, f\rangle u_n \langle \mathbf{1}_{B_n(x)} w, \varphi_f\rangle\big]^2 \Big\} \mathrm{d}x \mathrm{d}s + \mathcal{O}\big(n^{-1/3}\big) \\
&= 4u^2 n^{2/3} \int_0^t \int_{\mathbb{R}^d} \langle \overline{w}_s^n, f\rangle^2 \Big\{ \overline{w}_s^n(x) \langle \mathbf{1}_{B_n(x)}(1-w), \varphi_f\rangle^2 \\
&\qquad\qquad\qquad + \big(1-\overline{w}_s^n(x)\big) \langle \mathbf{1}_{B_n(x)} w, \varphi_f\rangle^2 \Big\} \mathrm{d}x \mathrm{d}s + \mathcal{O}\big(n^{-1/3}\big) \\
&= 4u^2 V_R^2 \int_0^t \langle \overline{w}_s^n, f\rangle^2 \langle \overline{w}_s^n(1-\overline{w}_s^n)^2 + (\overline{w}_s^n)^2(1-\overline{w}_s^n), f^2\rangle \, \mathrm{d}s + \mathcal{O}\big(n^{-1/3}\big) \\
&= 4u^2 V_R^2 \int_0^t \langle \overline{w}_s^n, f\rangle^2 \langle \overline{w}_s^n(1-\overline{w}_s^n), f^2\rangle \, \mathrm{d}s + \mathcal{O}\big(n^{-1/3}\big).
\end{aligned}$$

By the same argument as for the processes of the form $(\langle \overline{w}_t^n, f\rangle)_{t\geq 0}$, we obtain that any limit point of $(\overline{M}^n)_{n\geq 1}$ must satisfy that for every $f \in C_c^\infty(\mathbb{R}^d)$, there exists a Brownian motion $\tilde{B}^f$ for which

$$\begin{aligned}
\mathrm{d}\big(\langle w_t^\infty, f\rangle^2\big) &= \langle w_t^\infty, f\rangle \Big\{ u\Gamma_R \langle w_t^\infty, \Delta f\rangle - 2\sigma u V_R \langle w_t^\infty(1-w_t^\infty), f\rangle \Big\} \mathrm{d}t \\
&\quad + u^2 V_R^2 \langle w_t^\infty(1-w_t^\infty), f^2\rangle \mathrm{d}t + 2uV_R \langle w_t^\infty, f\rangle \sqrt{\langle w_t^\infty(1-w_t^\infty), f^2\rangle} \, \mathrm{d}\tilde{B}_t^f,
\end{aligned} \tag{39}$$

with initial value $\langle w_0, f\rangle^2$. Using the fact that $\langle w, f\rangle\langle w, g\rangle = (1/2)(\langle w, (f+g)^2\rangle - \langle w, f^2\rangle - \langle w, g^2\rangle)$, we can deduce that the quadratic variation process of $(\langle w_t^\infty, f\rangle)_{t\geq 0}$ with $(\langle w_t^\infty, g\rangle)_{t\geq 0}$ is given, for every time $t \geq 0$, by

$$u^2 V_R^2 \int_0^t \langle w_s^\infty(1-w_s^\infty), fg\rangle \, \mathrm{d}s, \tag{40}$$

as in the statement of Theorem 1.8(i).

To prove uniqueness of the solution to this system, we use Itô's Formula and (40) to extend (38) to functions of the product form $\prod_{i=1}^k \langle \cdot, f_i\rangle$ and then to the full class of functions considered in (15), by the same density argument as before. Again in Chapter 7 of [Lia09], it is proved that in one dimension and when $\sigma = 0$, any solution to these equations is dual, through the set of relations (15), to a system of independent Brownian motions with variance parameter $u\Gamma_R$, in which, this time, particles coalesce pairwise at an instantaneous rate given by $u^2 V_R^2$ times the local time at 0 of their separation. As we mentioned earlier, this is easily modified to cover the case $\sigma > 0$, by imposing that individual lineages should also branch into two at rate $u\sigma V_R$. By the same chain of arguments as in the case $d \geq 2$, we can therefore conclude that the system of equations (38) has a unique solution, to which the full sequence $(\overline{M}^n)_{n\geq 0}$ thus converges in distribution as n tends to infinity. Theorem 1.8(i) is proved. □

Remark 4.1. *Liang's notation is very different from ours. To see that his process (with selection added and the coalescence rate multiplied by $u^2V_R^2$) and our limiting process do coincide, notice that $m(\mathrm{d}x) = \mathrm{d}x$ in our case and $\hat{X}_t(x) = w_t^\infty(x)\delta_0 + (1 - w_t^\infty(x))\delta_1$. Hence, taking $\chi(\kappa) = \mathbf{1}_0(\kappa) = \rho(\kappa)$ and $\psi(x) = f(x), \phi(x) = g(x)$ in Proposition 7.2 in [Lia09] indeed leads to*

$$\begin{aligned} \mathrm{d}[\mathcal{Z}^f, \mathcal{Z}^g]_t &= u^2V_R^2 \int_{\mathbb{R}^d} w_t^\infty(x)f(x)g(x)\mathrm{d}x - u^2V_R^2 \int_{\mathbb{R}^d} \big(w_t^\infty(x)\big)^2 f(x)g(x)\mathrm{d}x \\ &= u^2V_R^2 \int_{\mathbb{R}^d} w_t^\infty(x)\big(1 - w_t^\infty(x)\big)f(x)g(x)\mathrm{d}x. \end{aligned}$$

4.2 Proof of Theorem 1.10.

We divide the proof into two parts. The first, and simpler, shows that the only possible limit for $(\Xi^n)_{n\geq 1}$ is the system of branching and coalescing Brownian motions Ξ^∞. The second part, tightness of the sequence $(\Xi_t^n)_{n\geq 1}$, is rather more involved and will be broken into a number of smaller steps.

Recall that Ξ^n takes its values in the set $\mathcal{M}_p(\mathbb{R}^d)$ of all finite point measures on $\mathbb{R}^d$. The set of test functions

$$\Xi \mapsto \prod_{i=1}^{|\Xi|} f(\xi^i) = \exp\left\{ \int_{\mathbb{R}^d} (\ln f(x))\Xi(\mathrm{d}x)\right\}, \tag{41}$$

where $f \in C^\infty(\mathbb{R}^d)$ takes values in $[0, 1]$, is thus sufficient to characterise the law of Ξ.

Let us start with the following result.

Lemma 4.2. *The finite dimensional distributions of the system of scaled processes Ξ^n converge as $n \to \infty$ to those of the system of branching and coalescing Brownian motions Ξ^∞, described in the statement of Theorem 1.10. In particular, the only possible limit point for the sequence $(\Xi^n)_{n\geq 1}$ is Ξ^∞.*

Proof of Lemma 4.2.

Suppose first that the density ψ of the locations of the atoms of Ξ_0^n can be factorised as $\psi(x_1, \ldots, x_k) = \psi_1(x_1)\cdots\psi_k(x_k)$, with ψ_i a probability density function of class C^3 on $\mathbb{R}^d$ for every i. By Theorem 1.8, the rescaled and locally averaged forwards-in-time process $(\overline{M}_t^n)_{t\geq 0}$ converges to the process $(M_t^\infty)_{t\geq 0}$ for which, for every $f \in C_c^\infty(\mathbb{R}^d)$ and every set $\{w_t^\infty, t \geq 0\}$ of representatives of the density of each M_t^∞,

$$\left(\langle w_t^\infty, f\rangle - \langle w_0^\infty, f\rangle - \int_0^t \left\{ \frac{u\Gamma_R}{2} \langle w_s^\infty, \Delta f\rangle - u\sigma V_R \langle w_s^\infty(1 - w_s^\infty), f\rangle \right\} \mathrm{d}s \right)_{t\geq 0}$$

is a martingale, with quadratic variation 0 when $d \geq 2$, and

$$2u^2V_R^2 \int_0^t \langle w_s^\infty(1 - w_s^\infty), f^2\rangle \,\mathrm{d}s$$

at time $t \geq 0$ when $d = 1$. Furthermore, this description can be extended to any product of the form $\langle w_t^\infty, f_1\rangle\langle w_t^\infty, f_2\rangle \cdots \langle w_t^\infty, f_m\rangle$, for any $m \geq 1$ and $f_1, \ldots, f_m \in C_c^\infty(\mathbb{R}^d)$ (see the paragraph on the uniqueness of the limit in the proof of Theorem 1.8).

Now, as we have argued in the proof of Theorem 1.8 (based on results from Chapter 7 of [Lia09]), any solution to these equations is dual, through the set of functions (15), to a system of independent Brownian motions with variance parameter $u\Gamma_R$, which branch into two at rate $u\sigma V_R$ independently of each other, and coalesce pairwise when $d=1$, at a rate proportional to $u^2V_R^2$ times the local time at 0 of their separation.

Let us write $(M_t^{(n)})_{t\geq 0}$ for the unscaled SLFVS with parameters s_n, u_n, and $w_t^{(n)}$ for a representative of the density of $M_t^{(n)}$, for every $t\geq 0$. Using the approximation (30) to replace $\langle \overline{w}_t^n, \psi_i\rangle$ by $\langle w_{nt}^{(n)}(n^{1/3}\cdot), \psi_i\rangle + \mathcal{O}(n^{-2/3})$ on the third line, and conversely on the last line, together with Fubini's Theorem and the duality formula (15), we can write that for any density w_0 (or measure M_0) to which the sequence of initial values of $\overline{M}^n$ converges as $n\to\infty$, we have

$$
\begin{aligned}
&\mathbb{E}\Big[\prod_{i=1}^k\Big(\int_{\mathbb{R}^d\times\{0,1\}}\psi_i(x_i)\mathbf{1}_{\{0\}}(\kappa_i)\overline{M}_t^n(\mathrm{d}x_i,\mathrm{d}\kappa_i)\Big)\Big]\\
&\quad=\mathbb{E}\Big[\int_{(\mathbb{R}^d)^k}\psi_1(x_1)\cdots\psi_k(x_k)\Big\{\prod_{i=1}^k\overline{w}_t^n(x_i)\Big\}\mathrm{d}x_1\ldots\mathrm{d}x_k\Big]\\
&\quad=\mathbb{E}\Big[\int_{(\mathbb{R}^d)^k}\psi_1(x_1)\cdots\psi_k(x_k)\Big\{\prod_{i=1}^k w_{nt}^{(n)}(n^{1/3}x_i)\Big\}\mathrm{d}x_1\ldots\mathrm{d}x_k\Big]+\mathcal{O}(n^{-2/3})\\
&\quad=n^{-dk/3}\mathbb{E}\Big[\int_{(\mathbb{R}^d)^k}\psi_1(n^{-1/3}x_1)\cdots\psi_k(n^{-1/3}x_k)\Big\{\prod_{i=1}^k w_{nt}^{(n)}(x_i)\Big\}\mathrm{d}x_1\ldots\mathrm{d}x_k\Big]+\mathcal{O}(n^{-2/3})\\
&\quad=n^{-dk/3}\int_{(\mathbb{R}^d)^k}\psi_1(n^{-1/3}x_1)\cdots\psi_k(n^{-1/3}x_k)\mathbb{E}_{\Xi[x_1,\ldots,x_k]}\Big[\prod_{j=1}^{N_{nt}}w_0^{(n)}(\xi_{nt}^j)\Big]\mathrm{d}x_1\ldots\mathrm{d}x_k+\mathcal{O}(n^{-2/3})\\
&\quad=\int_{(\mathbb{R}^d)^k}\psi_1(x_1)\cdots\psi_k(x_k)\mathbb{E}_{\Xi[n^{1/3}x_1,\ldots,n^{1/3}x_k]}\Big[\prod_{j=1}^{N_{nt}}w_0^{(n)}(n^{1/3}(n^{-1/3}\xi_{nt}^j))\Big]\mathrm{d}x_1\ldots\mathrm{d}x_k+\mathcal{O}(n^{-2/3})\\
&\quad=\int_{(\mathbb{R}^d)^k}\psi_1(x_1)\cdots\psi_k(x_k)\mathbb{E}_{\Xi[x_1,\ldots,x_k]}\Big[\prod_{j=1}^{N_t^n}\overline{w}_0^n(\xi_t^{n,j})\Big]\mathrm{d}x_1\ldots\mathrm{d}x_k+\mathcal{O}(n^{-2/3})\\
&\quad=\mathbb{E}_{\Xi_0^n}\Big[\prod_{j=1}^{N_t^n}\overline{w}_0^n(\xi_t^{n,j})\Big]+\mathcal{O}(n^{-2/3}).
\end{aligned}
\tag{42}
$$

Now the expression in the l.h.s. of (42) converges to the corresponding expression for M^∞ as $n\to\infty$. In addition, as mentioned in the paragraph around (41), test functions of the form used in the r.h.s. of (42), with $\overline{w}_0^n$ replaced by its limit w_0, are sufficient to characterise the law of a random particle configuration Ξ. We can therefore conclude that the one-dimensional distributions of $(\Xi_t^n)_{t\geq 0}$ converge to those of $(\Xi_t^\infty)_{t\geq 0}$. The generalisation to the finite-dimensional distributions is straightforward since the duality formula (15) holds on any time interval $[s,t]$ (if we replace w_0 by w_s and ξ_t^j by ξ_{t-s}^j).

Finally, since linear combinations of functions of the product form $x\mapsto\psi_1(x_1)\cdots\psi_k(x_k)$, with ψ_i a probability density function of class C^3 for every i, are dense in the set of continuous probability densities ψ on $(\mathbb{R}^d)^k$, an analogue of Relation (42) in the limit as $n\to\infty$ can be established for this more general class of initial densities ψ. The same chain of arguments are then sufficient to conclude the proof of Lemma 4.2. □

Tightness

We now show tightness of the sequence $(\Xi^n)_{n\geq 1}$. To ease the notation, we write $\mathbb{P}_\psi$ for the probability measure on $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$ under which the locations of the atoms of each Ξ_0^n have density ψ. We use the Aldous-Rebolledo criterion based on stopping times (see again [Ald78, Reb80]), with the family of real-valued functions described in (41).

Fix $T>0$ and $f\in C^\infty(\mathbb{R}^d)$ with values in $(0,1]$, and suppose that $(\tau_n)_{n\geq 1}$ is any sequence of stopping times bounded by $T-\delta_0$ for some small $\delta_0>0$. We must show that $(\Xi_t^n)_{n\geq 1}$ is tight for every $t\geq 0$, and that for every $\varepsilon>0$, there exists $\delta=\delta(f,T,\psi,\varepsilon)$ such that

$$\limsup_{n\to\infty}\mathbb{P}_\psi\left[\sup_{0\leq t\leq\delta}\left|\prod_{i=1}^{N^n_{\tau_n+t}} f(\xi^{n,i}_{\tau_n+t})-\prod_{i=1}^{N^n_{\tau_n}} f(\xi^{n,i}_{\tau_n})\right|>\varepsilon\right]\leq\varepsilon. \tag{43}$$

We shall proceed in a number of steps. First we control the maximum number of particles in Ξ_t^n up to time T. Conditional on this, it is easy to control the probability that there is a branch in an interval of length δ. If we can also show that with high probability there is no coalescence (so that the number of particles in the system does not change), then the problem is reduced to controlling the jumps in a random walk. The most involved step, which is the substance of Proposition 4.4, is showing that there is no accumulation of coalescence events. In what follows, we write $|\Xi_t^n|$ for the total mass of (or number of particles in) Ξ_t^n.

Lemma 4.3. *Given $\varepsilon>0$, there exists $K>0$ such that*

$$\mathbb{P}_\psi\left[\sup_{0\leq t\leq T}|\Xi_t^n|>K\right]\leq\frac{\varepsilon}{5}.$$

Proof of Lemma 4.3.

Recall that two particles are created when at least one of the extant particles is affected by a selective event. For a given particle of Ξ^n, this happens at rate $ns_nV_Ru_n=u\sigma V_R$. Furthermore, the presence of more than one particle in the area affected by the event does not speed up the branching. Consequently, the number of particles in $(\Xi_t^n)_{t\geq 0}$ is stochastically bounded by the number of particles in a continuous-time branching process in which particles split (independently of one another) into two offspring at rate $u\sigma V_R$. Since the initial value, Ξ_0^n, has $k<\infty$ particles, we conclude that there exists $K\in\mathbb{N}$ such that for every $n\in\mathbb{N}$,

$$\mathbb{P}_\psi\left[\sup_{0\leq t\leq T}|\Xi_t^n|>K\right]\leq\frac{\varepsilon}{5},$$

as required. □

From now on, all our calculations will proceed conditional on the event $A_n=\{\sup_{0\leq t\leq T}|\Xi_t^n|\leq K\}$. From our reasoning above, we already see that for any $t\in[0,T]$, conditional on A_n, the probability that at least one particle is created during the time interval $(t,t+\delta]$ is bounded by

$$K\,\mathbb{P}_\psi\big[\text{a given particle branches in }(t,t+\delta]\big]\leq K\big(1-e^{-u\sigma V_R\delta}\big)\leq u\sigma KV_R\,\delta.$$

This bound is uniform and so we see that there exists $\delta_1>0$ such that for every $n\geq 1$,

$$\mathbb{P}_\psi\big[\text{at least 1 particle created in }(\tau_n,\tau_n+\delta_1]\,;\,A_n\big]\leq\frac{\varepsilon}{5}. \tag{44}$$

We also want to control the probability of coalescence events. Because of the calculation above, it is enough to do so in the absence of branching.

Proposition 4.4. *Let B^c_δ denote the event that there is no branching event in $(\tau_n, \tau_n+\delta]$. There exists $\delta_2 \in (0, \delta_1]$ such that*

$$\mathbb{P}_\psi[\text{at least 1 coalescence in } (\tau_n, \tau_n + \delta_2]\,;\, A_n, B^c_{\delta_2}] \leq \frac{\varepsilon}{5}.$$

Before proving Proposition 4.4, let us turn to the final ingredient in the proof and control the jumps of a single particle.

From the description in Section 1.2, after rescaling of time and space, $\xi^{n,1}$ jumps at rate $nu_nV_R(1+s_n) = n^{2/3}uV_R(1+o(1))$, to a new location whose distribution is symmetric about its current location. Furthermore, the locations of the particle both before and after the jump belong to the same ball of radius $Rn^{-1/3}$, and so the length of the jump is bounded by $2Rn^{-1/3}$. Doob's Maximal Inequality and standard estimates for the variance of a compound Poisson process then imply that there exists $C_1 > 0$ such that for every n, any $s, \eta > 0$, and every stopping time T_n,

$$\mathbb{P}_\psi\left[\sup_{t\in[0,s]} \left|\xi^{n,1}_{T_n+t} - \xi^{n,1}_{T_n}\right| > \eta\right] \leq \frac{C_1}{\eta^2}\, s, \tag{45}$$

where we have used the strong Markov property of $\xi^{n,1}$ at time T_n. From this, we can draw two conclusions. First, taking $s = T$ and $T_n = 0$, we can find a compact set $E \subset \mathbb{R}^d$ such that for every $n \geq 1$,

$$\mathbb{P}_\psi\left[\sup_{t\in[0,T]} \Xi^n_t(E^c) > 0;\, A_n\right] \leq \frac{\varepsilon}{5}. \tag{46}$$

Indeed, since ψ is integrable, there exists a compact set $\tilde{E}$ such that $\mathbb{P}_\psi(\Xi^n_0(\tilde{E}^c) > 0) < \varepsilon/10$. Conditionally on all the initial particles belonging to $\tilde{E}$, by (45) we can then find a radius $\eta > 0$ such that the probability that any of the (at most) K particles leaves $E = \tilde{E} + B(0, \eta)$ is less than $\varepsilon/10$. Let us write E_n for the event that $\sup_{t\in[0,T]} \Xi^n_t(E^c) > 0$ (i.e., at least one of the particles escapes from E before time T). Note that the property that

$$\liminf_{n\geq 1} \mathbb{P}(A_n \cap E^c_n) \geq 1 - \frac{2\varepsilon}{5}$$

implies the tightness of $(\Xi^n_t)_{n\geq 1}$ for every $t \in [0, T]$.

Second, conditional on the number of individuals not changing during a time interval of length δ, we can index the particles of $\Xi^n_{\tau_n}$ and $\Xi^n_{\tau_n+\delta}$ by a common indexing set which we denote I_n. Under this assumption and on the event E_n, a Taylor expansion of f yields

$$\left|\prod_{i\in I_n} f(\xi^{n,i}_{\tau_n+t}) - \prod_{i\in I_n} f(\xi^{n,i}_{\tau_n})\right| \leq C\left(\sup_E \|\nabla f\|\right) \sum_{i\in I_n} \left|\xi^{n,i}_{\tau_n+t} - \xi^{n,i}_{\tau_n}\right|, \tag{47}$$

for some $C > 1$, where the supremum of $\|\nabla f\|$ over the compact set E is finite. Together with (45) and the choice $s = \delta$, $T_n = \tau_n$ and $\eta = \varepsilon/(KC\sup_E \|\nabla f\|)$, this shows that there exists $\delta_3 \in (0, \delta_2]$ such that for n large enough, writing C^c_δ for the event that there is no coalescence in $[\tau_n, \tau_n + \delta]$,

$$\mathbb{P}_\psi\left[\sup_{t\in[0,\delta_3]} \left|\prod_{i\in I_n} f(\xi^{n,i}_{\tau_n+t}) - \prod_{i\in I_n} f(\xi^{n,i}_{\tau_n})\right| > \varepsilon\,;\, A_n, B^c_{\delta_3}, C^c_{\delta_3}, E^c_n\right] \leq \frac{KC_1}{\eta^2}\,\delta_3 \leq \frac{\varepsilon}{5}. \tag{48}$$

Combining Lemma 4.3, (44), Proposition 4.4, (46) and (48), we obtain (43) with $\delta = \delta_3$.

It remains to prove Proposition 4.4. Let us remark that it is not enough to consider lineages at an initial separation of order $\mathcal{O}(1)$ (or $\mathcal{O}(n^{1/3})$ before rescaling). In particular, when two particles are created through a selective event, their (rescaled) initial distance is of order $\mathcal{O}(n^{-1/3})$ and so we also need to control the coalescence of particles starting from very small initial separations.

Proof of Proposition 4.4.

It suffices to consider just two particles and find $\delta_2 > 0$ such that the probability that they coalesce in a time interval of length δ_2 is bounded by $\varepsilon/(3K(K-1))$, irrespective of their initial separation. Once this bound has been established, we can write

$$\mathbb{P}_\psi\big[\text{at least 1 coalescence in } (\tau_n, \tau_n + \delta_2]\, ;\, A_n, B^c_{\delta_2}\big] \leq \frac{K(K-1)}{2} \frac{\varepsilon}{3K(K-1)} = \frac{\varepsilon}{6}, \tag{49}$$

since, on the event A_n, there are at most $K(K-1)/2$ pairs of particles at any time.

Recall that before scaling, each lineage jumps at rate proportional to $u_n = un^{-1/3}$. This makes it convenient to work in the timescale $(n^{1/3}t,\, t \geq 0)$ and without rescaling space. We shall write $\tilde{\xi}^{n,i}_t = \xi^i_{n^{1/3}t}$, $i \in \{1, 2\}$.

When $\tilde{\xi}^{n,1}$ and $\tilde{\xi}^{n,2}$ are separated by more than $2R$, they cannot be contained in the same reproduction event, and so they evolve independently of one another. The ith lineage jumps at rate $n^{1/3}u_n V_R(1 + s_n) = uV_R(1 + o(1))$ to a new location, which is uniformly distributed over the ball $B(Z, R)$, where Z itself is chosen uniformly at random from $B(\tilde{\xi}^{n,i}, R)$. In what follows, we only need that the jump made by each lineage is an independent realisation of a random variable X taking values in $B(0, 2R)$, whose distribution is symmetric about the origin.

On the other hand, when $|\tilde{\xi}^{n,1} - \tilde{\xi}^{n,2}| < 2R$, the two particles can both lie in a region affected by a given reproduction event and their jumps become correlated. In particular, if they are both affected by this event, they merge together. The generator of $((\tilde{\xi}^{n,1}_t, \tilde{\xi}^{n,2}_t))_{t\geq 0}$ takes the form

$$\begin{aligned}
&u(1 + s_n) \int_{B(\tilde{\xi}^1,R)\setminus B(\tilde{\xi}^2,R)} \int_{B(x,R)} \frac{1}{V_R} \big(f(z, \tilde{\xi}^2) - f(\tilde{\xi}^1, \tilde{\xi}^2)\big)\mathrm{d}z\mathrm{d}x \\
&+ u(1 + s_n) \int_{B(\tilde{\xi}^2,R)\setminus B(\tilde{\xi}^1,R)} \int_{B(x,R)} \frac{1}{V_R} \big(f(\tilde{\xi}^1, z) - f(\tilde{\xi}^1, \tilde{\xi}^2)\big)\mathrm{d}z\mathrm{d}x \\
&+ u(1 - un^{-1/3})(1 + s_n) \int_{B(\tilde{\xi}^1,R)\cap B(\tilde{\xi}^2,R)} \int_{B(x,R)} \frac{1}{V_R} \big(f(z, \tilde{\xi}^2) + f(\tilde{\xi}^1, z) - 2f(\tilde{\xi}^1, \tilde{\xi}^2)\big)\mathrm{d}z\mathrm{d}x \\
&+ u^2n^{-1/3}(1 + s_n) \int_{B(\tilde{\xi}^1,R)\cap B(\tilde{\xi}^2,R)} \int_{B(x,R)} \frac{1}{V_R} \big(f(z, z) - f(\tilde{\xi}^1, \tilde{\xi}^2)\big)\mathrm{d}z\mathrm{d}x.
\end{aligned}$$

We can think of this as composed of two parts: the process $((\hat{\xi}^{n,1}_t, \hat{\xi}^{n,2}_t))_{t\geq 0}$ whose generator is determined by the first three lines above, on top of which a coalescence event occurs at instantaneous rate $u^2n^{-1/3}(1 + s_n)V_R(0, \hat{\xi}^{n,1}_t - \hat{\xi}^{n,2}_t)$ (recall that $V_R(0, a)$ is the volume of the intersection $B(0, R) \cap B(a, R)$).

With this description, the probability that the two particles have not coalesced by time $\delta n^{2/3}$ (which corresponds to a time span of δ on the timescale of $\xi^{n,i}$) is given by

$$\mathbb{P}_\psi\big[\tilde{T} > \delta n^{2/3}\big] = \mathbb{E}_\psi\bigg[\exp\bigg\{ -\frac{u^2(1 + s_n)}{n^{1/3}} \int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}^{n,1}_s - \hat{\xi}^{n,2}_s)\, \mathrm{d}s\bigg\}\bigg], \tag{50}$$

where we have written $\tilde{T}$ for the coalescence time of the two particles.

Since $V_R(0,x) = 0$ when $x \geq 2R$, it just remains to establish how much time $\hat{\xi}^{n,1} - \hat{\xi}^{n,2}$ spends in the ball $B(0, 2R)$ by time $\delta n^{2/3}$. To do this, we define two sequences of stopping times, $(\sigma_k^n)_{k\geq 1}$ and $(\tau_k^n)_{k\geq 1}$ by

$$\sigma_1^n = \inf\{t \geq 0 \,:\, |\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}| \leq 2R\}, \qquad \tau_1^n = \inf\{t \geq \sigma_1^n \,:\, |\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}| > 2R\},$$

and for every $k \geq 1$,

$$\sigma_k^n = \inf\{t \geq \tau_{k-1}^n \,:\, |\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}| \leq 2R\}, \qquad \tau_k^n = \inf\{t \geq \sigma_k^n \,:\, |\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}| > 2R\}.$$

Now, we have the following result.

Lemma 4.5. *There exists $\mathcal{C} > 0$ such that for every $n, k \geq 1$,*

$$\mathbb{E}_\psi\big[\tau_k^n - \sigma_k^n\big] \leq \mathcal{C}.$$

In words, although the two particles are correlated when they are close together, each 'incursion' of $\hat{\xi}^{n,1} - \hat{\xi}^{n,2}$ inside $B(0, 2R)$ lasts only $\mathcal{O}(1)$ units of time, uniformly in n. The proof of Lemma 4.5 is similar to that of Lemma 6.6 in [BEV10] (based on the facts that the difference walk jumps at a rate bounded from below by a positive constant, independent of its current value, and that the probability that this jump leads to a sufficient increase of their separation for $\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}$ to leave $B(0, 2R)$ is also bounded from below by a positive constant). Therefore, we omit it here.

Outside $B(0, 2R)$, the difference $\hat{\xi}_t^{n,1} - \hat{\xi}_t^{n,2}$ has the same law as a symmetric random walk, with jumps of size at most $2R$, jumping at rate $2uV_R(1 + s_n)$. Its behaviour will be determined by the spatial dimension.

$\mathbf{d} \geq 3$: When $d \geq 3$, transience of the random walk guarantees that the number of times $\hat{\xi}^{n,1} - \hat{\xi}^{n,2}$ returns to $B(0, 2R)$ is a.s. finite. Since the parameter n appears only in the jump rates and not in the embedded chain of locations (during an excursion outside $B(0, 2R)$), the probability that the difference walk enters $B(0, 2R)$ at least k times decays to 0, uniformly in n, as $k \to \infty$. Together with Lemma 4.5 and the fact that $V_R(0, \cdot)$ is bounded, this shows that for every $\eta > 0$,

$$\lim_{n\to\infty} \mathbb{P}_\psi\bigg[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2})\, \mathrm{d}s > \eta\, \frac{n^{1/3}}{u^2(1 + s_n)}\bigg] = 0$$

As a consequence, coming back to (50) and choosing η small enough that $\mathbb{P}[\mathrm{Exp}(1) \leq \eta] \leq \varepsilon/(3K(K-1))$, we can conclude that for any $\delta > 0$,

$$\limsup_{n\to\infty} \mathbb{P}_\psi\big[\tilde{T} \leq \delta n^{2/3}\big] \leq \frac{\varepsilon}{3K(K-1)}. \tag{51}$$

$\mathbf{d} = 2$: When $d = 2$, we claim that there exists $\mathcal{C}' > 0$, independent of n, such that for every x_1, x_2 with $|x_1 - x_2| > 2R$,

$$\mathbb{P}_{\{x_1,x_2\}}\big[\sigma_1^n > \delta n^{2/3}\big] \geq \frac{\mathcal{C}'}{\log(\delta n^{2/3})},$$

where we have written $\mathbb{P}_{\{x_1,x_2\}}$ for the probability measure under which the two particles start at locations x_1, x_2. The proof of this claim is very similar to the beginning of the proof of Lemma 4.2 in [BEV12], and so we only sketch the main ideas. We can a.s. embed the trajectories of the difference process $\hat{\xi}^{n,1}_t - \hat{\xi}^{n,2}_t$ into the trajectories of a two-dimensional Brownian motion, in the same spirit as Skorokhod's embedding in one dimension (see e.g. [Bil95]). Now, since the jumps of the difference process (when outside $B(0, 2R)$) are rotationally invariant, we have

$$\inf_{|x_1-x_2|>2R} \mathbb{P}_{\{x_1,x_2\}}\big[\hat{\xi}^{n,1} - \hat{\xi}^{n,2} \text{ leaves } B(0,4R) \text{ before entering } B(0,2R)\big] > 0,$$

and the result then follows from that for Brownian motion, namely Theorem 2 in [RR66] applied with $a = 2R$ and $r \geq 4R$. As a consequence, the number N^n_E of excursions outside $B(0, 2R)$ that the difference walk makes before starting an excursion of (time) length at least $\delta n^{2/3}$ is stochastically bounded by a geometric random variable with success probability $C/\log(\delta n^{2/3})$. Now, once the difference walk has started such a long excursion (say, the kth one), it is sure not to come back within $B(0, 2R)$ before time $\delta n^{2/3}$ and the number of incursions in $B(0, 2R)$ in the time interval $[0, \delta n^{2/3}]$ is bounded by k. Thus, fixing $\eta > 0$ as before and observing that $V_R(x, y)$ is bounded by the volume V_R of a ball of radius R, we obtain that

$$\begin{aligned}
&\mathbb{P}_\psi\bigg[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}^{n,1}_s - \hat{\xi}^{n,2}_s)\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\bigg] \\
&\qquad \leq \mathbb{P}_\psi\big[N^n_E > C^n_E \log(\delta n^{2/3})\big] + \mathbb{P}_\psi\bigg[\sum_{k=1}^{\lceil C^n_E \log(\delta n^{2/3})\rceil} (\tau^n_k - \sigma^n_k) > \eta\,\frac{n^{1/3}}{u^2(1+s_n)V_R}\bigg] \\
&\qquad \leq e^{-C^n_E \mathcal{C}'} + \frac{u^2(1+s_n)V_R}{\eta n^{1/3}}\, C^n_E \log(\delta n^{2/3})\mathcal{C},
\end{aligned}$$

where the last inequality uses the stochastic bound of N^n_E first, and then Markov's inequality. Choosing $C^n_E = \log n$, for instance, we deduce that for any $\delta > 0$,

$$\lim_{n\to\infty} \mathbb{P}_\psi\bigg[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}^{n,1}_s - \hat{\xi}^{n,2}_s)\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\bigg] = 0,$$

and we conclude as in (51).

d = 1: Finally, when $d = 1$ it is shown in [PS71] that there exists $\mathcal{C}' > 0$ such that for every x_1, x_2 such that $|x_1 - x_2| > 2R$,

$$\mathbb{P}_{\{x_1,x_2\}}[\sigma^n_1 > \delta n^{2/3}] \geq \frac{\mathcal{C}'}{\sqrt{\delta}\, n^{1/3}}.$$

Proceeding as before, and with the same notation, we therefore have

$$\begin{aligned}
&\mathbb{P}_\psi\bigg[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}^{n,1}_s - \hat{\xi}^{n,2}_s)\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\bigg] \\
&\qquad \leq \mathbb{P}_\psi\big[N^n_E > C^n_E\sqrt{\delta} n^{1/3}\big] + \mathbb{P}_\psi\bigg[\sum_{k=1}^{\lceil C^n_E\sqrt{\delta} n^{1/3}\rceil} (\tau^n_k - \sigma^n_k) > \eta\,\frac{n^{1/3}}{u^2(1+s_n)\|V_R\|}\bigg] \\
&\qquad \leq e^{-C^n_E \mathcal{C}'} + \frac{u^2(1+s_n)\|V_R\|}{\eta n^{1/3}}\, C^n_E\sqrt{\delta} n^{1/3}\mathcal{C}.
\end{aligned}$$

Choosing C^n_E to be a constant large enough for the first term to be less than $\varepsilon/(9K(K-1))$, and then $\delta_3 > 0$ small enough for the second term to be less than $\varepsilon/(9(K(K-1))$, and finally taking η small enough, we obtain that for any $\delta \leq \delta_3$,

$$\mathbb{P}_\psi\big[\tilde{T} \leq \delta n^{2/3}\big] \leq \mathbb{P}_\psi\bigg[\int_0^{\delta n^{2/3}} V_R(0, \hat{\xi}_s^{n,1} - \hat{\xi}_s^{n,2})\,\mathrm{d}s > \eta\,\frac{n^{1/3}}{u^2(1+s_n)}\bigg] + \mathbb{P}[\mathrm{Exp}(1) \leq \eta]$$
$$\leq \frac{\varepsilon}{9K(K-1)} + \frac{\varepsilon}{9K(K-1)} + \frac{\varepsilon}{9K(K-1)} = \frac{\varepsilon}{3K(K-1)}. \tag{52}$$

We have now proved the desired bound for the probability of a coalescence in any dimension and the proof of Proposition 4.4 is complete. □

5 Convergence of the SLFVS and its dual - the stable radius case

Proof of Theorem 1.9.

1) Tightness.

We shall use the same method as in the proof of Theorem 1.8, but the computations required will be different. Recall the notation $\Theta^+_{x,r,u_n}(w)$ and $\Theta^-_{x,r,u_n}(w)$ from (5). The generator of the unscaled process with parameters acting on functions of the form $\Psi_{F,f}$ (see (7)) is given by

$$\mathcal{L}\Psi_{F,f}(M) = \int_{\mathbb{R}^d}\int_1^\infty\int_{B(x,r)^2}\frac{1}{V_r^2}\Big\{w(y)(1+s_nw(z))\big[F(\langle\Theta^+_{x,r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]$$
$$+ (1-w(y)+s_n(1-w(y)w(z)))\big[F(\langle\Theta^-_{x,r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]\Big\}\mathrm{d}y\mathrm{d}z\mu(\mathrm{d}r)\mathrm{d}x,$$

where, as usual now, w is a representative of the density of M. To make the expressions easier to read, we retain the notation β, γ and δ from (17). As we did in the fixed radius case, let us consider the process $(M_t^n)_{t\geq 0}$ whose density at time t is $w_t^n(\cdot) := w_{nt}(n^\beta\cdot)$. The generator of this Markov process is then given by

$$\begin{aligned}
&\mathcal{L}^n\Psi_{F,f}(M)\\
&= n\int_{\mathbb{R}^d}\int_1^\infty\int_{B(x,r)^2}\frac{1}{V_r^2}\Big\{w(n^{-\beta}y)\big(1+s_nw(n^{-\beta}z)\big)\big[F(\langle\Theta^+_{n^{-\beta}x,n^{-\beta}r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]\big)\\
&\quad + (1-w(n^{-\beta}y)+s_n(1-w(n^{-\beta}y)w(n^{-\beta}z)))\big[F(\langle\Theta^-_{n^{-\beta}x,n^{-\beta}r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]\Big\}\\
&\qquad\qquad \mathrm{d}y\mathrm{d}z\mu(\mathrm{d}r)\mathrm{d}x\\
&= n^{1-\beta\alpha}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r^2}\Big\{w(y)(1+s_nw(z))\big[F(\langle\Theta^+_{x,r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]\\
&\quad + \big(1-w(y)+s_n(1-w(y)w(z))\big)\big[F(\langle\Theta^-_{x,r,u_n}(w),f\rangle) - F(\langle w,f\rangle)\big]\Big\}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x.
\end{aligned} \tag{53}$$

To simplify notation, we shall show the tightness of $(F(\langle w^n,f\rangle))_{n\geq 1}$. This implies that the sequence $(F(\langle\overline{w}^n,f\rangle))_{n\geq 1}$ in which we are interested is tight on replacing f by φ_f defined by

$$\varphi_f(x) = \frac{n^{d\beta}}{V_1}\int_{B(x,n^{-\beta})} f(y)\,\mathrm{d}y, \tag{54}$$

and exploiting the bound

$$\langle\overline{w}^n,f\rangle = \langle w^n,\varphi_f\rangle = \langle w^n,f\rangle + \delta_n(w^n,f)\,n^{-2\beta}, \tag{55}$$

for every $f \in C_c^\infty(\mathbb{R}^d)$, where $|\delta_n(w,f)|$ is bounded by a constant $\eta(f) > 0$ uniformly in n, w.

Just as in the previous section, we write $(\mathcal{A}_t^n)_{t\geq 0}$ for the finite variation part of $(F(\langle w_t^n, f\rangle))_{t\geq 0}$ and $(\mathcal{Q}_t^n)_{t\geq 0}$ for the quadratic variation of its martingale part. As before, it is convenient to split $\mathcal{L}^n\Psi_{F,f}(M)$ into its neutral and selective components. Using a Taylor expansion of the function F, we obtain that the neutral part is equal to

$$
\begin{aligned}
n^{1-\beta\alpha}&F'(\langle w,f\rangle)\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)}\frac{1}{V_r}\Big[w(y)\big\langle\Theta^+_{x,r,u_n}(w)-w,f\big\rangle \\
&\qquad\qquad\qquad + (1-w(y))\big\langle\Theta^-_{x,r,u_n}(w)-w,f\big\rangle\Big]\,\mathrm{d}y\mathrm{d}r\mathrm{d}x \\
&+ n^{1-\beta\alpha}\frac{F''(\langle w,f\rangle)}{2}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)}\frac{1}{V_r}\Big[w(y)\big\langle\Theta^+_{x,r,u_n}(w)-w,f\big\rangle^2 \\
&\qquad\qquad\qquad + (1-w(y))\big\langle\Theta^-_{x,r,u_n}(w)-w,f\big\rangle^2\Big]\,\mathrm{d}y\mathrm{d}r\mathrm{d}x+\varepsilon_n \\
=\ & n^{1-\beta\alpha-\gamma}uF'(\langle w,f\rangle)\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\,w(y)(f(z)-f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x \\
&+ n^{1-\beta\alpha-2\gamma}\frac{u^2}{2}F''(\langle w,f\rangle)\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)}\frac{1}{V_r}\left\{w(y)\langle\mathbf{1}_{B(x,r)}(1-w),f\rangle^2\right. \\
&\qquad\qquad\qquad \left.+ (1-w(y))\langle\mathbf{1}_{B(x,r)}w,f\rangle^2\right\}\mathrm{d}y\mathrm{d}r\mathrm{d}x+\varepsilon_n, \qquad (56)
\end{aligned}
$$

with

$$|\varepsilon_n| \leq n^{1-\alpha\beta-3\gamma}\frac{u^3C_F}{3!}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)}\frac{2}{V_r}\langle\mathbf{1}_{B(x,r)},f\rangle^3\mathrm{d}y\mathrm{d}r\mathrm{d}x,$$

where the constant C_F is the supremum of $F^{(3)}$ over the bounded set in which its argument takes its values (recall that $f \in C_c^\infty(\mathbb{R}^d)$). Consider the first term in the right hand side of (56). Since $1-\alpha\beta-\gamma = 0$, $n^{1-\beta\alpha-\gamma} = 1$. We split the integral over the radii into the sum of the integrals over $[n^{-\beta}, 1]$ and $[1,\infty]$. By using a Taylor expansion of f and a symmetry argument to cancel the integral of $(z-y)\mathrm{d}z$, we obtain that

$$
\begin{aligned}
&\left|uF'(\langle w,f\rangle)\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{1}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\,w(y)(f(z)-f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\right| \\
&\qquad\leq C\left|\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{1}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\,|z-y|^2\mathbf{1}_{\{B(x,r)\cap S_f\neq\emptyset\}}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\right| \\
&\qquad\leq C'\mathrm{Vol}(S_f+B(0,1))\int_{n^{-\beta}}^{1}\frac{1}{r^{d+1+\alpha}}\,r^{d+2}\mathrm{d}r = C''(1-n^{-\beta(2-\alpha)})
\end{aligned}
$$

for some constants $C, C', C'' > 0$. To control the integral over radii in $[1,\infty)$, the cruder bound $|f(y)-f(z)| \leq 2\|f\|$ suffices and, using the fact that

$$\mathrm{Vol}\{x : S_f\cap B(x,r)\neq\emptyset\} \leq C_2(r^d\vee 1), \qquad (57)$$

we have

$$\begin{aligned}
&\left|uF'(\langle w,f\rangle)\int_{\mathbb{R}^d}\int_1^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\,w(y)(f(z)-f(y))\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\right| \\
&\qquad\le C\int\int_1^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\big(\mathbf{1}_{\{y\in S_f\}}+\mathbf{1}_{\{z\in S_f\}}\big)\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x \\
&\qquad\le C'\int_1^\infty \frac{1}{r^{d+1+\alpha}}\int \mathbf{1}_{\{B(x,r)\cap S_f\neq\emptyset\}}\mathrm{Vol}(S_f)\mathrm{d}x\mathrm{d}r \\
&\qquad\le C''\int_1^\infty \frac{1}{r^{d+1+\alpha}}\,r^d\mathrm{d}r\le C''',
\end{aligned}$$

again for some constants C, C', C'' and C''' which depend only on d, F and f.

To control the second term in the right hand side of (56), we use (57) together with the inequality

$$|\langle \mathbf{1}_{B(x,r)}w,f\rangle|\le \|f\|\mathrm{Vol}(S_f\cap B(x,r))\le C_1\|f\|(r^d\wedge 1), \tag{58}$$

to see that it is bounded by

$$\begin{aligned}
&n^{-\gamma}\frac{u^2C_F}{2}\times 2C_1^2\|f\|^2\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}(r^d\wedge 1)^2\mathbf{1}_{\{S_f\cap B(x,r)\neq\emptyset\}}\mathrm{d}r\mathrm{d}x \\
&\quad= C_3n^{-\gamma}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}(r^d\wedge 1)^2(r^d\vee 1)\mathrm{d}r \\
&\quad= C_3n^{-\gamma}\int_{n^{-\beta}}^1 \frac{r^{2d}}{r^{d+1+\alpha}}\mathrm{d}r+C_3n^{-\gamma}\int_1^\infty \frac{r^d}{r^{d+1+\alpha}}\mathrm{d}r=C_4n^{-\gamma}\big(1-n^{-\beta(d-\alpha)}\big).
\end{aligned}\tag{59}$$

When $d\ge 2$, $d-\alpha>0$ and so this bound tends to 0 as $n\to\infty$. When $d=1$, $(\alpha-1)\beta-\gamma=0$, and so this term is bounded by a constant as $n\to\infty$. The same calculation shows that $\varepsilon_n\to 0$, uniformly in w, as $n\to\infty$. As a consequence, in any dimension the absolute value of the neutral term of $\mathcal{L}^n\Psi_{F,f}(M)$ is bounded by a constant independent of n and M.

Proceeding in the same way as for the second term above, we obtain that the selection term of the generator is bounded by (recall that $1-\alpha\beta-\gamma=0$)

$$\begin{aligned}
&2u\sigma n^{1-\beta\alpha-\gamma-\delta}C_F\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^3}\frac{1}{V_r^2}\,|f(z')|\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}z'\mathrm{d}r\mathrm{d}x \\
&\qquad\le Cn^{-\delta}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d}\big(1\wedge r^d\big)\,\mathbf{1}_{\{B(x,r)\cap S_f\neq\emptyset\}}\mathrm{d}x\mathrm{d}r \\
&\qquad\le C'n^{-\delta}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\,\big(1\wedge r^d\big)\big(1\vee r^d\big)\mathrm{d}r\le C''n^{-\delta+\alpha\beta}=C'',
\end{aligned}$$

since $\alpha\beta-\delta=0$. Together with our bounds on the neutral term, just as in the corresponding part of the proof of Theorem 1.8, this shows the tightness of the sequence of finite variation parts of $(\Psi_{F,f}(M_t^n))_{t\ge 0}$.

For the quadratic variation of the martingale part, a similar analysis yields that the integrand in $\mathcal{Q}_t^n$ is bounded by

$$\begin{aligned}
&Cn^{1-\beta\alpha-2\gamma}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^3}\frac{1}{V_r}\,f(z)f(z')\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}z'\mathrm{d}r\mathrm{d}x \\
&\qquad\le C'n^{-\gamma}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\,\big(1\wedge r^d\big)^2\mathbf{1}_{\{B(x,r)\cap S_f\neq\emptyset\}}\mathrm{d}r\mathrm{d}x \\
&\qquad\le C''n^{-\gamma}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\,\big(1\wedge r^d\big)^2\big(1\vee r^d\big)\mathrm{d}r\le C'''n^{-\gamma}\big(1+n^{-\beta(d-\alpha)}\big),
\end{aligned}\tag{60}$$

which is bounded by a constant independent of n and M. As before, we conclude that the sequence of quadratic variations of $(\Psi_{F,f}(M^n_t))_{t\geq 0}$ is tight.

By the same arguments as in the proof of Theorem 1.8, taking f of the form (54), we conclude that the sequence $(\overline{M}^n)_{n\geq 1}$ is tight in $D_{\mathcal{M}_\lambda}[0,\infty)$.

2) Identifying the limit.

Let us first consider the generator of the Markov process $(M^n_t)_{t\geq 0}$ applied to functions of the form $\Psi_{\mathrm{Id},\varphi_f}$ (with φ_f defined in (54)). First, let us find the limit of its neutral component. Since $1-\alpha\beta-\gamma=0$, the prelimit takes the form

$$un^{1-\alpha\beta-\gamma}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r}\,w(y)(\varphi_f(z)-\varphi_f(y))\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x$$
$$=u\int_{\mathbb{R}^d}w(y)\int_{\mathbb{R}^d}\left(\int_{n^{-\beta}\vee\frac{|z-y|}{2}}^{\infty}\frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\;\mathrm{d}r\right)(\varphi_f(z)-\varphi_f(y))\,\mathrm{d}z\mathrm{d}y.$$

Now, a simple Taylor expansion gives us that

$$\varphi_f(z)-\varphi_f(y)=f(z)-f(y)+\mathcal{O}(n^{-2\beta})\big(\mathbf{1}_{\{B_n(z)\cap S_f\neq\emptyset\}}+\mathbf{1}_{\{B_n(y)\cap S_f\neq\emptyset\}}\big),$$

where $B_n(\cdot)=B(\cdot,n^{-\beta})$ and the error term is uniform in y and z. Since

$$n^{-2\beta}\int_{\mathbb{R}^d}w(y)\int_{\mathbb{R}^d}\left(\int_{n^{-\beta}\vee\frac{|z-y|}{2}}^{\infty}\frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\,\mathrm{d}r\right)\big(\mathbf{1}_{\{B_n(z)\cap S_f\neq\emptyset\}}+\mathbf{1}_{\{B_n(y)\cap S_f\neq\emptyset\}}\big)\mathrm{d}z\mathrm{d}y$$
$$\leq Cn^{-2\beta}\int_{S_f+B_n(0)}\int_{\mathbb{R}^d}\left(n^{-\beta}\vee\frac{|z-y|}{2}\right)^{-d-\alpha}\mathrm{d}z\mathrm{d}y\leq C'n^{-\beta(2-\alpha)}\to 0$$

as $n\to\infty$, we can conclude that up to a vanishing error term, the neutral part of $\mathcal{L}^n\Psi_{F,\varphi_f}(M)$ is given by

$$u\int_{\mathbb{R}^d}w(y)\int_{\mathbb{R}^d}\left(\int_{n^{-\beta}\vee\frac{|z-y|}{2}}^{\infty}\frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\;\mathrm{d}r\right)(f(z)-f(y))\,\mathrm{d}z\mathrm{d}y.\tag{61}$$

Now, our computations in the proof of tightness imply that the function

$$a_n(y):y\mapsto\int_{\mathbb{R}^d}\left(\int_{n^{-\beta}\vee\frac{|z-y|}{2}}^{\infty}\frac{1}{r^{d+1+\alpha}}\frac{V_r(y,r)}{V_r}\;\mathrm{d}r\right)(f(z)-f(y))\,\mathrm{d}z$$

is a continuous function, uniformly bounded in y and n. Hence, up to a vanishing error term we can first replace w by $\overline{w}$ (the local average of w over a ball of radius $n^{-\beta}$) in (61) and, second, use dominated convergence to pass to the limit as $n\to\infty$ in the expression for a_n. Doing so, we obtain that the limit of the neutral term in $\mathcal{L}^n\Psi_{F,\varphi_f}(M)$ is equal to

$$u\int_{\mathbb{R}^d}\overline{w}(y)\int_{\mathbb{R}^d}\Phi(|z-y|)(f(z)-f(y))\mathrm{d}z\mathrm{d}y,$$

where, as in (19),

$$\Phi(|z-y|):=\int_{\frac{|z-y|}{2}}^{\infty}\frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\,\mathrm{d}r.$$

Lemma 5.1. *Writing*

$$\mathcal{D}^\alpha f(y) = u \int_{\mathbb{R}^d} \Phi(|z-y|)(f(z)-f(y))\mathrm{d}z, \tag{62}$$

$\mathcal{D}^\alpha$ is the generator of a symmetric α-stable process $(\zeta_t)_{t\geq 0}$.

Proof of Lemma 5.1.

It is reassuring to first check that this is the generator of a well-defined Lévy process:

$$\int_{\mathbb{R}^d} (1\wedge |y|^2) \int_0^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y)}{V_r}\,\mathrm{d}r\,\mathrm{d}y \leq C\int_0^1 \frac{1}{r^{d+1+\alpha}} \int_{B(0,2r)} |y|^2\,\mathrm{d}y\mathrm{d}r + C' \int_1^\infty \frac{1}{r^{d+1+\alpha}}\mathrm{d}r < \infty.$$

To verify that the associated Lévy process is a symmetric stable process, we check the scaling property. The generator of $(b^{-1/\alpha}\zeta_{bt})$ is given by

$$\begin{aligned}\mathcal{D}^\alpha_b f(y) &= bu \int_{\mathbb{R}^d} \Phi(|z - b^{1/\alpha}y|)(f(b^{-1/\alpha}z) - f(y))\,\mathrm{d}z \\ &= ub^{1+d/\alpha} \int_{\mathbb{R}^d} \Phi(|b^{1/\alpha}z - b^{1/\alpha}y|)(f(z)-f(y))\,\mathrm{d}z.\end{aligned}$$

But a simple change of variables gives us that

$$\begin{aligned}\Phi(|b^{1/\alpha}z - b^{1/\alpha}y|) &= \int_{\frac{b^{1/\alpha}|z-y|}{2}}^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(b^{1/\alpha}y, b^{1/\alpha}z)}{V_r}\,\mathrm{d}r \\ &= b^{-1-d/\alpha} \int_{\frac{|z-y|}{2}}^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r}\,\mathrm{d}r,\end{aligned}$$

and so $\mathcal{D}^\alpha_b = \mathcal{D}^\alpha$ for all $b>0$. This shows the desired property of $\mathcal{D}^\alpha$. □

Having identified the neutral part of the limit, we now turn to the selection part. It is given by

$$u\sigma n^{1-\beta\alpha-\gamma-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)^3} \frac{1}{V_r^2}\,(w(y)w(z) - w(z'))\varphi_f(z')\,\mathrm{d}y\mathrm{d}z\mathrm{d}z'\mathrm{d}r\mathrm{d}x. \tag{63}$$

Now, the term which is linear in w is easy to deal with: by Fubini's Theorem, it is equal to

$$\begin{aligned}&u\sigma n^{-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}} \int_{B(x,r)} w(z')\varphi_f(z')\,\mathrm{d}z'\mathrm{d}r\mathrm{d}x \\ &\quad = u\sigma n^{-\delta} \int_{\mathbb{R}^d} w(z')\varphi_f(z') \left(\int_{n^{-\beta}}^\infty \frac{V_1 r^d}{r^{d+1+\alpha}}\,\mathrm{d}r \right) \mathrm{d}z' = \frac{u\sigma V_1}{\alpha}\,\langle \overline{w}, f\rangle,\end{aligned} \tag{64}$$

where the last equality uses the fact that $\alpha\beta - \delta = 0$.

Similar calculations show that the 'quadratic' term in (63) is equal to

$$u\sigma n^{-\delta} \int_{\mathbb{R}^d} \int_{n^{-\beta}}^{n^{-\beta}\log n} \frac{1}{r^{d+1+\alpha}} \left(\int_{B(x,r)} \frac{1}{V_r}\,w(y)\mathrm{d}y \right)^2 \int_{B(x,r)} \varphi_f(z)\mathrm{d}z\mathrm{d}r\mathrm{d}x + \mathcal{O}\big((\log n)^{-\alpha}\big).$$

In contrast with the fixed radius case, here we first have to show that up to a vanishing error term, along the trajectories of the process M^n we can replace the average of the density w^n over a ball of radius at most $n^{-\beta}\log n$ by $\overline{w}^n$, the average over a ball of radius $n^{-\beta}$ centered at the same point. In a second step, we use the same method as in the fixed radius case to prove that if we consider a subsequence of $\overline{M}^n$ converging to some limit with density w^∞, then for every $t\geq 0$, $(\overline{w}^n_t)^2$ converges to $(w^\infty_t)^2$ in the appropriate sense.

Concerning the first point, we have

$$\left(\int_{B(x,r)}\frac{1}{V_r}\,w(y)\mathrm{d}y\right)^2=\left(\int_{B(x,r)}\frac{1}{V_r}\,w(y)\mathrm{d}y+\overline{w}(x)\right)\left(\int_{B(x,r)}\frac{1}{V_r}\,w(y)\mathrm{d}y-\overline{w}(x)\right)+\overline{w}(x)^2.$$

Suppose we have the following lemma (whose proof is quite technical and is given in Appendix C).

Lemma 5.2. *Under the conditions of Theorem 1.9, for every $x\in\mathbb{R}^d$ and $r\in[n^{-\beta},n^{-\beta}\log n]$,*

$$\lim_{n\to\infty}\mathbb{E}\left[\left|\int_{B(x,r)}\frac{w^n_t(y)}{V_r}\,\mathrm{d}y-\overline{w}^n_t(x)\right|\right]=0$$

uniformly in $t\leq T$.

From this result, we can conclude from a simple dominated convergence argument, and a Taylor expansion of φ_f, that along any trajectory of M^n, the 'quadratic' part of (63) is equal to

$$u\sigma n^{-\delta}\int_{\mathbb{R}^d}\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{V_r}{r^{d+1+\alpha}}\,\overline{w}^n_\cdot(x)^2f(x)\,\mathrm{d}r\mathrm{d}x+o(1)=\frac{u\sigma V_1}{\alpha}\,\langle(\overline{w}^n_\cdot)^2,f\rangle+o(1).$$

As concerns the second point, we proceed as in (35) and below. Using Proposition C.1(ii) in Appendix C, the facts that the support of f is bounded, that p_ε is supported in $B(0,\varepsilon)$ (so that τ_2 in (96) is bounded by $\varepsilon^{\alpha/(d+1)}$ when $|z_1-z_2|\leq\varepsilon$ and n is sufficiently large), we obtain that the first term in the decomposition (35) of $\langle(\overline{w}^n_s)^2,f\rangle$ is bounded by a constant (independent of n,ε) times

$$n^{-a}+\varepsilon^{\alpha/(d+1)}+\varepsilon^{1/4}+\varepsilon^{\alpha/(2d+2)}+n^{-\beta(d-1)}\varepsilon^{(\alpha-d)/(2d+2)}.$$

Letting n tend to infinity in the above expression, we can write that the third term in the decomposition (35) is bounded by a constant times

$$\varepsilon^{\alpha/(d+1)}+\varepsilon^{1/4}+\varepsilon^{\alpha/(2d+2)}+\varepsilon^{(\alpha-1)/4}\mathbf{1}_{\{d=1\}}.$$

Finally, the second term in the decomposition (35) tends to 0 by the assumption that $\overline{M}^n$ converges to M^∞. As in the fixed radius case, we can therefore conclude that $\langle(\overline{w}^n_s)^2,f\rangle$ converges to $\langle(\overline{w}^\infty_s)^2,f\rangle$ as n tends to infinity, uniformly in s and whatever the representatives of the different densities that we choose.

Combining the above, we obtain that any limit point $(M^\infty_t)_{t\geq 0}$ of $(\overline{M}^n_t)_{t\geq 0}$ should satisfy: for every $f\in C^\infty_c$, every choice of representative w^∞_s of the density of M^∞_s at every time s, and for every $t\geq 0$,

$$\langle w^\infty_t,f\rangle-\langle w^\infty_0,f\rangle=\int_0^t\left\{\langle w^\infty_s,\mathcal{D}^\alpha f\rangle-\frac{u\sigma V_1}{\alpha}\,\langle w^\infty_s(1-w^\infty_s),f\rangle\right\}\mathrm{d}s+M_t(f),\tag{65}$$

where $(M_t(f))_{t\geq 0}$ is a continuous zero-mean martingale. As in the fixed radius case, when $d\geq 2$ our computations in the paragraph on tightness show that $M(f)\equiv 0$. By another modification of the results of Chapter 7 in [Lia09], we can also conclude that there exists a unique $\mathcal{M}_\lambda$-valued (deterministic) solution to the system of equations (65) and that $\overline{M}^n$ converges weakly to it as n tends to infinity.

When $d=1$, the integrand in the expression for the quadratic variation of the martingale part of $\Psi_{\mathrm{Id},\varphi_f}(M^n)=\langle\overline{w}^n,f\rangle$ is given by

$$u^2n^{-\gamma}\int_{\mathbb{R}}\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{1}{r^{2+\alpha}V_r}\int_{B(x,r)^3}\left\{w_s^n(y)(1-w_s^n(z))(1-w_s^n(z'))+(1-w_s^n(y))w_s^n(z)w_s^n(z')\right\}$$
$$\times\varphi_f(z)\varphi_f(z')\mathrm{d}y\mathrm{d}z\mathrm{d}z'\mathrm{d}r\mathrm{d}x+\mathcal{O}\big((\log n)^{-\alpha}\big),$$

where we have used the estimates obtained in the paragraph on tightness, which show that selective events and neutral events with radii greater than $n^{-\beta}\log n$ do not contribute.

Writing $\overline{w}_s^{n,r}(x)$ for the average value of w_s^n over $B(x,r)$ (so that $\overline{w}_s^{n,n^{-\beta}}(x)=\overline{w}_s^n(x)$), essentially the same techniques show that the integral above is equal to

$$\begin{aligned}&u^2n^{-\gamma}\int_{\mathbb{R}}f(x)^2\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{V_r^2}{r^{2+\alpha}}\left\{\overline{w}_s^{n,r}(x)(1-\overline{w}_s^{n,r}(x))^2+(1-\overline{w}_s^{n,r}(x))\overline{w}_s^{n,r}(x)^2\right\}\mathrm{d}r\mathrm{d}x+o(1)\\&=u^2n^{-\gamma}\int_{\mathbb{R}}f(x)^2\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{V_r^2}{r^{2+\alpha}}\,\overline{w}_s^n(x)(1-\overline{w}_s^n(x))\,\mathrm{d}r\mathrm{d}x\\&\quad+u^2n^{-\gamma}\int_{\mathbb{R}}f(x)^2\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{V_r^2}{r^{2+\alpha}}\big(\overline{w}_s^{n,r}(x)-\overline{w}_s^n(x)\big)\big(1-\overline{w}_s^{n,r}(x)\big)\,\mathrm{d}r\mathrm{d}x\\&\quad+u^2n^{-\gamma}\int_{\mathbb{R}}f(x)^2\int_{n^{-\beta}}^{n^{-\beta}\log n}\frac{V_r^2}{r^{2+\alpha}}\,\overline{w}_s^n(x)\big(\overline{w}_s^n(x)-\overline{w}_s^{n,r}(x)\big)\,\mathrm{d}r\mathrm{d}x+o(1).\end{aligned}\tag{66}$$

But the first term in the right hand side of (66) is equal to

$$\frac{4u^2}{\alpha-1}\,\langle\overline{w}_s^n(1-\overline{w}_s^n),f^2\rangle+\mathcal{O}\big((\log n)^{-\alpha}\big),$$

while the other two terms tend to 0 by Lemma 5.2 and the Dominated Convergence Theorem. Using again the continuity estimates stated in Proposition C.1(ii) in Appendix C and Theorem 6.3.4 in [EK86], we find that any limit of $(\overline{M}^n)_{n\geq1}$ satisfies the martingale problem stated in Theorem 1.9. In addition, considering test functions of the form $\Psi_{(\cdot)^2,\varphi}$ and carrying out the same analysis as above, we can show that any limit of $(\overline{M}^n)_{n\geq1}$ must satisfy that for every $f\in C_c^\infty(\mathbb{R}^d)$,

$$\Bigg(\langle w_t^\infty,f\rangle^2-\langle w_0^\infty,f\rangle^2-2\int_0^t\langle w_s^\infty,f\rangle\bigg\{\langle w_s^\infty,\mathcal{D}^\alpha f\rangle-\frac{2u\sigma}{\alpha}\langle w_s^\infty(1-w_s^\infty),f\rangle\bigg\}\mathrm{d}s$$
$$-\frac{4u^2}{\alpha-1}\int_0^t\langle w_s^\infty(1-w_s^\infty),f^2\rangle\mathrm{d}s\Bigg)_{t\geq0}$$

is a continuous zero-mean martingale. From this, we can deduce that the quadratic variation process of $(\langle w_t^\infty,f\rangle)_{t\geq0}$ with $(\langle w_t^\infty,g\rangle)_{t\geq0}$ is given by

$$\big[\mathcal{Z}^f,\mathcal{Z}^g\big]_t=\frac{4u^2}{\alpha-1}\int_0^t\langle w_s^\infty(1-w_s^\infty),fg\rangle\mathrm{d}s,$$

as stated in Theorem 1.9(i).

Uniqueness follows from duality with the limiting process of Theorem 1.11, proved by the appropriate modification of the construction in Chapter 7 of [Lia09]. This completes the proof of Theorem 1.9. □

Finally, let us prove the convergence of the rescaled dual process.

Proof of Theorem 1.11.

Most of the proof is identical to that of Theorem 1.10. That the only possible limit for $(\Xi_t^n)_{t\geq 0}$ is the system of branching (and in one dimension coalescing) symmetric α-stable processes described in the theorem, again follows from an adaptation of Chapter 7 of [Lia09], in which the only change is that Brownian motion is replaced by the stable process generated by $\mathcal{D}^\alpha$ (see (62)) and we have added natural selection/branching of particles.

Next, we have to show that the sequence $(\Xi^n)_{n\geq 1}$ is tight. Recall the notation $\mathbb{P}_\psi$ for the probability measure on $D_{\mathcal{M}_p(\mathbb{R}^d)}[0,\infty)$ under which for each $n\geq 1$, the locations of the atoms of Ξ_0^n have density ψ. As in the proof of Theorem 1.10, after proving that $(\Xi_t^n)_{n\geq 1}$ is tight for every $t\geq 0$, we shall use the Aldous-Rebolledo criterion based on stopping times. That is, we fix $T>0$, $f\in C^\infty(\mathbb{R}^d)$ with values in $[0,1]$ and show that if $(\tau_n)_{n\geq 1}$ is a sequence of stopping times bounded by $T-\delta_0$ for some small δ_0, then for every $\varepsilon>0$ there exists $\delta=\delta(f,T,\psi,\varepsilon)>0$ such that

$$\limsup_{n\to\infty} \mathbb{P}_\psi\left[\sup_{0\leq t\leq \delta}\left|\prod_{i=1}^{N^n_{\tau_n+t}} f\big(\xi^{n,i}_{\tau_n+t}\big)-\prod_{i=1}^{N^n_{\tau_n}} f\big(\xi^{n,i}_{\tau_n}\big)\right|>\varepsilon\right]\leq \varepsilon. \tag{67}$$

Again, we proceed in four steps. First, by exactly the same arguments as in the proof of Theorem 1.10, there exists $K>0$ such that for every $n\in\mathbb{N}$ we have

$$\mathbb{P}_\psi(A_n):=\mathbb{P}_\psi\left[\sup_{0\leq s\leq T}|\Xi^n|\leq K\right]\geq 1-\frac{\varepsilon}{5}. \tag{68}$$

Furthermore, there exists $\delta_1>0$, independent of the subinterval of $[0,T]$ considered, such that

$$\mathbb{P}_\psi\big[\text{at least 1 particle created in } [\tau_n,\tau_n+\delta_1]\,;\,A_n\big]\leq\frac{\varepsilon}{5}. \tag{69}$$

As before, the difficulty will be to control the coalescence, but suppose for a moment that there is no change in the number of particles in the interval $[\tau_n,\tau_n+\delta_2]$ and write I_n for the indexing set of the particles in $\Xi^n_{\tau_n}$. Then, exactly as before, assuming that there exists a compact $E\subset\mathbb{R}^d$ such that

$$\mathbb{P}_\psi\left[\sup_{t\in[0,T]}\Xi_t^n(E^c)>0;\,A_n\right]\leq\frac{\varepsilon}{5} \tag{70}$$

we can then write

$$\left|\prod_{i\in I_n} f\big(\xi^{n,i}_{\tau_n+t}\big)-\prod_{i\in I_n} f\big(\xi^{n,i}_{\tau_n}\big)\right|\leq C\bigg(\sup_E\|\nabla f\|\bigg)\sum_{i\in I_n}\big|\xi^{n,i}_{\tau_n+t}-\xi^{n,i}_{\tau_n}\big|, \tag{71}$$

and it suffices to consider the motion of a single lineage to control the evolution of the whole set of potential lineages. This is slightly more involved than in the fixed radius case.

Let $(Z_t^n)_{t\geq 0}$ be a Lévy process, independent of $(\xi_t^n)_{t\geq 0}$ and with generator

$$D^nf(x):=u(1+s_n)\int_0^{n^{-\beta}}\frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d}\frac{V_r(x,y)}{V_r}\big(f(y)-f(x)\big)\,\mathrm{d}y\,\mathrm{d}r.$$

Then the process $(X_t)_{t\geq 0}$ defined by $X_t = \xi_t^n + Z_t^n$ has generator $(1+s_n)\mathcal{D}^\alpha$, where $\mathcal{D}^\alpha$ was shown in Lemma 5.1 to be the generator of a symmetric stable process (indeed, observe that the jump rates of ξ^n and Z^n depend only on the jump size $|y-x|$, hence the fact that the intensity measure of the jumps of X is the sum of the intensity measures of ξ^n and Z^n). Using the strong Markov property and standard results on the growth of Lévy processes, see e.g. [Pru81], we have for any $\eta, \delta > 0$, and any stopping time T_n

$$\mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|X_{T_n+t} - X_{T_n}\big| > \eta\Big] < C\frac{\delta}{\eta^\alpha}$$

for a constant C which is independent of η, δ and T_n.

Since

$$\mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|\xi^n_{T_n+t} - \xi^n_{T_n}\big| > \eta\Big] \leq \mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|X_{T_n+t} - X_{T_n}\big| > \eta\Big] + \mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|Z^n_{T_n+t} - Z^n_{T_n}\big| > \eta\Big], \quad (72)$$

it remains to show that

$$\mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|Z^n_{T_n+t} - Z^n_{T_n}\big| > \eta\Big] \to 0, \qquad \text{as } n\to\infty.$$

Now, by construction, $(Z_t^n)_{t\geq 0}$ is a Lévy process whose generator D^n satisfies, for $f \in C^\infty(\mathbb{R}^d)$,

$$\begin{aligned} |D^n f(x)| &= \left| u\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r}\big[(y-x).\nabla f(x) + \mathcal{O}(|y-x|^2)\big]\,\mathrm{d}y\mathrm{d}r\right| \qquad (73)\\ &\leq C\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{B(x,2r)} |y-x|^2\,\mathrm{d}y\mathrm{d}r = C' n^{-\beta(2-\alpha)}, \end{aligned}$$

where the Taylor expansion is justified since $V_r(x,y)=0$ if $|x-y|>2r$ and we are concentrating on radii $r \leq n^{-\beta}$, and the first integral in the right hand side of (73) vanishes by rotational symmetry.

The process $(Z_t^n)_{t\geq 0}$ has finite quadratic variation, whose time derivative when $Z_t^n = x$ is

$$\begin{aligned} &(1+s_n)u\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r}\big(f(y)-f(x)\big)^2\,\mathrm{d}y\mathrm{d}r\\ &\quad = (1+s_n)u\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{V_r(x,y)}{V_r}\big((y-x).\nabla f(x) + \mathcal{O}(|y-x|^2)\big)^2\,\mathrm{d}y\mathrm{d}r\\ &\quad \leq C\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{B(x,2r)} |y-x|^2\,\mathrm{d}y\mathrm{d}r = C' n^{-\beta(2-\alpha)}. \end{aligned}$$

Hence, we can conclude that for any η, δ,

$$\lim_{n\to\infty} \mathbb{P}\Big[\sup_{t\in[0,\delta]} \big|Z^n_{T_n+t} - Z^n_{T_n}\big| > \eta\Big] = 0.$$

Coming back to (72), and taking $T_n = 0$, $\delta = T$ and η large enough, we first obtain (70). Together with (68), this shows the tightness of $(\Xi_t^n)_{n\geq 1}$ for every $t \in [0,T]$. Next, taking $T_n = \tau_n$ and η fixed, we can also conclude that there exists $\delta_3 \in (0, \delta_2]$ such that for n large enough,

$$\mathbb{P}\Big[\sup_{t\in[0,\delta_3]} \big|\xi^n_{\tau_n+t} - \xi^n_{\tau_n}\big| > \eta\Big] \leq \frac{\varepsilon}{5K}.$$

Choosing $\eta = \varepsilon/(KC\sup_E \|\nabla f\|)$ and recalling (71), we obtain that for all sufficiently large n,

$$\mathbb{P}_\psi\bigg[\sup_{t\in[0,\delta_3]}\bigg|\prod_{i\in I_n} f\big(\xi^{n,i}_{\tau_n+t}\big) - \prod_{i\in I_n} f\big(\xi^{n,i}_{\tau_n}\big)\bigg| > \varepsilon\,;\, A_n, B^c_{\delta_3}, C^c_{\delta_3}, E^c_n\bigg] \leq \frac{\varepsilon}{5}, \tag{74}$$

where as in the fixed radius case, B^c_δ is the event that there is no branching event in $[\tau^n, \tau^n+\delta]$, C^c_δ is the event that there is no coalescence in $[\tau^n, \tau^n+\delta]$ and $E^c_n = \{\sup_{t\in[0,T]} \Xi^n_t(E^c) = 0\}$.

Finally, tightness will be proven if we can show that coalescence events cannot accumulate. In particular, since we have controlled the total number of particles and the probability of branching, we just need to control the probability that two lineages coalesce. The result will be based on the following lemma.

Lemma 5.3. *Let $(\hat{\xi}^1_{n^\gamma t})_{t\geq 0}$ and $(\hat{\xi}^2_{n^\gamma t})_{t\geq 0}$ be two independent copies of the motion of a single (unscaled) lineage on the timescale $(n^\gamma t,\, t\geq 0)$, and let $\zeta^n_t = \hat{\xi}^2_{n^\gamma t} - \hat{\xi}^1_{n^\gamma t}$ denote their difference. Then, for every $t \geq 0$ we have:*

(i) When $d=1$, there exists $C(t)>0$ such that

$$\limsup_{n\to\infty} \mathbb{E}\bigg[\frac{1}{n^\gamma}\int_0^{n^{1-\gamma}t} \frac{1}{2^\alpha \vee |\zeta^n_s|^\alpha}\,\mathrm{d}s\bigg] \leq C(t).$$

Furthermore, the function $t\mapsto C(t)$ can be chosen such that $C(t)\downarrow 0$ as $t\to 0$.

(ii) When $d\geq 2$,

$$\lim_{n\to\infty} \mathbb{E}\bigg[\frac{1}{n^\gamma}\int_0^{n^{1-\gamma}t} \frac{1}{2^\alpha \vee |\zeta^n_s|^\alpha}\,\mathrm{d}s\bigg] = 0.$$

We defer the proof of Lemma 5.3 until after the end of the proof of Theorem 1.11.

Suppose that we start with a sample of two (non independent) lineages at some (unscaled) separation $z_0\in\mathbb{R}^d$. As before, we work on the timescale n^γ so that a single lineage jumps at rate $\mathcal{O}(1)$ and we suppose the two lineages ξ^1 and ξ^2 are currently at locations 0 and z (in fact, only their separation matters). Then, the generator Γ of the difference walk $(\xi^2_{n^\gamma t} - \xi^1_{n^\gamma t})_{t\geq 0}$ is equal to

$$\begin{aligned}
\Gamma f(z) = 2u(1+s_n)\int_{\mathbb{R}^d}\bigg\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d} \mathbf{1}_{\{0\notin B(x,r)\}}\mathbf{1}_{\{z\in B(x,r)\}}\frac{\mathbf{1}_{\{y\in B(x,r)\}}}{V_r}\,\mathrm{d}x\mathrm{d}r \\
+ (1-u_n)\int_1^\infty \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d} \mathbf{1}_{\{0\in B(x,r)\}}\mathbf{1}_{\{z\in B(x,r)\}}\frac{\mathbf{1}_{\{y\in B(x,r)\}}}{V_r}\,\mathrm{d}x\mathrm{d}r\bigg\}(f(y)-f(z))\mathrm{d}y \\
= 2u(1+s_n)\int_{\mathbb{R}^d}\bigg\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(y,z)}{V_r}\,\mathrm{d}r\bigg\}(f(y)-f(z))\,\mathrm{d}y \\
- 2u^2 n^{-\gamma}(1+s_n)\int_{\mathbb{R}^d}\bigg\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y,z)}{V_r}\,\mathrm{d}r\bigg\}(f(y)-f(z))\,\mathrm{d}y \qquad (75)\\
+ u^2 n^{-\gamma}(1+s_n)\int_{\mathbb{R}^d}\bigg\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y,z)}{V_r}\,\mathrm{d}r\bigg\}(f(\Delta)-f(z))\,\mathrm{d}y, \qquad (76)
\end{aligned}$$

where Δ is a cemetery state, corresponding to the two walks having coalesced, and $V_r(0,y,z)$ denotes the volume of the intersection $B(0,r)\cap B(y,r)\cap B(z,r)$.

As a consequence, until coalescence we can couple the difference walk (on the timescale n^γ) with the difference $(\zeta^n_t)_{t\geq 0}$ between two independent random walks, each jumping according to the law of a single lineage but with each jump $z \mapsto y$ 'cancelled' with probability

$$\Delta_n(z,y) = \frac{2u^2 n^{-\gamma}(1+s_n)\int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(0,y,z)}{V_r}\,\mathrm{d}r}{2u(1+s_n)\int_1^\infty \frac{1}{r^{d+1+\alpha}} \frac{V_r(y,z)}{V_r}\,\mathrm{d}r}.$$

(One can check that these two descriptions give rise to the same jump times and embedded chain.) Each time we cancel a jump, with probability one half it was a coalescence in the original system, but the key point is that if there are no cancelled jumps, then there was no coalescence.

It therefore suffices to show that we can find $\delta_2 \in (0, \delta_1]$ such that, for sufficiently large n, the probability that an event is cancelled in the interval $[0, \delta_2 n^{1-\gamma}]$ is smaller than $\varepsilon/(5K(K-1))$.

Now, according to the expression in the right hand side of (75), when the two lineages lie at separation $z \in \mathbb{R}^d$, a cancelled event occurs at instantaneous rate

$$\begin{aligned} &2u^2 n^{-\gamma}(1+s_n)\int_{\mathbb{R}^d}\left\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y,z)}{V_r}\,\mathrm{d}r\right\}\mathrm{d}y \\ &\qquad \leq 2u^2 n^{-\gamma}(1+s_n)\int_{\mathbb{R}^d}\int_{1\vee\frac{|z|}{2}\vee\frac{|y|}{2}} \frac{1}{r^{d+1+\alpha}}\,\mathrm{d}r\mathrm{d}y = C_1 n^{-\gamma}\big(2\vee|z|\big)^{-\alpha}. \end{aligned}$$

Hence, (using the coupling with $(\zeta_t^n)_{t\geq 0}$), the probability of having no event cancelled up to time $n^{1-\gamma}t$ (corresponding to time nt in original units) is equal to

$$\begin{aligned} &\mathbb{E}\bigg[\exp\bigg\{-\int_0^{n^{1-\gamma}t} 2u^2 n^{-\gamma}(1+s_n)\int_{\mathbb{R}^d}\bigg\{\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y,\zeta_s^n)}{V_r}\,\mathrm{d}r\bigg\}\mathrm{d}y\mathrm{d}s\bigg\}\bigg] \\ &\quad \geq \mathbb{E}\bigg[\exp\bigg\{-C_1 n^{-\gamma}\int_0^{n^{1-\gamma}t}\big(2\vee|\zeta_s^n|\big)^{-\alpha}\,\mathrm{d}s\bigg\}\bigg] \geq 1 - C_1\,\mathbb{E}\bigg[n^{-\gamma}\int_0^{n^{1-\gamma}t}\frac{\mathrm{d}s}{\big(2\vee|\zeta_s^n|\big)^{\alpha}}\bigg]. \end{aligned}$$

But Lemma 5.3 shows that we can indeed find $\delta_2 > 0$ such that

$$\limsup_{n\to\infty}\mathbb{E}\bigg[n^{-\gamma}\int_0^{n^{1-\gamma}\delta_2}\frac{\mathrm{d}s}{\big(2\vee|\zeta_s^n|\big)^{\alpha}}\bigg] \leq \frac{\varepsilon}{5C_1K(K-1)}. \tag{77}$$

This completes the proof of tightness and therefore of Theorem 1.11. □

Proof of Lemma 5.3.

As before, we shall exploit the fact that $(\zeta_t^n)_{t\geq 0}$ is 'nearly' a symmetric α-stable process. Indeed, the intensity at which $(\zeta_t^n)_{t\geq 0}$ jumps by some vector y is independent of its current location and equal to

$$2(1+s_n)\bigg(\int_1^\infty \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y)}{V_r}\,\mathrm{d}r\bigg)\,\mathrm{d}y.$$

Writing $(Z_t^n)_{t\geq 0}$ for a jump process, independent of $(\zeta_t^n)_{t\geq 0}$, starting at 0 and with jump intensity

$$2(1+s_n)\bigg(\int_0^1 \frac{1}{r^{d+1+\alpha}}\frac{V_r(0,y)}{V_r}\,\mathrm{d}r\bigg)\,\mathrm{d}y,$$

then the generator of the process $(X_t)_{t\geq 0}$, where $X_t = \zeta_t^n + Z_t^n$, is precisely $2(1+s_n)$ times the operator $\mathcal{D}^\alpha$ defined in (62), which we already checked corresponds to a symmetric α-stable process. Once again, the idea is that the jumps of $(Z_t^n)_{t\geq 0}$ (which are bounded by 2) do not contribute much to the evolution of $(X_t)_{t\geq 0}$. More precisely, let us show that there exists $C>0$ such that for every n large enough and every $s\geq 1$,

$$\mathbb{P}\bigg[\frac{|Z^n_s|}{\sqrt{s}} > (\log n)^2\bigg] \leq Ce^{-(\log n)^2/d}. \tag{78}$$

To this end, observe first that since the law of Z^n_s is invariant under rotation, we can write that

$$\mathbb{P}\bigg[\frac{|Z^n_s|}{\sqrt{s}} > (\log n)^2\bigg] \leq d\,\mathbb{P}\bigg[\frac{|Z^{n(1)}_s|}{\sqrt{s}} > \frac{(\log n)^2}{d}\bigg] = 2d\,\mathbb{P}\bigg[\frac{Z^{n(1)}_s}{\sqrt{s}} > \frac{(\log n)^2}{d}\bigg], \tag{79}$$

where $Z^{n(1)}_s$ denotes the first coordinate of Z^n_s. Now, $(Z^{n(1)}_s)_{s\geq 0}$ is again a symmetric Lévy process with jumps bounded by 2, and so Theorem 25.3 in [Sat99] shows that for every $s,q\geq 0$, $\mathbb{E}[\exp(qZ^{n(1)}_s)]<\infty$. In this case, it is known that the characteristic exponent Ψ^n of $(Z^{n(1)}_s)_{s\geq 0}$, given here by a formula of the form

$$\Psi^n(q) = \int_{[-2,2]} \big(1-e^{iqx}+iqx\mathbf{1}_{\{|x|<1\}}\big)\,\mathfrak{m}^n(\mathrm{d}x),$$

has an analytic extension to the half-plane with negative imaginary part, and we have

$$\mathbb{E}\Big[e^{qZ^{n(1)}_s}\Big] = e^{s\psi^n(q)}, \qquad \text{with} \quad \psi^n(q) = -\Psi^n(-iq).$$

As a consequence, the Markov inequality gives us that

$$\mathbb{P}\bigg[\frac{Z^{n(1)}_s}{\sqrt{s}} > \frac{(\log n)^2}{d}\bigg] \leq e^{-(\log n)^2/d + s\psi^n(1/\sqrt{s})}. \tag{80}$$

Since the measure $\mathfrak{m}^n$ has support in $[-2,2]$, we can write that when q is small

$$\begin{aligned}\psi^n(q) &= -\int_{[-2,2]} \big(1-[1+qx+q^2x^2/2+\mathcal{O}(q^3x^3)]+qx\mathbf{1}_{\{|x|<1\}}\big)\,\mathfrak{m}^n(\mathrm{d}x)\\ &= q\int_{[-2,2]} x\mathbf{1}_{\{|x|\geq 1\}}\mathfrak{m}^n(\mathrm{d}x) + \frac{q^2}{2}\int_{[-2,2]} x^2\mathfrak{m}^n(\mathrm{d}x) + \mathcal{O}(q^3),\end{aligned}$$

where the first term on the right is zero, by symmetry. Furthermore, $s_n\to 0$ and so $\mathfrak{m}^n$ converges to some finite $\mathfrak{m}$. Consequently, there exists a constant $C>0$ such that for every $s\geq 1$, $\psi^n(1/\sqrt{s})\leq C/s$. Together with (79) and (80), this gives us (78).

It will be convenient to suppose that $\zeta_0=0$, but notice that there will be no loss of generality in so-doing, since for n sufficiently large, ζ_0 will be bounded by $(\log n)^2$ and so, for $s>1$, can be absorbed into our bound for Z_s. Similarly, we can, and do, replace $2^\alpha\wedge|\zeta^n_s|^\alpha$ by $1\wedge|\zeta^n_s|^\alpha$ in the denominator of our integrand.

Based on these considerations, let us return to the integral of interest when $d\geq 2$. Fixing $a\in(0,\gamma)$ and splitting the integral with respect to time into $\int_{[0,n^a]}+\int_{[n^a,n^{1-\gamma}t]}$, we obtain

$$\begin{aligned}\mathbb{E}\bigg[\frac{1}{n^\gamma}\int_0^{n^{1-\gamma}t}\frac{1}{1\vee|\zeta^n_s|^\alpha}\,\mathrm{d}s\bigg] &= \mathcal{O}(n^{a-\gamma}) + \frac{1}{n^\gamma}\int_{n^a}^{n^{1-\gamma}t}\mathbb{E}\bigg[\frac{1}{1\vee|X_s-Z^n_s|^\alpha}\bigg]\mathrm{d}s\\ &\leq Cn^{a-\gamma} + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t}\mathbb{P}\bigg[\frac{|Z^n_s|}{\sqrt{s}}>(\log n)^2\bigg]\mathrm{d}s + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t}\mathbb{E}\bigg[\frac{\mathbf{1}_{\{|Z^n_s|\leq(\log n)^2\sqrt{s}\}}}{1\vee|X_s-Z^n_s|^\alpha}\bigg]\mathrm{d}s\\ &\leq Cn^{a-\gamma} + Cn^{1-2\gamma}te^{-(\log n)^2/d} + n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t}\mathbb{E}\bigg[\frac{\mathbf{1}_{\{|Z^n_s|\leq(\log n)^2\sqrt{s}\}}}{1\vee|X_s-Z^n_s|^\alpha}\bigg]\mathrm{d}s.\end{aligned} \tag{81}$$

Since the first two terms on the right tend to 0 as $n \to \infty$, it now suffices to show that the last term remains bounded when n is large.

By Lemma 5.3 in [BW98], if $(p_s^\alpha)_{s\geq 0}$ denotes the transition density of $(X_t)_{s\geq 0}$, we have, for every $s > 0$ and $x \in \mathbb{R}^d$,

$$p_s^\alpha(0,x) =: p_s^\alpha(x) = s^{-d/\alpha} p_1^\alpha(xs^{-1/\alpha}) \tag{82}$$

and there exists $C_{d,\alpha} > 0$ (independent of x) such that

$$0 \leq p_1^\alpha(x) \leq C_{d,\alpha}\big(1 + |x|^{d+\alpha}\big)^{-1}. \tag{83}$$

Hence, for any $s \geq n^a$ and any $z \in \mathbb{R}^d$ such that $|z| \leq (\log n)^2\sqrt{s}$, we can write

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{1 \vee |X_s - z|^\alpha}\right] \leq s^{-d/\alpha}\int_{\mathbb{R}^d} \frac{1}{(1 \vee |x-z|^\alpha)(1 + |xs^{-1/\alpha}|^{d+\alpha})}\,\mathrm{d}x \\
&\leq s^{-d/\alpha}\int_{B(z,1)} \frac{1}{1 + |xs^{-1/\alpha}|^{d+\alpha}}\,\mathrm{d}x + s^{-d/\alpha}\int_{B(z,1)^c} \frac{1}{|x-z|^\alpha(1 + |xs^{-1/\alpha}|^{d+\alpha})}\,\mathrm{d}x \\
&\leq Cs^{-d/\alpha} + C's^{-d/\alpha}\int_{B(0,s^{1/\alpha})\setminus B(z,1)} \frac{\mathrm{d}x}{|x-z|^\alpha} + C''s^{-d/\alpha}\int_{B(0,s^{1/\alpha})^c} \frac{\mathrm{d}x}{|x-z|^\alpha |xs^{-1/\alpha}|^{d+\alpha}}.
\end{aligned}$$

But since $s \geq n^a$ and $|z| \leq (\log n)^2\sqrt{s}$, we have

$$|z|s^{-1/\alpha} \leq (\log n)^2 s^{\frac{1}{2}-\frac{1}{\alpha}} \leq (\log n)^2 n^{-a(2-\alpha)/(2\alpha)} \to 0,$$

and so the second term on the right is bounded (after a change to polar coordinates) by

$$C's^{-d/\alpha}\int_1^{s^{1/\alpha}} \rho^{d-1-\alpha} d\rho = C's^{-1},$$

while the third term is bounded by

$$C''s^{-d/\alpha}s^{1+d/\alpha}\int_{s^{1/\alpha}}^\infty \rho^{d-1-2\alpha-d} d\rho = C''s^{-1}.$$

Since all the constants depend on neither z (in the range considered) nor s, we deduce that the right hand side of (81) is bounded by

$$C'n^{a-\gamma} + Cn^{1-2\gamma}te^{-(\log n)^2/d} + C''n^{-\gamma}\big(n^{-a(d-\alpha)/d} + \log n + \log t\big) \to 0 \qquad \text{as } n \to \infty,$$

which proves (ii).

The only point that differs when $d = 1$ is that $1 - d/\alpha > 0$ and so

$$n^{-\gamma}\int_{n^a}^{n^{1-\gamma}t} s^{-1/\alpha}\mathrm{d}s \leq Cn^{-\gamma} n^{(1-\frac{1}{\alpha})(1-\gamma)} t^{1-\frac{1}{\alpha}}.$$

An easy check confirms that $(1 - \frac{1}{\alpha})(1-\gamma) - \gamma = 0$, and so $C(t)$ exists and is proportional to $t^{1-\frac{1}{\alpha}}$. Since $\alpha > 1$, we also have that $C(t) \downarrow 0$ as $t \to 0$. □

References

[AS72] M. Abramowitz and I.A. Stegun, eds. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover Publications.

[Ald78] D. Aldous (1978). Stopping times and tightness. *Ann. Probab.*, 6:335–340.

[BP15] B. Bah and E. Pardoux (2015). Lambda-lookdown model with selection. *Stoch. Process. Appl.*, 125:1089–1126.

[BD01] E. Brunet and B. Derrida (201). Effect of microscopic noise on front propagation. *J. Stat. Phys.* 103:269–282.

[BDE02] N.H. Barton, F. Depaulis and A.M. Etheridge (2002). Neutral evolution in spatially continuous populations. *Theor. Pop. Biol.*, 61(1):31–48.

[BEK18] N. Biswas, A.M. Etheridge and A. Klimek (2018). The spatial Lambda-Fleming-Viot process with fluctuating selection. *arXiv:1802.08188 [math.PR].*

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

[BEV13] N.H. Barton, A.M. Etheridge and A. Véber (2013). Modelling evolution in a spatial continuum. *JSTAT*, P01002.

[BEV12] N. Berestycki, A.M. Etheridge and A. Véber (2013). Large-scale behaviour of the spatial Λ-Fleming-Viot process. *Ann. Inst. H. Poincaré Probab. Statist.*, 49:374–401.

[BW98] P. Biler and W.A. Woyczynski (1998). Global and exploding solutions for nonlocal quadratic evolution problems. *SIAM J. Appl. Math.*, 59:845–869.

[Bil95] P. Billingsley. *Probability and Measure.* Wiley, New York, 1995.

[Bur73] D.L. Burkholder (1973). Distribution function inequalities for martingales. *Ann. Probab.*, 1:19–42.

[CD05] J.G. Conlon and C.R. Doering (2005). On travelling waves for the stochastic Fisher-Kolmogorov-Petrovsky-Piscunov equation. *J. Stat. Phys.*, 120:421–477.

[CR13] X. Cabré and J.-M. Roquejoffre (2013). The influence of fractional diffusion in Fisher-KPP equations. *Communications in Mathematical Physics*, 320:679–722.

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

[Eth11] A.M., Etheridge (2011). *Some Mathematical Models from Population Genetics: École d'été de Probabilités de Saint-Flour XXXIX-2009.* Springer.

[EFP17] A. M. Etheridge, N. Freeman and S. Penington (2017). Branching Brownian motion, mean curvature flow and the motion of hybrid zones. *Electron. J. Probab.*, 22:103.

[EFPS17] A. M. Etheridge, N. Freeman, S. Penington and D. Straulino (2017). Branching Brownian motion and selection in the spatial Lambda-Fleming-Viot process. *Ann. Applied Probab.*, 27:2605–2645.

[EFS17] A. M. Etheridge, N. Freeman and D. Straulino (2017). Branching Brownian motion, the Brownian net and selection in the spatial Λ-Fleming-Viot process. *Electron. J. Probab.*, 22:39.

[EK18] A.M. Etheridge and T.G. Kurtz (2018). Genealogical constructions of population models. To appear in *Ann. Applied Probab.*

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

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

[Fis37] R. Fisher (1937). The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369.

[FP17] R. Forien and S. Penington (2017). A central limit theorem for the spatial Lambda-Fleming-Viot process with selection. *Electron. J. Probab.*, 22:5.

[Fou13] C.Foucart (2013). The impact of selection in the Λ-Wright-Fisher model. *Electron. Commun. Probab.* 18:1–10. Erratum: 2014. *Electron. Commun. Probab.* 15:1–3.

[Kal02] O. Kallenberg (2002). *Foundations of Modern Probability*, 2nd edition. Springer, New York.

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

[KPP37] A.N. Kolmogorov, I. Petrovsky and N. Piscounov (1937). Étude de l'equation de la diffusion avec croissance de la quantité de mati'ere et son application à un problème biologique. *Moscow Univ. Math. Bull.* 1:1–25.

[KN97] S.M. Krone and C. Neuhauser (1997). Ancestral processes with selection. *Theor. Pop. Biol.*, 51:210–237.

[Lia09] R.H. Liang (2009). *Two continuum-sites stepping stone models in population genetics with delayed coalescence.* PhD Thesis, University of California, Berkeley.

[Mil15] L.R. Miller (2015). *Evolution of highly fecund organisms.* PhD Thesis, University of Oxford.

[MMQ11] C. Mueller, L. Mytnik and J. Quastel (2011). Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. Math.* 184:405–453.

[MS95] C. Mueller and R. Sowers (1995). Random travelling waves for the KPP equation with noise. *J. Functional Analysis*, 128:439–498.

[MT95] C. Mueller and R. Tribe (1995). Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Relat. Fields*, 102: 519–545.

[NK97] C. Neuhauser and S.M. Krone (1997). Genealogies of samples in models with selection. *Genetics*, 145:519–534.

[PS71] S.C. Port and C.J. Stone (1971). Infinitely divisible processes and their potential theory. II. *Ann. Instit. Fourier*, 21:179–265.

[Pru81] W.E. Pruitt (1981). The growth of random walks and Lévy processes. *Ann. Probab.*, 9:948–956.

[Reb80] R. Rebolledo (1980). Sur l'existence de solutions à certains problèmes de semimartingales. *C.R. Acad. Sci. Paris*, 290.

[RR66] C.J. Ridler-Rowe (1966). On first hitting times of some recurrent two-dimensional random walks. *Z. Wahr. verw. Geb.*, 5:187–201.

[Sat99] K. Sato (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press.

[SS80] T. Shiga and A. Shimizu (1980). Infinite dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* 20:395–416.

[SW71] E. Stein and G. Weiss (1971). *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press.

[VW15] A. Véber and A. Wakolbinger (2015). The spatial Λ-Fleming-Viot process: An event-based construction and a lookdown representation. *Ann. Instit. H. Poincaré Probab. Statist.*, 51:570–598.

A Proof of Theorem 1.3

In this section, we provide a rather elementary proof of existence of the SLFVS as the unique solution to a martingale problem. Because we can obtain uniform bounds on the values of $\mathcal{L}\Psi$ for every function Ψ in our set of test functions (see (85) below), Fubini's Theorem will then be sufficient to conclude that the solution to the martingale problem indeed has generator $\mathcal{L}$ as stated in Theorem 1.3. We include the proof here as what it lacks in sophistication, it makes up for in flexibility.

Suppose that instead of $\mathbb{R}^d$, the geographical space $E \subset \mathbb{R}^d$ in which the population evolves is a hypercube of finite sidelength. Suppose also that μ and μ' have finite masses. In this case, the events fall globally at a finite rate and the SLFVS is well-defined. To extend to arbitrary measures that satisfy condition (4), it is convenient to proceed in two steps:

- (i) We show existence when E has finite sidelength but μ and μ' are only σ-finite.
- (ii) Given (i), we extend to $\mathbb{R}^d$ by proving tightness of a sequence of processes obtained by restricting to an increasing family of hypercubes $(E_n)_{n\geq 1}$ which exhaust the space.

Recall the definitions of $\Theta^+_{x,r,u}(w)$ and $\Theta^-_{x,r,u}(w)$ given in (5), and set

$$\mathcal{D}(\mathcal{L}) := \big\{F(\langle \cdot, f\rangle) : \, f \in C_c(E), F \in C^1(\mathbb{R})\big\}.$$

For the estimates carried out in the proof of (ii), it will be convenient that the test functions f should have compact support. The result can then be extended to functions that are only continuous and integrable by a simple density argument.

Proof of (i)**.**

Let E be some hypercube with sidelength L, and let μ, μ' be σ-finite measures on $(0,\infty)$. Let also $(\mu^n)_{n\geq 1}$, $(\mu^{'n})_{n\geq 1}$ be two sequences of finite measures on $(0,\infty)$, and $(\{\nu_r^n,\, r>0\})_{n\geq 1}$, $(\{\nu_r^{'n},\, r>0\})_{n\geq 1}$ be two sequences of probability measures on $[0,1]$ for which, as $n\to\infty$,

$$\mu^n(\mathrm{d}r)\nu_r^n(du)\nearrow \mu(\mathrm{d}r)\nu_r(du) \qquad \text{and} \qquad \mu^{'n}(\mathrm{d}r)\nu_r^{'n}(du)\nearrow \mu'(\mathrm{d}r)\nu_r'(du). \tag{84}$$

Finally, let us write $(M_t^{(n)})_{t\geq 0}$ for the SLFVS with parameters μ^n, $\mu^{'n}$, ν^n and $\nu^{'n}$.

The strategy of the proof is as follows. First we show that the sequence $(M^{(n)})_{n\in\mathbb{N}}$ is tight in the space $D_{\mathcal{M}_\lambda(E\times\{0,1\})}[0,\infty)$ of all càdlàg paths with values in $\mathcal{M}_\lambda(E\times\{0,1\})$ (here we specify the space $E\times\{0,1\}$ on which the measures $M_t^{(n)}$ are defined for the sake of clarity). Second, writing $\mathcal{L}^{(n)}$ for the generator of $(M_t^{(n)})_{t\geq 0}$, we check that $\mathcal{L}^{(n)}\Psi_{F,f}\to\mathcal{L}\Psi_{F,f}$ as $n\to\infty$, uniformly in M for each $f\in C_c(E)$, $F\in C^1(\mathbb{R})$. Now if there is a solution to the martingale problem associated with $(\mathcal{L},\mathcal{D}(\mathcal{L}))$, by Proposition 1.5 it is dual to the system of branching and coalescing lineages of Proposition 1.5 and so it is unique (since by Lemma 1.1 of [VW15] the test functions of Proposition 1.5 form a separating class). This uniqueness, plus the uniform convergence of $\mathcal{L}^{(n)}\Psi_{F,f}$ to $\mathcal{L}\Psi_{F,f}$ is enough to apply Theorem 4.8.10 of [EK86] to deduce that the solution to the $(\mathcal{L},\mathcal{D}(\mathcal{L}))$-martingale problem exists and that $(M^{(n)})_{n\geq 1}$ converges to it as $n\to\infty$.

Let us then check tightness and uniform convergence.

Tightness of $(M^{(n)})_{n\geq 1}$.

First recall that $\mathcal{M}_\lambda(E\times\{0,1\})$, equipped with the topology of vague convergence, is a compact space (c.f. Lemma 1.1 in [VW15]). Therefore, by the Aldous-Rebolledo criterion [Ald78, Reb80], we only have to show that for every $f\in C_c(E)$ and every $F\in C^1(\mathbb{R})$, both the finite variation part and the quadratic variation of the martingale part of the real-valued jump processes $(F(\langle w_\cdot^{(n)},f\rangle))_{n\geq 1}$ are tight (here and below, for every time t we fix a representative $w_t^{(n)}$ of the density of $M_t^{(n)}$, e.g. as suggested in Remark 1.2). Since each $(M_t^{(n)})_{t\geq 0}$ is a finite rate jump process, the finite variation part of $F(\langle w_t^{(n)},f\rangle)$ is given by

$$\Phi^n(t)=\int_0^t \mathcal{L}^{(n)}\Psi_{F,f}(M_s^{(n)})\mathrm{d}s$$

and the quadratic variation is

$$\begin{aligned}
&Q^n(t)\\
&=\int_0^t\int_E\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_{x,r}}\Big\{w_{s-}^{(n)}(y)\big[F(\langle\Theta^+_{x,r,u}(w_{s-}^{(n)}),f\rangle)-F(\langle w_{s-}^{(n)},f\rangle)\big]^2\\
&\qquad\qquad+(1-w_{s-}^{(n)}(y))\big[F(\langle\Theta^-_{x,r,u}(w_{-s}^{(n)}),f\rangle)-F(\langle w_{s-}^{(n)},f\rangle)\big]^2\Big\}\,\mathrm{d}y\,\nu_r^n(\mathrm{d}u)\,\mu^n(\mathrm{d}r)\,\mathrm{d}x\,\mathrm{d}s\\
&\quad+\int_0^t\int_E\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_{x,r}^2}\Big\{w_{s-}^{(n)}(y)w_{s-}^{(n)}(z)\big[F(\langle\Theta^+_{x,r,u}(w_{s-}^{(n)}),f\rangle)-F(\langle w_{s-}^{(n)},f\rangle)\big]^2\\
&\quad+(1-w_{s-}^{(n)}(y)w_{s-}^{(n)}(z))\big[F(\langle\Theta^-_{x,r,u}(w_{s-}^{(n)}),f\rangle)-F(\langle w_{s-}^{(n)},f\rangle)\big]^2\Big\}\,\mathrm{d}y\,\mathrm{d}z\,\nu^{'n}(\mathrm{d}u)\,\mu^{'n}(\mathrm{d}r)\,\mathrm{d}x\,\mathrm{d}s,
\end{aligned}$$

where $B(x,r)$ should be understood as $B(x,r)\cap E$ and $V_{x,r}$ is the volume of this ball in E. Using the expression for $\mathcal{L}^{(n)}$ given in (8), a Taylor expansion of F (of class C^1), and the fact that the increments of $\langle w^{(n)}, f\rangle$ during an event are proportional to u, we obtain that for every $M \in \mathcal{M}_\lambda(E\times\{0,1\})$,

$$\begin{aligned}
&\big|\mathcal{L}^{(n)}\Psi_{F,f}(M)\big| \\
&\qquad \leq \|F'\|\|f\|\bigg(\int_E\int_0^\infty\int_0^1 ur^d\nu_r^n(\mathrm{d}u)\mu^n(\mathrm{d}r)\mathrm{d}x + \int_E\int_0^\infty\int_0^1 ur^d\nu_r'^n(\mathrm{d}u)\mu'^n(\mathrm{d}r)\mathrm{d}x\bigg) \\
&\qquad \leq C\bigg(\int_0^\infty\int_0^1 ur^d\nu_r(\mathrm{d}u)\mu(\mathrm{d}r) + \int_0^\infty\int_0^1 ur^d\nu_r'(\mathrm{d}u)\mu'(\mathrm{d}r)\bigg), \qquad (85)
\end{aligned}$$

where we have used that E has finite volume and, by assumption, $\mu^n(\mathrm{d}r)\nu_r^n(\mathrm{d}u) \nearrow \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ (and the corresponding statement with primes). By Condition (4), the expression in the right hand side is finite (and independent of M), and so we conclude that for every ε, there exists $\delta > 0$ such that for every $T > 0$

$$\limsup_{n\to\infty}\ \mathbb{P}\big[\omega'(\Phi^n,\delta,T) > \varepsilon\big] \leq \varepsilon,$$

where ω' is the Skorokhod modulus of continuity defined by

$$\omega'(f,\delta,T) = \inf_{\{t_i\}}\max_i \sup_{s,t\in[t_{i-1},t_i)} |f(s)-f(t)|$$

and the infimum is taken over all finite partitions of $[0,T]$ such that $t_i - t_{i-1} > \delta$ for every i. The sequence $(\Phi^n)_{n\geq 1}$ is thus tight.

Similarly, the increments of Q^n are bounded by $\|F'\|\|f\|\,u^2r^d$. But $u^2 \leq u$ and so the same reasoning shows that the sequence $(Q^n)_{n\geq1}$ is also tight. Tightness of $(M^{(n)})_{n\geq1}$ in $D_{\mathcal{M}_\lambda(E\times\{0,1\})}[0,\infty)$ now follows from the Aldous-Rebolledo criterion.

Convergence of $\mathcal{L}^{(n)}$ to $\mathcal{L}$.

Using (84) and (85), we see that $\mathcal{L}^{(n)}$ converges (in a uniformly bounded way) to the operator $\mathcal{L}$ given by

$$\begin{aligned}
\mathcal{L}\Psi_{F,f}(M) = &\int_E\int_0^\infty\int_0^1\int_{B(x,r)}\frac{1}{V_{x,r}}\Big[w(y)F(\langle\Theta^+_{x,r,u}(w),f\rangle) + (1-w(y))F(\langle\Theta^-_{x,r,u}(w),f\rangle) \\
&\qquad\qquad - F(\langle w,f\rangle)\Big]\,\mathrm{d}y\,\nu_r(\mathrm{d}u)\,\mu(\mathrm{d}r)\,\mathrm{d}x \\
&+ \int_E\int_0^\infty\int_0^1\int_{B(x,r)^2}\frac{1}{V_{x,r}^2}\Big[w(y)w(z)F\big(\langle\Theta^+_{x,r,u}(w),f\rangle\big) \\
&\qquad + (1-w(y)w(z))F(\langle\Theta^-_{x,r,u}(w),f\rangle) - F(\langle w,f\rangle)\Big]\,\mathrm{d}y\,\mathrm{d}z\,\nu_r'(\mathrm{d}u)\,\mu'(\mathrm{d}r)\,\mathrm{d}x. \qquad (86)
\end{aligned}$$

Applying the estimate (85) with the measures $\mu(\mathrm{d}r)\nu_r(\mathrm{d}u) - \mu^n(\mathrm{d}r)\nu_r^n(\mathrm{d}u)$ and $\mu'(\mathrm{d}r)\nu_r'(\mathrm{d}u) - \mu'^n(\mathrm{d}r)\nu_r'^n(\mathrm{d}u)$, we see that the convergence is uniform, as required to prove item (b') of Theorem 4.8.10 in [EK86], and any limit point of $(M^{(n)})_{n\geq1}$ must satisfy the martingale problem associated to $(\mathcal{L},\mathcal{D}(\mathcal{L}))$. The proof of (i) is thus complete.

Proof of (ii)**.**

The proof of (ii) follows exactly the same pattern, but now the task of bounding the integrals defining Φ^n and Ψ^n becomes more delicate. The resolution is to exploit the fact that f has compact support S_f.

Fix μ, μ', ν and ν' satisfying (4) and let $\{E_n\}_{n\geq 1}$ be a sequence of hypercubes increasing to $\mathbb{R}^d$. We embed each $\mathcal{M}_\lambda(E_n \times \{0,1\})$ into $\mathcal{M}_\lambda(\mathbb{R}^d \times \{0,1\})$ by setting $w(x) \equiv 0$ outside E_n. In this way, we have a sequence of SLFVS processes, which by an abuse of notation we also denote $(M_t^{(n)})_{t\geq 0}$, with generators $\mathcal{L}_{(n)}$ given by (recall that $S_f + B(0,r) = \{x+y : x \in S_f, y \in B(0,r)\}$)

$$\begin{aligned}\mathcal{L}_{(n)}\Psi_{F,f}(M) = &\int_0^\infty \int_{(S_f+B(0,r))\cap E_n} \int_0^1 \int_{B(x,r)} \frac{1}{V_{x,r}} \Big[w(y)F(\langle \Theta^+_{x,r,u}(w), f\rangle) \\ &\qquad + (1-w(y))F(\langle \Theta^-_{x,r,u}(w), f\rangle) - F(\langle w, f\rangle)\Big] \mathrm{d}y\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \\ &+ \int_0^\infty \int_{(S_f+B(0,r))\cap E_n} \int_0^1 \int_{B(x,r)^2} \frac{1}{V_{x,r}} \Big[w(y)w(z)F\big(\langle \Theta^+_{x,r,u}(w), f\rangle\big) \\ &\qquad + (1-w(y)w(z))F\big(\langle \Theta^-_{x,r,u}(w), f\rangle\big) - F(\langle w, f\rangle)\Big] \mathrm{d}y\, \mathrm{d}z\, \nu'_r(\mathrm{d}u)\, \mathrm{d}x\, \mu'(\mathrm{d}r). \end{aligned} \tag{87}$$

The key observation is that

$$|\langle \mathbf{1}_{B(x,r)} w, f\rangle| \leq \|f\| \mathrm{Vol}(S_f \cap B(x,r)) \leq C_1 \|f\| (r^d \wedge 1), \tag{88}$$

and

$$\mathrm{Vol}\{x : S_f \cap B(x,r) \neq \emptyset\} \leq C_2 (r^d \vee 1), \tag{89}$$

where C_1 and C_2 are independent of r and depend only on the support of f. Moreover, the estimate (88) is uniform in w and, in particular, the same bound holds if we replace w by $1-w$.

To see how to apply this, consider the neutral part of (87). We split the integral over $(0,\infty)$ at some radius $R_0 > 1$. We have that

$$\begin{aligned} &\bigg| \int_{R_0}^\infty \int_{(S_f+B(0,r))\cap E_n} \int_0^1 \int_{B(x,r)} \frac{1}{V_{x,r}} \Big[w(y)F(\langle \Theta^+_{x,r,u}(w), f\rangle) \\ &\qquad + (1-w(y))F(\langle \Theta^-_{x,r,u}(w), f\rangle) - F(\langle w, f\rangle)\Big] \mathrm{d}y\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \bigg| \\ &\leq \|F'\| \|f\| \int_{R_0}^\infty \int_{(S_f+B(0,r))\cap E_n} \int_0^1 u \mathrm{Vol}(B(x,r) \cap S_f)\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \\ &\leq C \|F'\| \|f\| \mathrm{Vol}(S_f) \int_{R_0}^\infty \int_0^1 u r^d\, \nu_r(\mathrm{d}u) \mu(\mathrm{d}r). \end{aligned} \tag{90}$$

To control the second part of the integral corresponding to the neutral part, notice that a simple estimate using the fact that the corresponding events have radius bounded above by R_0, yields

$$\begin{aligned} &\bigg| \int_0^{R_0} \int_{(S_f+B(0,r))\cap E_n} \int_0^1 \int_{B(x,r)} \frac{1}{V_{x,r}} \Big[w(y)F(\langle \Theta^+_{x,r,u}(w), f\rangle) \\ &\qquad + (1-w(y))F(\langle \Theta^-_{x,r,u}(w), f\rangle) - F(\langle w, f\rangle)\Big] \mathrm{d}y\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \bigg| \\ &\leq \|F'\| \|f\| \int_0^{R_0} \int_{(S_f+B(0,r))\cap E_n} \int_0^1 u \mathrm{Vol}(B(x,r) \cap S_f)\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \\ &\leq C \mathrm{Vol}(S_f + B(0,R_0)) \|F'\| \|f\|, \end{aligned}$$

where we have bounded $\mathrm{Vol}(B(x,r) \cap S_f)$ by $c_d r^d$, which is independent of x, and then we have bounded the remaining integral of $\mathrm{d}x$ over $(S_f + B(0,r)) \cap E_n$ by $\mathrm{Vol}(S_f + B(0,R_0))$.

Exactly the same arguments control the selection part of the generator $\mathcal{L}_{(n)}$ and, combining with the above, this gives the tightness of the finite variation parts of the processes. As in (i), since $u^2 < u$, tightness of the quadratic variation of the martingale parts follows easily and the Aldous-Rebolledo criterion yields tightness of the sequence of processes $(M^{(n)})_{n\geq 1}$.

To check the uniform convergence of $\mathcal{L}_{(n)}\Psi_{F,f}$ to $\mathcal{L}\Psi_{F,f}$ (which takes the form of $\mathcal{L}_{(n)}$ but with E_n replaced by $\mathbb{R}^d$ in the domain of integration), notice that by Condition (4), by taking R_0 sufficiently large, the right hand side of (90) can be made arbitrarily small, independent of M (or w). This is enough to ensure that the missing contribution of the events centered *outside* E_n is negligible, that is that

$$\Big| \int_0^\infty \int_{(S_f+B(0,r))\cap E_n^c} \int_0^1 \int_{B(x,r)} \frac{1}{V_{x,r}} \Big[w(y)F(\langle \Theta^+_{x,r,u}(w), f\rangle) \\ + (1-w(y))F(\langle \Theta^-_{x,r,u}(w), f\rangle) - F(\langle w, f\rangle)\Big] \mathrm{d}y\, \nu_r(\mathrm{d}u)\, \mathrm{d}x\, \mu(\mathrm{d}r) \Big| \tag{91}$$

converges to zero uniformly in M as $n \to \infty$, and so does the selection term.

By the same duality argument (based on Proposition 1.5) that we used in (i), the martingale problem associated with $\mathcal{L}$ has at most one solution, and so Theorem 4.8.10 in [EK86] yields the desired convergence. Theorem 1.3 is proved. □

B Continuity estimates in the fixed radius case

In this section, we state the continuity estimates for the scaled measures M^n_T required in the proof of Theorem 1.8. Because their proof is an adaptation of the (long and slightly more involved) proof of Proposition C.1(ii), we do not give it here and instead refer to Appendix C. These estimates have the same flavour as the one dimensional estimates derived in [MT95] for the convergence of the local densities of 1's in the long range voter or contact process.

Proposition B.1. *Under the conditions of Theorem 1.8, for every $T > 0$ there exists $a, \lambda, v, C > 0$ such that for every $n \geq 1$, $z_1, z_2 \in \mathbb{R}^d$ such that $|z_1 - z_2| < 1$ and $\epsilon \in (0,1)$,*

$$\mathbb{E}\left[\left| \frac{1}{V_\epsilon} \int_{\mathbb{R}^d} w^n_T(x)(\mathbf{1}_{\{|x-z_1|<\epsilon\}} - \mathbf{1}_{\{|x-z_2|<\epsilon\}})\mathrm{d}x \right|\right] \\ \leq Cn^{-a} + C\tau + C\big(|z_1-z_2|^{1/4} + \tau^{1/2}\big)e^{\lambda(|z_1|+\epsilon)} + Cn^{-(d-1)/6}\tau^{(2-d)/4} \tag{92}$$

where

$$\tau = \tau(n, z_1, z_2) = n^{-v} \vee |z_1 - z_2|^{2/(d+1)},$$

and ϵ can depend on n (as long as $\epsilon_n \leq 1$).

C Continuity estimates in the stable radius case

As in the previous section, our aim in this section is to obtain some continuity estimates for the measure M^n_T (this time in the stable radius case), which are valid for fixed (large) n. Since in the stable radius case, we also need to compare the local densities of type-1 individuals over balls of radius $n^{-\beta}$ to the densities over balls of radius $\mathcal{O}(\log n)n^{-\beta}$, Proposition C.1 below is more complete than Proposition B.1. Lemma 5.2 will then follow as a corollary of item (i).

Proposition C.1. *Suppose the conditions of Theorem 1.9 are satisfied. Fix $T > 0$. Then,*
(i) For every $z \in \mathbb{R}^d$, $t \in [0,T]$ and $n^{-\beta} \le \epsilon_n < \epsilon'_n \le 1$,

$$\mathbb{E}\left[\left|\frac{1}{V_{\epsilon_n}}\int_{B(z,\epsilon_n)} w_t^n(y)\mathrm{d}y - \frac{1}{V_{\epsilon'_n}}\int_{B(z,\epsilon'_n)} w_t^n(y)\mathrm{d}y\right|\right]$$
$$\le Cn^{-a} + C\tau_1 + C\epsilon'_n(\log n)^d \tau_1^{1-\frac{d+1}{\alpha}} + (\log n)^{d/2} n^{\frac{\beta(\alpha-d)-\gamma}{2}}\Big[\epsilon'^2_n \tau_1^{1-\frac{2(d+1)}{\alpha}} + \epsilon'_n n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\tau_1^{1-\frac{d+1}{\alpha}}\Big]^{1/2} \tag{93}$$

for some $a, C > 0$ independent of n (but dependent on T), where

$$\tau_1 = \tau_1(n) = n^{-\beta(2-\alpha)/(2(d+1))}. \tag{94}$$

(ii) For every $|z_1 - z_2| < 1$, $t \in [0,T]$ and $\epsilon \in (0,1)$,

$$\mathbb{E}\left[\left|\frac{1}{V_\epsilon}\int_{\mathbb{R}^d} w_t^n(x)(\mathbf{1}_{\{|x-z_1|<\epsilon\}} - \mathbf{1}_{\{|x-z_2|<\epsilon\}})\mathrm{d}x\right|\right]$$
$$\le Cn^{-a} + C\tau_2 + C\big(|z_1-z_2|^{1/4} + (\tau_2)^{1/2}\big)e^{\lambda(|z_1|+\epsilon)} + C\big(n^{-\beta(d-1)}\tau_2^{1-d/\alpha}\big)^{1/2} \tag{95}$$

for some $a, \lambda, C > 0$ independent of n (but dependent on T), where

$$\tau_2 = \tau_2(n, z_1, z_2) = n^{-\beta(2-\alpha)d/(4(d+1))} \vee |z_1 - z_2|^{\alpha/(d+1)}, \tag{96}$$

and ϵ can depend on n (as long as $\epsilon_n \le 1$).

In particular, (ii) implies uniform continuity of the limiting process of allele frequencies. That is:

Corollary C.2. *Suppose the conditions of Theorem 1.9 are satisified and fix $T > 0$. Then for every $|z_1 - z_2| < 1$, $t \in [0,T]$ and $\epsilon \in (0,1)$,*

$$\limsup_{n\to\infty} \mathbb{E}\left[\left|\frac{1}{V_\epsilon}\int_{\mathbb{R}^d} w_t^n(x)(\mathbf{1}_{\{|x-z_1|<\epsilon\}} - \mathbf{1}_{\{|x-z_2|<\epsilon\}})\mathrm{d}x\right|\right]$$
$$\le C|z_1-z_2|^{(\alpha-1)/4}\mathbf{1}_{\{d=1\}} + C|z_1-z_2|^{\alpha/(d+1)}\mathbf{1}_{\{d\ge 2\}} + C\big(|z_1-z_2|^{1/4} + |z_1-z_2|^{\alpha/(2(d+1))}\big)e^{\lambda(|z_1|+\epsilon)},$$

where C depends on T.

Before proving Proposition C.1, let us show how it implies Lemma 5.2.

Proof of Lemma 5.2.
Set $\epsilon_n = n^{-\beta}$ and $\epsilon'_n \in [n^{-\beta}, n^{-\beta}\log n]$ in (i). Then

$$\epsilon'^2_n \tau_1^{1-\frac{2(d+1)}{\alpha}} = (\log n)^2 n^{-2\beta} n^{\frac{\beta(2-\alpha)}{2(d+1)} - \frac{\beta(2-\alpha)}{\alpha}},$$

and it is straightforward to check that the exponent of n in the right hand side is negative for any $\alpha \in (1,2)$. Moreover,

$$\epsilon'_n(\log n)^d \tau_1^{1-\frac{d+1}{\alpha}} \le (\log n)^a n^{-\beta(1-\frac{2-\alpha}{2}(\frac{1}{\alpha}-\frac{1}{d+1}))}$$

for some $a > 0$, and again one can check that the exponent of n is negative in all dimensions. Thus the right hand side of (93) tends to zero and the lemma follows. □

The rest of this section is devoted to the proof of Proposition C.1. Note that the different lemmas that appear in this proof will be shown later in Appendix C.3.

Proof of Proposition C.1.

We define for $x \in \mathbb{R}^d$,

$$\sqcap_r(x) = \frac{1}{V_r}\mathbf{1}_{\{|x|\leq r\}},$$

$\sqcap_r^{*k}$ to be the k-fold convolution of $\sqcap_r$ and $\tilde{w}^n(x;r) = \frac{1}{V_r}\int_{B(x,r)} w^n(y)\mathrm{d}y$. Recall the expression (53) for the generator of M^n. For $\varphi \in \mathbb{L}^1(\mathbb{R}^d)$, we follow our usual strategy of writing the value of $\langle w_T^n, \varphi\rangle$ as a sum of drift and martingale terms: for any representative w_t^n of the density of each M_t^n, we have

$$\begin{aligned}
\langle w_T^n, \varphi\rangle &= \langle w_0^n, \varphi\rangle + \mathcal{M}_T^{n,\varphi} + u_n n^{1-\beta\alpha}\int_0^T\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2} \\
&\qquad \frac{1}{V_r^2}\Big\{w_t^n(y)(1+s_n w_t^n(z))\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \varphi\rangle \\
&\qquad\qquad -\big(1-w_t^n(y)+s_n(1-w_t^n(y)w_t^n(z))\big)\langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle\Big\}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\mathrm{d}t \\
&= \langle w_0^n, \varphi\rangle + \mathcal{M}_T^{n,\varphi} + u\int_0^T\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)^2}\frac{1}{V_r^2}\big\{w_t^n(y)\langle \mathbf{1}_{B(x,r)}, \varphi\rangle \qquad (97)\\
&\qquad -\langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle + s_n(w_t^n(y)w_t^n(z)\langle \mathbf{1}_{B(x,r)}, \varphi\rangle - \langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle)\big\}\,\mathrm{d}y\mathrm{d}z\mathrm{d}r\mathrm{d}x\mathrm{d}t
\end{aligned}$$

(since $u_n n^{1-\beta\alpha} = u$) where $(\mathcal{M}_T^{n,\varphi})_{T\geq 0}$ is a mean zero martingale. The first term in the integrand in (97) is equal to:

$$\begin{aligned}
&\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{B(x,r)}\frac{1}{V_r}\{w_t^n(y)\langle \mathbf{1}_{B(x,r)}, \varphi\rangle - \langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle\}\mathrm{d}y\mathrm{d}r\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\frac{1}{V_r}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_{\{|x-y|\leq r\}}\mathbf{1}_{\{|x-z|\leq r\}}\{w_t^n(y)\varphi(z) - w_t^n(z)\varphi(z)\}\mathrm{d}z\mathrm{d}y\mathrm{d}r\mathrm{d}x \\
&= \int_{n^{-\beta}}^\infty\int_{\mathbb{R}^d}\frac{V_r}{r^{d+1+\alpha}}\{(\sqcap_r^{*2} * w_t^n)(z)\varphi(z) - w_t^n(z)\varphi(z)\}\mathrm{d}z\mathrm{d}r \\
&= \int_{\mathbb{R}^d} w_t^n(z)\int_{n^{-\beta}}^\infty \frac{V_r}{r^{d+1+\alpha}}\{(\sqcap_r^{*2} * \varphi)(z) - \varphi(z)\}\mathrm{d}r\mathrm{d}z. \qquad (98)
\end{aligned}$$

The second term in the integrand in (97) is equal to

$$\begin{aligned}
&s_n\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}(\tilde{w}_t^n(x;r)^2\langle \mathbf{1}_{B(x,r)}, \varphi\rangle - \langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle)\mathrm{d}r\mathrm{d}x \\
&\quad = s_n\int_{(\mathbb{R}^d)^2}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\mathbf{1}_{\{|x-y|<r\}}(\tilde{w}_t^n(x;r)^2 - w_t^n(y))\varphi(y)\mathrm{d}y\mathrm{d}x\mathrm{d}r.
\end{aligned}$$

Since $u_n^2 n^{1-\beta\alpha} = u^2 n^{-\gamma} = u^2 n^{-(\alpha-1)/(2\alpha-1)}$, the martingale term in (97) has quadratic variation

$$\begin{aligned}
[\mathcal{M}^{n,\varphi}]_T &= u^2 n^{-\gamma}\int_0^T\int_{\mathbb{R}^d}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\Big\{\tilde{w}_t^n(x;r)(1+s_n\tilde{w}_t^n(x;r))\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \varphi\rangle^2 \\
&\qquad +\big(1-\tilde{w}_t^n(x;r)+s_n(1-\tilde{w}_t^n(x;r)^2)\big)\langle \mathbf{1}_{B(x,r)}w_t^n, \varphi\rangle^2\Big\}\,\mathrm{d}r\mathrm{d}x\mathrm{d}t.
\end{aligned}$$

It is convenient to replace this martingale problem by a mild version, obtained by replacing φ by the time dependent function $\zeta_t^n(x,z,\epsilon)$ chosen to solve

$$\partial_t \zeta_t^n(x;z,\epsilon) = \int_{n^{-\beta}}^{\infty} \frac{uV_r}{r^{d+1+\alpha}} \Big[(\sqcap_r^{*2} * \zeta_t^n(\cdot;z,\epsilon))(x) - \zeta_t^n(x;z,\epsilon) \Big] \mathrm{d}r$$

with initial condition $\zeta_0^n(\cdot;z,\epsilon)$. That is $\zeta_t^n(\cdot;z,\epsilon)$ is the density at time t of the d-dimensional Lévy process, $(X_t^n)_{t\geq 0}$, with initial distribution $\zeta_0^n(\cdot;z,\epsilon)$, zero drift, no Brownian component, and Lévy measure

$$\nu^n(\mathrm{d}x) = \int_{n^{-\beta}}^{\infty} \frac{uV_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x) \mathrm{d}r \mathrm{d}x$$

for $x \in \mathbb{R}^d$ (in particular, $\zeta_t^n(x,z,\epsilon) \in \mathbb{L}^1(\mathbb{R}^d)$). Here we assume that for any $n \in \mathbb{N}$, $z \in \mathbb{R}^d$ and $\epsilon > 0$, $\zeta_0^n(\cdot;z,\epsilon) = \zeta_0^n(\cdot - z;0,\epsilon)$ and that the support of $\zeta_0^n(\cdot;0,\epsilon)$ is included in $B(0,\epsilon)$. Of course, the particular example we have in mind is $\zeta_0^n(\cdot;z,\epsilon) = \frac{1}{V_\epsilon}\mathbf{1}_{\{|\cdot - z|<\epsilon\}}$. The parameter ϵ can be taken to depend on n. We observe that ν^n is radially symmetric. Let

$$\begin{aligned}
a^n(x;r) &= r^{-d} \int_{\mathbb{R}^d} \mathbf{1}_{\{|x-y|<r\}} (\tilde{w}^n(y;r)^2 - w^n(x)) \mathrm{d}y \\
b^n(x;r) &= \tilde{w}^n(x;r)(1 + s_n \tilde{w}^n(x;r)) \\
c^n(x;r) &= 1 - \tilde{w}^n(x;r) + s_n(1 - \tilde{w}^n(x;r)^2).
\end{aligned}$$

Notice that a^n, b^n and c^n are all uniformly (in n, x and r) bounded between constants. Suppose that we know the exponential decay of $\zeta_{T-t}^n(\cdot;z,\epsilon)$ (which we prove in Lemma C.3), then substituting in the martingale problem in the usual way, we obtain

$$\begin{aligned}
\langle w_T^n, \zeta_0^n(\cdot;z,\epsilon)\rangle &= \langle w_0^n, \zeta_T^n(\cdot;z,\epsilon)\rangle + \mathcal{M}_T^{n,\zeta_0^n(\cdot;z,\epsilon)} \\
&\quad + us_n \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x;r) \zeta_{T-t}^n(x;z,\epsilon) \mathrm{d}x \mathrm{d}r \mathrm{d}t \qquad (99) \\
\Big[\mathcal{M}^{n,\zeta_0^n(\cdot;z,\epsilon)}\Big]_T &= u^2 n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \Big\{ b_t^n(x;r) \langle \mathbf{1}_{B(x,r)}(1-w_t^n), \zeta_{T-t}^n(\cdot;z,\epsilon)\rangle^2 \\
&\quad + c_t^n(x;r) \langle \mathbf{1}_{B(x,r)} w_t^n, \zeta_{T-t}^n(\cdot;z,\epsilon)\rangle^2 \Big\} \mathrm{d}x \mathrm{d}r \mathrm{d}t \\
&= u^2 n^{-\gamma} \int_0^T \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \Big\{ b_t^n(x;r) \left(\int_{B(x,r)} (1 - w_t^n(y)) \zeta_{T-t}^n(y;z,\epsilon) \mathrm{d}y \right)^2 \\
&\quad + c_t^n(x;r) \left(\int_{B(x,r)} w_t^n(y) \zeta_{T-t}^n(y;z,\epsilon) \mathrm{d}y \right)^2 \Big\} \mathrm{d}x \mathrm{d}r \mathrm{d}t. \qquad (100)
\end{aligned}$$

In order to control the different terms appearing in (99) and (100), we are going to need to establish continuity estimates for ζ^n. In preparation for this, note that $(X_t^n)_{t\geq 0}$ is a continuous time random walk with jump rate

$$A = \int_{n^{-\beta}}^{\infty} \frac{uV_r}{r^{d+1+\alpha}} \mathrm{d}r = V_1 n^{\alpha\beta}.$$

To describe the corresponding jump chain, let R_k be i.i.d. $\mathbb{R}^d$-valued random variables distributed according to $\frac{V_1}{A} r^{-(1+\alpha)} \mathbf{1}_{\{r>n^{-\beta}\}} \mathrm{d}r$, $Z_{1,k}$ and $Z_{2,k}$ be independent uniformly distributed random variables in $B(0,1)$, and $Y_k = R_k(Z_{1,k} + Z_{2,k})$. Then we can write

$$X_t^n = X_0^n + \sum_{k=1}^{K_t} Y_k, \tag{101}$$

where K_t is a Poisson random variable with parameter At. We define f_Y as the density of Y_1, f_Y^{*k} to be the k-fold convolution of f_Y,

$$\begin{aligned} q_t^{n,\{k\}}(x) &= f_Y^{*k}(x)\mathbb{P}[K_t = k] = e^{-At}\frac{(At)^k}{k!} f_Y^{*k}(x) \\ q_t^n(x) &= \sum_{k=1}^{\infty} q_t^{n,\{k\}}(x). \end{aligned} \tag{102}$$

Then,

$$\zeta_t^n(x; z, \epsilon) = \zeta_0^n(x; z, \epsilon)e^{-At} + (\zeta_0^n(\cdot; z, \epsilon) * q_t^n(\cdot))(x).$$

Our estimates will involve splitting into two cases, according to whether the walk has taken greater or fewer than L steps in the interval $[0, t]$ and so it will be convenient to define $q_t^{n,I} = \sum_{k\in I} q^{n,\{k\}}$ for $I \subset [1,\infty)$, $\zeta_t^{n,\{k\}}(\cdot; z, \epsilon) = \zeta_0^n(\cdot; z, \epsilon) * q_t^{n,\{k\}}(\cdot)$, and $\zeta_t^{n,I} = \sum_{k\in I} \zeta_t^{n,\{k\}}$ for $I \subset [0,\infty)$.

Since the number of jumps made by the walk in $[0, t]$ has mean proportional to $n^{\alpha\beta}$, with probability tending to one as $n \to \infty$ it will take at least $n^{c\alpha\beta}$ steps for any $c \in (0, 1)$. We define $c_1 := (\alpha - 1)/(2\alpha) \in (0, 1)$ and set

$$L = n^{c_1\alpha\beta/2}.$$

In Section C.3, we shall prove a sequence of lemmas that control the behaviour of the random walk. In particular, we establish the following. For every $t \geq 0$, let q_t be the density function of value at time t of the symmetric α-stable process starting at 0 and with Laplace exponent

$$\psi(\theta) := \int_{\mathbb{R}^d} \left(e^{i\theta\cdot x} - 1\right)\nu(\mathrm{d}x),$$

where

$$\nu(\mathrm{d}x) := \int_0^\infty \frac{uV_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x)\mathrm{d}r\mathrm{d}x.$$

(Note that this process is the one appearing in Lemma 5.1.)

Lemma C.3. *Let $\|f\|_\lambda = \sup_x |f(x)|e^{\lambda|x|}$. Let $c_2 \in (0, \alpha)$ be a constant. Recall $L = n^{c_1\alpha\beta/2}$, with $c_1 = \frac{\alpha-1}{2\alpha}$. For $x, y, z \in \mathbb{R}^d$ and n,*

(i) If $M \geq 2$ and $t \in [n^{-c_2\beta(2-\alpha)/(2(d+1))}, T]$, then

$$|q_t^{n,[M,\infty)}(x) - q_t(x)| \leq C_{d,T} n^{-\beta(2-\alpha)d/(2(d+1))} + C_d n^{\beta d}(a^{M-1} + \mathbb{P}[K_t < M])$$

for some $a \in (0, 1)$ independent of M and T. Furthermore,

$$|q_t^{n,[L,\infty)}(x) - q_t(x)| \leq C_{d,T} n^{-\beta(2-\alpha)d/(2(d+1))}.$$

(ii) If $t > 0$, then $|q_t(x) - q_t(y)| \leq Ct^{-(d+1)/\alpha}|x - y|$.

(iii) *If* $t \in [n^{-c_2\beta(2-\alpha)/(2(d+1))}, T]$*, then*

$$|q_t^{n,[L,\infty)}(x) - q_t^{n,[L,\infty)}(y)| \leq Ct^{-(d+1)/\alpha}|x-y| + C_{d,T}n^{-\beta(2-\alpha)d/(2(d+1))}.$$

(iv) *If* $\lambda > 0$*,* $t \leq T$ *and* $|x| \geq 1$*, then* $q_t^{n,[1,\infty)}(x) \leq C_{\lambda,T}e^{-\lambda(|x|-1)}$*.*

(v) *If* $\lambda > 0$*,* $t \in [n^{-c_2\beta(2-\alpha)/(2(d+1))}, T]$ *and* $|y-z| \leq 1$*, then*

$$\|\zeta_t^{n,[L,\infty)}(\cdot;y,\epsilon) - \zeta_t^{n,[L,\infty)}(\cdot;z,\epsilon)\|_\lambda \leq C_{\lambda,d,T}e^{\lambda\epsilon}(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))})e^{\lambda|z|},$$

where ϵ *can depend on* n*.*

Recall the definitions

$$\begin{aligned} \tau_1 &= n^{-\beta(2-\alpha)/(2(d+1))}, \\ \tau_2 &= n^{-\beta(2-\alpha)d/(4(d+1))} \vee |z_1 - z_2|^{\alpha/(d+1)}, \end{aligned}$$

The quantity τ_1 (resp., τ_2) will be used in the bounds needed to prove Proposition C.1(i) (resp., (ii)). Observe that for $t \geq \tau_2$ and $|z_1 - z_2| < 1$, the estimate in the right hand side of Lemma C.3(v) is

$$\leq C_{\lambda,d,T}(|z_1 - z_2|^{1/2} + \tau_2)e^{\lambda\epsilon}e^{\lambda|z_1|}.$$

Since the organisations of the proofs are similar, we shall show Proposition C.1(i) and (ii) in parallel. In both cases, we set

$$\zeta_0^n(\cdot;z,\epsilon) := \frac{1}{V_\epsilon}\mathbf{1}_{B(z,\epsilon)}$$

(although most of the proof does not require a specific form for ζ_0^n), and we estimate

$$\begin{aligned} &(i) \qquad \langle w_T^n, \zeta_0^n(\cdot;z,\epsilon_n) - \zeta_0^n(\cdot;z,\epsilon'_n)\rangle, \\ &(ii) \qquad \langle w_T^n, \zeta_0^n(\cdot;z_1,\epsilon) - \zeta_0^n(\cdot;z_2,\epsilon)\rangle \end{aligned}$$

for the range of parameters stated in Proposition C.1, using (99) and (100).

C.1 Drift terms

Let us split the different terms into the case where K_t, the number of jumps of X^n by time t, is less than or larger than L. This first gives (using the fact that the function a_t^n is bounded uniformly in n, t, x, r):

$$\begin{aligned} &\left| us_n \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} a_t^n(x;r)\big(\zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n) - \zeta_{T-t}^{n,[0,L)}(x;z,\epsilon'_n)\big)\mathrm{d}x\mathrm{d}r\mathrm{d}t \right| \\ &\quad \leq Cus_n \int_0^T \int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}} \int_{\mathbb{R}^d} \big(\zeta_{T-t}^{n,[0,L)}(x;z,\epsilon_n) + \zeta_{T-t}^{n,[0,L)}(x;z,\epsilon'_n)\big)\mathrm{d}x\mathrm{d}r\mathrm{d}t \\ &\quad \leq Cus_n n^{\alpha\beta} \int_0^T \mathbb{P}[K_t < L]\,\mathrm{d}t \leq Cn^{-(1-c_1)\alpha\beta} \end{aligned} \tag{103}$$

by Lemma C.6 (which controls $\mathbb{P}[K_t < L]$) and the fact that, by definition, $s_n n^{\alpha\beta} \equiv \sigma$. The same estimate holds for (ii) and the corresponding integral.

Next, let us split the remaining integral into an integral over large and small times. We can write

$$\begin{aligned}\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon_n)-\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon'_n)&=\int_{\mathbb{R}^d}\big(\zeta^n_0(x';z,\epsilon_n)-\zeta^n_0(x';z,\epsilon'_n)\big)q^{n,[L,\infty)}_{T-t}(x-x')\mathrm{d}x'\\&=\int_{\mathbb{R}^d}\zeta^n_0(x';z,\epsilon_n)\big(q^{n,[L,\infty)}_{T-t}(x-x')-q^{n,[L,\infty)}_{T-t}(x-z)\big)\mathrm{d}x'\\&\quad-\int_{\mathbb{R}^d}\zeta^n_0(x';z,\epsilon'_n)\big(q^{n,[L,\infty)}_{T-t}(x-x')-q^{n,[L,\infty)}_{T-t}(x-z)\big)\mathrm{d}x'.\end{aligned}$$

(The extra terms cancel since $\int_{\mathbb{R}^d}\zeta^n_0(x',z',\epsilon_n)\mathrm{d}x'=1$ for all choices of ϵ_n.) Since the second term above will be bounded in the same way as the first term, let us just consider the first one. We have by Lemma C.3(iii) and (iv) (recalling also that the support of $\zeta^n_0(\cdot;z,\epsilon)$ is contained in $B(z,\epsilon)$):

$$\begin{aligned}&Cus_n\int_0^{T-\tau_1}\int_{n^{-\beta}}^\infty\frac{1}{r^{1+\alpha}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\zeta^n_0(x';z,\epsilon_n)\big|q^{n,[L,\infty)}_{T-t}(x-x')-q^{n,[L,\infty)}_{T-t}(x-z)\big|\mathrm{d}x'\mathrm{d}x\mathrm{d}r\mathrm{d}t\\&\le Cs_nn^{\alpha\beta}\int_{\tau_1}^T\int_{B(z,\log n)}\int_{\mathbb{R}^d}\zeta^n_0(x';z,\epsilon_n)\big(t^{-(d+1)/\alpha}|z-x'|+C_{d,T}n^{-\beta(2-\alpha)d/(2(d+1))}\big)\mathrm{d}x'\mathrm{d}x\mathrm{d}t\\&\quad+C's_nn^{\alpha\beta}\int_{\tau_1}^T\int_{B(z,\log n)^c}\int_{\mathbb{R}^d}\zeta^n_0(x';z,\epsilon_n)e^{-|x-z|}\mathrm{d}x'\mathrm{d}x\mathrm{d}t\\&\le C\epsilon_n(\log n)^d\int_{\tau_1}^T t^{-(d+1)/\alpha}\mathrm{d}t+C_{d,T}T(\log n)^dn^{-\beta(2-\alpha)d/(2(d+1))}\big)+C'T\frac{(\log n)^{d-1}}{n}\\&\le C\Big(\epsilon_n(\log n)^d\tau_1^{1-\frac{d+1}{\alpha}}+(\log n)^dn^{-\beta(2-\alpha)d/(2(d+1))}+\frac{(\log n)^{d-1}}{n}\Big).\end{aligned}\tag{104}$$

For (ii), the corresponding calculation is different and uses Lemma C.3(v) with an arbitrary $\lambda>0$:

$$\begin{aligned}&\left|us_n\int_0^{T-\tau_2}\int_{n^{-\beta}}^\infty\frac{1}{r^{1+\alpha}}\int_{\mathbb{R}^d}a^n_t(x;r)\big[\zeta^{n,[L,\infty)}_{T-t}(x;z_1,\epsilon)-\zeta^{n,[L,\infty)}_{T-t}(x;z_2,\epsilon)\big]\mathrm{d}x\mathrm{d}r\mathrm{d}t\right|\\&\le Cus_nn^{\alpha\beta}\sup_{t\in[0,T-\tau_2]}\big\|\zeta^{n,[L,\infty)}_{T-t}(x;z_1,\epsilon)-\zeta^{n,[L,\infty)}_{T-t}(x;z_2,\epsilon)\big\|_\lambda\int_0^{T-\tau_2}\int_{\mathbb{R}^d}e^{-\lambda|x|}\mathrm{d}x\mathrm{d}t\\&\le C_{\lambda,d,T}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda\epsilon}e^{\lambda|z_1|}.\end{aligned}\tag{105}$$

Finally, it remains to bound the integral corresponding to small $(T-t)$'s. For (i), we obtain

$$\begin{aligned}&\left|us_n\int_{T-\tau_1}^T\int_{n^{-\beta}}^\infty\frac{1}{r^{1+\alpha}}\int_{\mathbb{R}^d}a^n_t(x;r)\big[\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon_n)-\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon'_n)\big]\mathrm{d}x\mathrm{d}r\mathrm{d}t\right|\\&\le\ Cs_n\int_{T-\tau_1}^T\int_{n^{-\beta}}^\infty\frac{1}{r^{1+\alpha}}\int_{\mathbb{R}^d}\big(\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon_n)+\zeta^{n,[L,\infty)}_{T-t}(x;z,\epsilon'_n)\big)\mathrm{d}x\mathrm{d}r\mathrm{d}t\\&\le\ Cs_nn^{\alpha\beta}\tau_1=C\tau_1.\end{aligned}\tag{106}$$

The same result obviously holds for (ii), with τ_1 replaced by τ_2.

Likewise, for the terms involving the initial condition w^n_0, similar arguments using Lemma C.3(i) and (v), and Lemma C.6 lead to

$$\big|\langle w^n_0,\zeta^n_T(\cdot;z,\epsilon_n)-\zeta^n_T(\cdot;z,\epsilon'_n)\rangle\big|\le Ce^{-n^{c_1\alpha\beta/2}}+Cn^{-\beta(2-\alpha)d/(2(d+1))},$$

and

$$\left|\langle w_0^n, \zeta_T^n(\cdot; z_1, \epsilon) - \zeta_T^n(\cdot; z_2, \epsilon)\rangle\right| \leq Ce^{-n^{c_1\alpha\beta/2}} + Ce^{\lambda(|z_1|+\epsilon)}\big(\tau_2 + |z_1 - z_2|^{1/2}\big).$$

C.2 Martingale terms

Now we turn to the martingale terms. As before, we first consider the case $K_t < L$. We shall estimate the term involving b^n, but the same approach can also be applied to the terms involving c^n. We have

$$\begin{aligned}
&\left|u^2 n^{-\gamma}\int_0^T\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d} b_t^n(x;r)\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \zeta_{T-t}^{n,[0,L)}(\cdot; z, \epsilon_n)\rangle^2 \mathrm{d}z\mathrm{d}r\mathrm{d}t\right| \\
&\leq Cn^{-\gamma}\int_0^T\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int \zeta_t^{n,[0,L)}(y;z,\epsilon_n)\int \mathbf{1}_{\{|y-x|<r\}}\int \mathbf{1}_{\{|y'-x|<r\}}\zeta_t^{n,[0,L)}(y';z,\epsilon_n)\mathrm{d}y'\mathrm{d}x\mathrm{d}y\mathrm{d}r\mathrm{d}t \\
&\leq Cn^{-\gamma}\int_0^T\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int \zeta_t^{n,[0,L)}(y;z,\epsilon)\int \mathbf{1}_{\{|x-y|<r\}}\mathrm{d}x\mathrm{d}y\mathrm{d}r\mathrm{d}t \\
&\leq Cn^{-\gamma}\int_0^T\int_{n^{-\beta}}^\infty \frac{1}{r^{1+\alpha}}\int \zeta_t^{n,[0,L)}(y;z,\epsilon)\mathrm{d}y\mathrm{d}r\mathrm{d}t \\
&\leq Cn^{-\gamma}n^{\alpha\beta}\int_0^T \mathbb{P}[K_t < L]\mathrm{d}t \leq Cn^\beta n^{-(1-(\alpha-1)/(2\alpha))\alpha\beta} = Cn^{-(\alpha-1)\beta/2} \qquad (107)
\end{aligned}$$

by Lemma C.6. Of course, this inequality holds for (i) and (ii).

Now we turn to

$$\left|u^2 n^{-\gamma}\int_0^{T-\tau_1}\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d} b_t^n(x;r)\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot;z,\epsilon_n) - \zeta_{T-t}^{n,[L,\infty)}(\cdot;z,\epsilon_n')\rangle^2 \mathrm{d}x\mathrm{d}r\mathrm{d}t\right|.$$

Once again we write

$$\begin{aligned}
\zeta_t^{n,[L,\infty)}(y;z,\epsilon_n) - \zeta_t^{n,[L,\infty)}(y;z,\epsilon_n') &= \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n)\big(q_t^{n,[L,\infty)}(y-x') - q_t^{n,[L,\infty)}(y-z)\big)\mathrm{d}x' \\
&\quad - \int_{\mathbb{R}^d} \zeta_0^n(x';z,\epsilon_n')\big(q_t^{n,[L,\infty)}(y-x') - q_t^{n,[L,\infty)}(y-z)\big)\mathrm{d}x'.
\end{aligned}$$

This gives us

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d} b_{T-t}^n(x;r)\langle \mathbf{1}_{B(x,r)}(1-w_{T-t}^n), \zeta_t^{n,[L,\infty)}(\cdot;z,\epsilon_n) - \zeta_t^{n,[L,\infty)}(\cdot;z,\epsilon_n')\rangle^2 \mathrm{d}x\Big| \\
&\leq C\int_{(\mathbb{R}^d)^3} \mathbf{1}_{\{|x-y|\leq r\}}\mathbf{1}_{\{|x-y'|\leq r\}}\Bigg[\bigg(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)\big|q_t^{n,[L,\infty)}(y-x') - q_t^{n,[L,\infty)}(y-z)\big|\mathrm{d}x'\bigg) \\
&\qquad \times \bigg(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)\big|q_t^{n,[L,\infty)}(y'-x') - q_t^{n,[L,\infty)}(y'-z)\big|\mathrm{d}x'\bigg) + S_t^n\Bigg]\mathrm{d}y'\mathrm{d}y\mathrm{d}x,
\end{aligned}$$

where S_t^n is the sum of the remaining three terms comprising the squared integral on the first line. Since all these terms behave in the same way, we shall only bound the first one. Writing as before $V_r(y,y')(\leq C_d r^d)$ for the volume of $B(y,r)\cap B(y',r)$, and using Fubini's theorem, we can replace the integral over x by $V_r(y,y')$. Next, as in our estimates of the drift, we split the integrals over y, y' according to whether or not $y, y' \in B(z, \log n)$. This gives us the following first bound, using Lemma C.3(iii):

$$\begin{aligned}
&\int_{B(z,\log n)^2} V_r(y,y')\Big(\int_{\mathbb{R}^d}\zeta_0^n(x;z,\epsilon_n)\Big(t^{-\frac{d+1}{\alpha}}|z-x| + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)\mathrm{d}x\Big)\\
&\qquad\qquad \times\Big(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)\Big(t^{-\frac{d+1}{\alpha}}|z-x'| + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)\mathrm{d}x'\Big)\mathrm{d}y'\mathrm{d}y\\
&\le r^d\int_{B(z,\log n)^2}\mathbf{1}_{\{|y-y'|\le 2r\}}\Big(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)^2\mathrm{d}y'\mathrm{d}y\\
&\le r^d(r\wedge\log n)^d(\log n)^d\Big(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)^2.
\end{aligned}$$

Integrating over t and r, we obtain

$$\begin{aligned}
&n^{-\gamma}\int_{\tau_1}^T\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}r^d(r\wedge\log n)^d(\log n)^d\Big(\epsilon_n t^{-\frac{d+1}{\alpha}} + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)^2\mathrm{d}r\mathrm{d}t\\
&\le Cn^{-\gamma}(\log n)^d\Big(\int_{n^{-\beta}}^{\log n} r^{d-1-\alpha}\mathrm{d}r + (\log n)^d\int_{\log n}^\infty r^{-1-\alpha}\mathrm{d}r\Big)\\
&\qquad \times\Big(\epsilon_n^2\int_{\tau_1}^T t^{-\frac{2(d+1)}{\alpha}}\mathrm{d}t + 2\epsilon_n n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\int_{\tau_1}^T t^{-\frac{d+1}{\alpha}}\mathrm{d}t + Tn^{-\frac{\beta(2-\alpha)d}{(d+1)}}\Big)\\
&\le Cn^{-\gamma}(\log n)^d\big((\log n)^{d-\alpha} + n^{\beta(\alpha-d)}\big)\Big[\epsilon_n^2\tau_1^{1-\frac{2(d+1)}{\alpha}} + 2\epsilon_n n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\tau_1^{1-\frac{d+1}{\alpha}} + n^{-\frac{\beta(2-\alpha)d}{(d+1)}}\Big]. \qquad (108)
\end{aligned}$$

Secondly, considering the case where $y \in B(z, \log n)$ and $y' \in B(z, \log n)^c$ and using Lemma C.3(iii) and (iv), the corresponding integral is bounded by

$$\begin{aligned}
&\int_{B(z,\log n)}\int_{B(z,\log n)^c} V_r(y,y')\Big(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)\Big(t^{-\frac{d+1}{\alpha}}|z-x| + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)\mathrm{d}x\Big)\\
&\qquad\qquad \times\Big(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)e^{-|z-y'|}\mathrm{d}x'\Big)\mathrm{d}y'\mathrm{d}y\\
&\le C\Big(t^{-\frac{d+1}{\alpha}}\epsilon_n + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)\int_{B(z,\log n)^c}\int_{B(z,\log n)\cap B(y',2r)} r^d e^{-|z-y'|}\mathrm{d}y\mathrm{d}y'\\
&\le C\Big(t^{-\frac{d+1}{\alpha}}\epsilon_n + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)r^d(r\wedge\log n)^d\frac{(\log n)^{d-1}}{n}.
\end{aligned}$$

Integrating over t and r as well, we obtain

$$\begin{aligned}
&n^{-\gamma}\int_{\tau_1}^T\int_{n^{-\beta}}^\infty\frac{1}{r^{d+1+\alpha}}\Big(t^{-\frac{d+1}{\alpha}}\epsilon_n + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big)r^d(r\wedge\log n)^d\frac{(\log n)^{d-1}}{n}\mathrm{d}r\mathrm{d}t\\
&\qquad \le n^{-\gamma}\frac{(\log n)^{d-1}}{n}\Big[\epsilon_n\tau_1^{1-\frac{d+1}{\alpha}} + C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\Big]\Big[(\log n)^{d-\alpha} + n^{\beta(\alpha-d)}\Big]. \qquad (109)
\end{aligned}$$

The case where $y \in B(z, \log n)^c$ and $y' \in B(z, \log n)$ is treated in the same way. Finally, if

$y, y' \in B(z, \log n)^c$, Lemma C.3(iv) gives us the bound

$$\int_{(B(z,\log n)^c)^2} V_r(y,y')\Big(\int_{\mathbb{R}^d}\zeta_0^n(x;z,\epsilon_n)e^{-|z-y|}\mathrm{d}x\Big)\Big(\int_{\mathbb{R}^d}\zeta_0^n(x';z,\epsilon_n)e^{-|z-y'|}\mathrm{d}x'\Big)\mathrm{d}y'\mathrm{d}y$$
$$\le Cr^d\int_{B(z,\log n)^c}\int_{B(z,\log n)^c\cap B(y,2r)} e^{-|z-y|}e^{-|z-y'|}\mathrm{d}y'\mathrm{d}y$$
$$\le Cr^d\big(1\wedge r^d\big)\frac{(\log n)^{d-1}}{n}.$$

Integrating over t and r gives the bound

$$n^{-\gamma}\int_{\tau_1}^T\int_{n^{-\beta}}^\infty \frac{1}{r^{d+1+\alpha}}\,r^d\big(1\wedge r^d\big)\frac{(\log n)^{d-1}}{n}\,\mathrm{d}r\mathrm{d}t \le CTn^{-\gamma}\frac{(\log n)^{d-1}}{n}\,(n^{\beta(\alpha-d)}+1). \tag{110}$$

For the corresponding bound for (ii), the argument is again much shorter thanks to Lemma C.3(v):

$$\Big|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_{\{|y-x|<r\}}\mathbf{1}_{\{|z-x|<r\}}b_t^n(x;r)(1-w_t^n(y))(1-w_t^n(z))$$
$$(\zeta_{T-t}^{n,[L,\infty)}(y;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(y;z_2,\epsilon))(\zeta_{T-t}^{n,[L,\infty)}(z;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(z;z_2,\epsilon))\mathrm{d}z\mathrm{d}y\mathrm{d}x\Big|$$
$$\le \int_{\mathbb{R}^d}(\zeta_{T-t}^{n,[L,\infty)}(y;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(y;z_2,\epsilon))\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_{\{|y-x|<r\}}\mathbf{1}_{\{|z-x|<r\}}$$
$$|\zeta_{T-t}^{n,[L,\infty)}(z;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(z;z_2,\epsilon)|\mathrm{d}z\mathrm{d}x\mathrm{d}y$$
$$\le 2\sup_{t\in[0,T-\tau_2]}\|\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_2,\epsilon)\|_\lambda$$
$$\int_{\mathbb{R}^d}\zeta_{T-t}^{n,[L,\infty)}(y;z_1,\epsilon)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_{\{|y-x|<r\}}\mathbf{1}_{\{|z-x|<r\}}e^{-\lambda|z|}\mathrm{d}z\mathrm{d}x\mathrm{d}y$$
$$\le C\sup_{t\in[0,T-\tau_2]}\|\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_2,\epsilon)\|_\lambda\int_{\mathbb{R}^d}\zeta_{T-t}^{n,[L,\infty)}(y;z_1,\epsilon)(r^{2d}\wedge r^d)\mathrm{d}y$$
$$\le C_{\lambda,d,T}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)}(r^{2d}\wedge r^d),$$

which yields

$$\Big|u^2n^{-\gamma}\int_0^{T-\tau_2}\int_{n^{-\beta}}^\infty\frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d}b_t^n(x;r)\langle\mathbf{1}_{B(x,r)}(1-w_t^n),\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_1,\epsilon)-\zeta_{T-t}^{n,[L,\infty)}(\cdot;z_2,\epsilon)\rangle^2\mathrm{d}x\mathrm{d}r\mathrm{d}t\Big|$$
$$\le C_{\lambda,d,T}n^{-\gamma}\int_0^{T-\tau_2}\int_{n^{-\beta}}^\infty\frac{1}{r^{d+1+\alpha}}(r^{2d}\wedge r^d)(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)}\mathrm{d}r\mathrm{d}t$$
$$\le C_{\lambda,d,T}n^{-\gamma}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)}\left(\int_{n^{-\beta}}^1 r^{d-1-\alpha}\mathrm{d}r+\int_1^\infty r^{-1-\alpha}\mathrm{d}r\right)$$
$$\le C_{\lambda,d,T}n^{-\gamma}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)}(n^{(\alpha-d)\beta}+C)$$
$$\le C_{\lambda,d,T}n^{-(d-1)\beta}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)} \tag{111}$$

since $n^{(\alpha-1)\beta}n^{-\gamma}=1$.

For $t \in (T-\tau_1, T)$, we apply Lemma C.7 to obtain

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} b_t^n(x;r)\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot; z, \epsilon_n)\rangle^2 \mathrm{d}x\right| \\
&\quad\leq C\int \zeta_{T-t}^{n,[L,\infty)}(y;z,\epsilon_n)\int\int \mathbf{1}_{\{|y-x|<r\}}\mathbf{1}_{\{|y'-x|<r\}}\zeta_{T-t}^{n,[L,\infty)}(y';z,\epsilon_n)\mathrm{d}y'\mathrm{d}x\mathrm{d}y \\
&\quad\leq C_d\int \zeta_{T-t}^{n,[L,\infty)}(y;z,\epsilon_n)\int \mathbf{1}_{\{|y-x|<r\}}(1\wedge(((T-t)^{-d/\alpha}+e^{-n^{c_5}})r^d))\mathrm{d}x\mathrm{d}y \\
&\quad\leq C_d(r^d\wedge(((T-t)^{-d/\alpha}+e^{-n^{c_5}})r^{2d})),
\end{aligned}$$

which implies that

$$\begin{aligned}
&\left|u^2 n^{-\gamma}\int_{T-\tau_1}^{T}\int_{n^{-\beta}}^{\infty}\frac{1}{r^{d+1+\alpha}}\int_{\mathbb{R}^d} b_t^n(x;r)(\langle \mathbf{1}_{B(x,r)}(1-w_t^n), \zeta_{T-t}^{n,[L,\infty)}(\cdot;z,\epsilon_n)\rangle^2\mathrm{d}x\mathrm{d}r\mathrm{d}t\right| \\
&\leq Cn^{-\gamma}\int_0^{\tau_1}\int_{n^{-\beta}}^{\infty}\Big[r^{-1-\alpha}\wedge\big(t^{-d/\alpha}+e^{-n^{c_5}}\big)r^{d-1-\alpha}\Big]\mathrm{d}r\mathrm{d}t \\
&= Cn^{-\gamma}\int_{n^{-\beta}}^{\infty}\int_0^{r^\alpha\wedge\tau_1} r^{-1-\alpha}\mathrm{d}t\mathrm{d}r + Cn^{-\gamma}\int_{n^{-\beta}}^{\infty}\int_{r^\alpha\wedge\tau_1}^{\tau_1}\big(t^{-d/\alpha}+e^{-n^{c_5}}\big)r^{d-1-\alpha}\mathrm{d}t\mathrm{d}r \\
&= Cn^{-\gamma}\int_{n^{-\beta}}^{\tau_1^{1/\alpha}} r^{-1}\mathrm{d}r + C\tau_1 n^{-\gamma}\int_{\tau_1^{1/\alpha}}^{\infty} r^{-1-\alpha}\mathrm{d}r + Cn^{-\gamma}\int_{n^{-\beta}}^{\tau_1^{1/\alpha}}\int_{r^\alpha}^{\tau_1} t^{-d/\alpha}r^{d-1-\alpha}\mathrm{d}t\mathrm{d}r \\
&\leq Cn^{-\gamma}\log n + Cn^{-\gamma+\beta(\alpha-d)}\tau_1^{1-d/\alpha}. \qquad (112)
\end{aligned}$$

The same bound holds for (ii), with τ_1 replaced by τ_2.

Combining (107), (108), (109), (110) and (112) yields (recall that $\epsilon_n \leq \epsilon_n'$)

$$\begin{aligned}
\Big[\mathcal{M}^{n,\zeta_0^n(\cdot;z,\epsilon_n)-\zeta_0^n(\cdot;z,\epsilon_n')}\Big]_T &\leq Cn^{-(\alpha-1)\beta/2} + n^{-\gamma+\beta(\alpha-d)}\tau_1^{1-d/\alpha} \\
&\quad + n^{-\gamma}(\log n)^d\big((\log n)^{d-\alpha}+n^{\beta(\alpha-d)}\big)\Big[\epsilon_n'^2\,\tau_1^{1-\frac{2(d+1)}{\alpha}} \\
&\quad + 2\epsilon_n' n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}\tau_1^{1-\frac{d+1}{\alpha}} + n^{-\frac{\beta(2-\alpha)d}{(d+1)}}\Big],
\end{aligned}$$

while combining (107), (111) and (112) gives us

$$\begin{aligned}
\Big[\mathcal{M}^{n,\zeta_0^n(\cdot;z_1,\epsilon)-\zeta_0^n(\cdot;z_2,\epsilon)}\Big]_T &\leq Cn^{-(\alpha-1)\beta/2} + C_{\lambda,d,T}(|z_1-z_2|^{1/2}+\tau_2)e^{\lambda(|z_1|+\epsilon)} \\
&\quad + C_d n^{-\gamma+\beta(\alpha-d)}\tau_2^{(\alpha-d)/\alpha}.
\end{aligned}$$

Now, by the Burkholder-Davis-Gundy inequality ([Bur73]),

$$\mathbb{E}\left[\sup_{t\leq T}\left|\mathcal{M}_t^{n,\zeta_0^n(\cdot;z,\epsilon_n)-\zeta_0^n(\cdot;z,\epsilon_n')}\right|\right] \leq \Big[\mathcal{M}^{n,\zeta_0^n(\cdot;z,\epsilon_n)-\zeta_0^n(\cdot;z,\epsilon_n')}\Big]_T^{1/2}.$$

Combining this and the estimate for the drift term yields the desired result. □

C.3 Lemmas

We define for $\theta \in \mathbb{R}^d$,

$$\tilde{q}_t^{n,\{k\}}(\theta) = \mathbb{E}\big[e^{i\theta\cdot(X_t^n - X_0^n)}\mathbf{1}_{\{K_t=k\}}\big],$$

and correspondingly $\tilde{q}_t^{n,I}(\theta)$ for $I \subset [0,\infty)$, as well as $\tilde{q}_t^n(\theta) = \tilde{q}_t^{n,[0,\infty)}(\theta)$. Recall the representation of X^n using random walks in (101). As X^n has independent and stationary increments, the Lévy-Khintchine Formula (see e.g. Theorems 2.7.10 and 2.8.1 of [Sat99]) implies that

$$\tilde{q}_t^{n,[0,\infty)}(\theta) = \mathbb{E}\big[e^{i\theta\cdot(X_t^n - X_0^n)}\big] = e^{t\psi^n(\theta)},$$

where

$$\psi^n(\theta) = \int_{\mathbb{R}^d} (e^{i\theta\cdot x} - 1)\nu^n(\mathrm{d}x). \tag{113}$$

Similarly, we define the limiting Lévy measure

$$\nu(\mathrm{d}x) = \int_0^\infty \frac{uV_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x)\mathrm{d}r\mathrm{d}x,$$

as well as the corresponding function ψ,

$$\psi(\theta) = \int_{\mathbb{R}^d} (e^{i\theta\cdot x} - 1)\nu(\mathrm{d}x). \tag{114}$$

We observe that for all $t > 0$, $|e^{t\psi^n(\theta)}| \le 1$ and hence $|e^{t\psi(\theta)}| \le 1$.

Lemma C.4. *For all n, we have:*

(i) For all $\theta \in \mathbb{R}^d$, $|\psi^n(\theta) - \psi(\theta)| \le \frac{4^d}{3} n^{-\beta(2-\alpha)}|\theta|^2$.

(ii) For $|\theta| \le n^\beta$, $-\psi^n(\theta) \ge c|\theta|^\alpha$ for some positive constant $c = c_d$ independent of n. Hence $-\psi(\theta) \ge c|\theta|^\alpha$ for all θ.

Proof. Since ν is radially symmetric, (113) implies

$$\begin{aligned}
\psi^n(\theta) &= \frac{1}{2}\int_{\mathbb{R}^d} (e^{i\theta\cdot x} - 2 + e^{-i\theta\cdot x})\nu^n(\mathrm{d}x) = \frac{1}{2}\int_{\mathbb{R}^d} (e^{i\theta\cdot x/2} - e^{-i\theta\cdot x/2})^2 \nu^n(\mathrm{d}x) \\
&= -2\int_{\mathbb{R}^d} \sin^2(\theta\cdot x/2)\nu^n(\mathrm{d}x) = -2\int_{\mathbb{R}^d} \sin^2(\theta\cdot x/2)\int_{n^{-\beta}}^\infty \frac{V_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x)\mathrm{d}r\mathrm{d}x.
\end{aligned}$$

The calculations above can easily be repeated for X and ψ, then

$$\begin{aligned}
\frac{1}{2}|\psi^n(\theta) - \psi(\theta)| &= \left|\int_0^{n^{-\beta}} \frac{V_r}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \sin^2(\theta\cdot x/2)\int \sqcap_r(y)\sqcap_r(x-y)\mathrm{d}y\mathrm{d}x\mathrm{d}r\right| \\
&= \int_0^{n^{-\beta}} \frac{V_r}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \frac{\sin^2(\theta\cdot x/2)}{V_r^2} \int \mathbf{1}_{\{|y|<r\}}\mathbf{1}_{\{|x-y|<r\}}\mathrm{d}y\mathrm{d}x\mathrm{d}r \\
&\le \int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{\mathbb{R}^d} \sin^2(\theta\cdot x/2)\int \mathbf{1}_{\{|x|<2r\}}\mathrm{d}x\mathrm{d}r \\
&= \int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{|x|<2r} \sin^2(\theta\cdot x/2)\mathrm{d}x\mathrm{d}r.
\end{aligned}$$

Since $|\sin(x)| \leq |x|$ for all x, we have

$$\begin{aligned}\frac{1}{2}|\psi^n(\theta) - \psi(\theta)| &\leq \frac{1}{4}\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{|x|<2r} \left(\sum_{i=1}^{d} \theta_i x_i\right)^2 \mathrm{d}x\mathrm{d}r \\ &\leq \frac{1}{4}\int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{-2r}^{2r} \dots \int_{-2r}^{2r} \left(\sum_{i=1}^{d} \theta_i x_i\right)^2 \mathrm{d}x_1 \dots \mathrm{d}x_d \mathrm{d}r \\ &\leq \frac{1}{4}\sum_{i=1}^{d} \theta_i^2 \int_0^{n^{-\beta}} \frac{1}{r^{d+1+\alpha}} \int_{-2r}^{2r} \dots \int_{-2r}^{2r} x_i^2 \mathrm{d}x_1 \dots \mathrm{d}x_d \mathrm{d}r.\end{aligned}$$

The $(d+1)$-dimensional integral above is the same for all i (by symmetry), and is equal to

$$\int_0^{n^{-\beta}} \frac{(4r)^{d-1}}{r^{d+1+\alpha}} \int_{-2r}^{2r} x_1^2 \mathrm{d}x_1 \mathrm{d}r = \int_0^{n^{-\beta}} \frac{4^{d-1}}{r^{2+\alpha}} \frac{2}{3}(2r)^3 \mathrm{d}r = \frac{4^{d+1}}{3}\int_0^{n^{-\beta}} r^{1-\alpha}\mathrm{d}r = \frac{4^{d+1}}{3} n^{-\beta(2-\alpha)}.$$

Hence

$$|\psi^n(\theta) - \psi(\theta)| \leq \frac{4^d}{3} n^{-\beta(2-\alpha)} |\theta|^2,$$

as required by (i).

For (ii), we have

$$\begin{aligned}-\frac{1}{2}\psi^n(\theta) &= \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{V_r}{r^{d+1+\alpha}} \sqcap_r^{*2}(x) \mathrm{d}r\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha} V_r} \int_{\mathbb{R}^d} \mathbf{1}_{\{|y|<r\}} \mathbf{1}_{\{|x-y|<r\}} \mathrm{d}y\mathrm{d}r\mathrm{d}x \\ &\geq c_0 \int_{\mathbb{R}^d} \sin^2(\theta \cdot x/2) \int_{n^{-\beta}}^{\infty} \frac{1}{r^{d+1+\alpha}} \mathbf{1}_{\{|x|<r\}} \mathrm{d}r\mathrm{d}x,\end{aligned}$$

since the intersection of the disc $\{y : |y| < r\}$ and $\{y : |y-x| < r\}$ has volume larger than $c_0 V_r$ for some positive constant c_0 (dependent on d) if $|x| < r$. For $d = 1$ and $\theta_1 > 0$, we have

$$\begin{aligned}-\frac{1}{2}\psi^n(\theta_1) &\geq 2c_0 \int_{n^{-\beta}}^{\infty} \int_0^r \sin^2(\theta_1 x/2) \frac{1}{r^{2+\alpha}} \mathrm{d}x\mathrm{d}r \\ &= 2c_0 \int_0^{\infty} \int_{n^{-\beta}\vee x}^{\infty} \sin^2(\theta_1 x/2) \frac{1}{r^{2+\alpha}} \mathrm{d}r\mathrm{d}x \\ &= \frac{2c_0}{1+\alpha} \int_0^{\infty} \sin^2(\theta_1 x/2)(n^{-\beta} \vee x)^{-(1+\alpha)} \mathrm{d}x \\ &\geq \frac{2c_0}{1+\alpha} \int_{n^{-\beta}}^{\infty} \sin^2(\theta_1 x/2) x^{-(1+\alpha)} \mathrm{d}x \\ &= \frac{2c_0}{1+\alpha} \theta_1^{\alpha} \int_{\theta_1 n^{-\beta}}^{\infty} \sin^2(y/2) y^{-(1+\alpha)} \mathrm{d}y.\end{aligned}$$

Since $\theta_1 \leq n^{\beta}$, the integral in the above is bounded below by a constant. By symmetry, with thus obtain that for any θ such that $|\theta| \leq n^{\beta}$,

$$-\psi^n(\theta) \geq c|\theta|^{\alpha} \tag{115}$$

for some $c>0$. We can carry out a similar calculation for $d\geq 2$. Since ψ^n is radially symmetric, it suffices to consider $\theta=(\theta_1,0,\ldots,0)$ with $\theta_1>0$:

$$\begin{aligned}
-\frac{1}{2}\psi^n(\theta) &\geq c_0\int_{n^{-\beta}}^{\infty}\int_0^r\int_{|x|=\rho}\sin^2(\theta_1x_1/2)\frac{1}{r^{d+1+\alpha}}\mathrm{d}x\mathrm{d}\rho\mathrm{d}r\\
&= c_0\int_0^{\infty}\int_{n^{-\beta}\vee\rho}^{\infty}\int_{|x|=\rho}\sin^2(\theta_1x_1/2)\frac{1}{r^{d+1+\alpha}}\mathrm{d}x\mathrm{d}r\mathrm{d}\rho\\
&= \frac{c_0}{d+\alpha}\int_0^{\infty}\int_{|x|=\rho}\sin^2(\theta_1x_1/2)(n^{-\beta}\vee\rho)^{-(d+\alpha)}\mathrm{d}x\mathrm{d}\rho\\
&\geq \frac{c_0}{d+\alpha}\int_{n^{-\beta}}^{\infty}\int_{|x|=\rho}\sin^2(\theta_1x_1/2)\rho^{-(d+\alpha)}\mathrm{d}x\mathrm{d}\rho\\
&= \frac{c_0}{d+\alpha}\int_{n^{-\beta}}^{\infty}\int_{|y|=1}\sin^2(\rho\theta_1y_1/2)\rho^{-(1+\alpha)}\mathrm{d}y\mathrm{d}\rho\\
&= \frac{c_0}{d+\alpha}\theta_1^{\alpha}\int_{\theta_1n^{-\beta}}^{\infty}\int_{|y|=1}\sin^2(ry_1/2)r^{-(1+\alpha)}\mathrm{d}y\mathrm{d}r.
\end{aligned}$$

Since $\theta_1=|\theta|\leq n^{\beta}$, the double integral in the above is bounded below by a constant. Therefore we arrive at the same estimate as in (115) and we have proved (ii). □

Lemma C.5. *(i) Let $c_2\in(0,\alpha)$ be a constant. If $n^{-c_2\beta(2-\alpha)/(2(d+1))}\leq t\leq T$, then*

$$\int_{|\theta|\leq n^{\beta}}|(e^{t(\psi^n(\theta)-\psi(\theta))}-1)e^{t\psi(\theta)}|d\theta\leq C_{d,T}n^{-\beta(2-\alpha)d/(2(d+1))}.$$

(ii) Let Z_r be a uniform random variable on $B(0,r)\subset\mathbb{R}^d$, then

$$\mathbb{E}\big[e^{i\theta\cdot Z_r}\big]=\frac{2^{d/2}\Gamma(d/2+1)}{|r\theta|^{d/2}}J_{d/2}(|r\theta|),$$

where $J_{d/2}$ is the Bessel function of the first kind of order $d/2$.

(iii) If $M\geq 2$, then under the assumptions of (i) there exist positive a (with $a<1$) and C_d, independent of M, such that for all $t>0$,

$$\int_{|\theta|\geq n^{\beta}}|\tilde{q}_t^{n,[M,\infty)}(\theta)|d\theta\leq C_dn^{\beta d}a^{M-1}.$$

Proof. Let $\epsilon=n^{-\beta(2-\alpha)d/(d+1)}$. For $|\theta|\leq\sqrt{\epsilon n^{\beta(2-\alpha)}}=n^{\beta(2-\alpha)/(2(d+1))}$, Lemma C.4$(i)$ implies for $t\leq T$ and sufficiently large n,

$$|e^{t(\psi^n(\theta)-\psi(\theta))}-1|\leq Ct\big|\psi^n(\theta)-\psi(\theta)\big|\leq C_{d,T}\epsilon.$$

Hence

$$\begin{aligned}
&\int_{|\theta|\leq n^{\beta}}|(e^{t(\psi^n(\theta)-\psi(\theta))}-1)e^{t\psi(\theta)}|d\theta\\
&\leq \int_{|\theta|\leq\sqrt{\epsilon n^{\beta(2-\alpha)}}}|(e^{t(\psi^n(\theta)-\psi(\theta)}-1)e^{t\psi(\theta)}|d\theta+\int_{\sqrt{\epsilon n^{\beta(2-\alpha)}}<|\theta|\leq n^{\beta}}(e^{t\psi^n(\theta)}+e^{t\psi(\theta)})d\theta\\
&\leq C_{d,T}(\epsilon n^{\beta(2-\alpha)})^{d/2}\epsilon+C_d\int_{\sqrt{\epsilon n^{\beta(2-\alpha)}}}^{n^{\beta}}r^{d-1}e^{-ctr^{\alpha}}\mathrm{d}r
\end{aligned}$$

by Lemma C.4(ii). The first term is equal to $C_{d,T}n^{-\frac{\beta(2-\alpha)d}{2(d+1)}}$. Since $tr^\alpha \geq n^{(\alpha-c_2)\beta(2-\alpha)/(2(d+1))}$ in the integral, the second term is bounded by $Cn^{\beta d}e^{-cn^b}$ (with $b=(\alpha-c_2)\beta(2-\alpha)/(2(d+1))>0$). Both estimates combined give us (i).

For (ii), we use Theorem 4.15 of [SW71], which states that the Fourier transform of the indicator function on the unit ball in d dimensions is

$$\int_{\mathbb{R}^d} \mathbf{1}_{[0,1]}(|x|)e^{i\theta\cdot x}\mathrm{d}x = \left|\frac{\theta}{2\pi}\right|^{-d/2} J_{d/2}(|\theta|).$$

Hence, dividing by the volume of the unit ball in d dimensions, which is $\pi^{d/2}/\Gamma(d/2+1)$, yields

$$\mathbb{E}\big[e^{i\theta\cdot Z_1}\big] = \frac{2^{d/2}\Gamma(d/2+1)}{|\theta|^{d/2}} J_{d/2}(|\theta|). \tag{116}$$

Scaling Z_1 by a factor of r gives us the desired result.

For (iii), we recall from (101) the representation of X^n using random walks with step size Y_k. Let R be an $\mathbb{R}$-valued random variable distributed according to $\frac{V_1}{A}r^{-(1+\alpha)}\mathbf{1}_{\{r>n^{-\beta}\}}\mathrm{d}r$, Z be a uniformly distributed random variable in $B(0,1)$ and $\tilde{\rho}(\theta)=\mathbb{E}[e^{i\theta\cdot Z}]$. Then $\tilde{\rho}$ is given by (116), is real and

$$\tilde{q}_t^{n,[M,\infty)}(\theta) = \mathbb{E}_{K_t}\big[(\mathbb{E}_R[\tilde{\rho}(R\theta)^2])^{K_t}\mathbf{1}_{\{K_t\geq M\}}\big],$$

where $\mathbb{E}_{K_t}$ and $\mathbb{E}_R$ are expectations taken with respect to K_t and R, respectively. Observe that

$$\mathbb{E}_R[\tilde{\rho}(R\theta)^2] = n^{-\alpha\beta}\int_{n^{-\beta}}^{\infty} r^{-(1+\alpha)}\tilde{\rho}(r\theta)^2\mathrm{d}r.$$

First, we show $|\tilde{\rho}(v)| = |\mathbb{E}[e^{iv\cdot Z}]|$ is bounded above by a constant $a\in(0,1)$ for $|v|\geq 1$ uniformly. Since Z is radially symmetric about 0, we have $\tilde{\rho}(v)=\mathbb{E}[\cos(v\cdot Z)]=\mathbb{E}[\cos(v_1Z^{(1)})]$, where v_1 and $Z^{(1)}$ denote the first coordinate of v and Z, respectively. It suffices to consider $v_1\geq 1$. Let δ_1 be a small positive constant. If $|v_1Z^{(1)}-n\pi|\geq\delta_1$ for all $n\in\mathbb{Z}$, then $|\cos(v_1Z^{(1)})|\leq \cos\delta_1<1$. Let $I_n=((n\pi-\delta_1)/v_1,(n\pi+\delta_1)/v_1)$, then

$$\mathbb{P}\big[|v_1Z^{(1)}-n\pi|<\delta_1 \text{ for some } n\in\mathbb{Z}\big] = \sum_{n=-\infty}^{\infty}\mathbb{P}\big[Z^{(1)}\in I_n\big].$$

Since $-1\leq Z^{(1)}\leq 1$, the intervals I_n for which the probabilities in the right hand side above are non-empty and have total length $\leq 2\delta_1$. These intervals do not overlap. The way to arrange non-overlapping intervals J_n of total length $2\delta_1$ so that the probability $\sum_n\mathbb{P}[Z^{(1)}\in J_n]$ is maximised is to take $J_1=[-1,-1+\delta_1]$, $J_2=[1-\delta_1,1]$ and $J_n=\emptyset$ otherwise. Therefore

$$\mathbb{P}\big[|v_1Z^{(1)}-n\pi|<\delta_1 \text{ for some } n\in\mathbb{Z}\big]\leq 2\mathbb{P}\big[Z^{(1)}\geq[1-\delta_1,1]\big]\leq 2\delta_2$$

for some $\delta_2\in(0,1/4)$ if we pick a sufficiently small δ_1. This implies

$$\begin{aligned}\mathbb{E}\big[\cos(v_1Z^{(1)})\big] &= \mathbb{E}\big[\cos(v_1Z^{(1)})\mathbf{1}_{|v_1Z^{(1)}-n\pi|\geq\delta_1}\big]+\mathbb{E}\big[\cos(v_1Z^{(1)})\mathbf{1}_{|v_1Z^{(1)}-n\pi|<\delta_1}\big]\\ &\leq (\cos\delta_1)\mathbb{P}\big[|v_1Z^{(1)}-n\pi|\geq\delta_1\big]+\mathbb{P}\big[|v_1Z^{(1)}-n\pi|<\delta_1\big]\leq a\end{aligned}$$

for some $a\in(0,1)$. This estimate implies

$$\mathbb{E}_R[\tilde{\rho}(R\theta)^2]\leq a$$

for $|\theta| \geq n^\beta$.

Second, plugging in $\theta = n^\beta \xi$ yields

$$\begin{aligned}\mathbb{E}_R[\tilde{\rho}(Rn^\beta\xi)^2] &= n^{-\alpha\beta}\int_{n^{-\beta}}^\infty r^{-(1+\alpha)}\tilde{\rho}(rn^\beta\xi)^2 \mathrm{d}r = \int_1^\infty x^{-(1+\alpha)}\tilde{\rho}(x\xi)^2\mathrm{d}x \\ &= 2^d\Gamma(d/2+1)^2\int_1^\infty x^{-(1+\alpha)}\frac{J_{d/2}(|x\xi|)^2}{|x\xi|^d}\mathrm{d}x \leq C_d\int_1^\infty x^{-(1+\alpha)}|x\xi|^{-(d+1)}\mathrm{d}x \\ &\leq C_d|\xi|^{-(d+1)},\end{aligned}$$

where we use the fact $|J_\nu(z)| < Cz^{-1/2}$ for $\nu > 0$ ([AS72], p. 362, 9.1.61). The two estimates above imply that there exist $a \in (0,1)$ and $C_d > 0$ (both independent of M) such that for $|\xi| \geq n^{-\beta}$,

$$\mathbb{E}_R[\tilde{\rho}(Rn^\beta\xi)^2] \leq a \wedge C_d|\xi|^{-(d+1)}.$$

We use this to estimate

$$\begin{aligned}\int_{|\theta|\geq n^\beta}|\tilde{q}_t^{n,[M,\infty)}(\theta)|d\theta &= n^{\beta d}\int_{|\xi|\geq 1}|\tilde{q}_t^{n,[M,\infty)}(n^\beta\xi)|d\xi \\ &\leq n^{\beta d}\int_{|\xi|\geq 1}(a\wedge C_d|\xi|^{-(d+1)})^M d\xi \\ &\leq n^{\beta d}\left(\int_{1\leq|\xi|\leq(C_d/a)^{1/(d+1)}}a^M d\xi + \int_{|\xi|>(C_d/a)^{1/(d+1)}}(C_d|\xi|^{-(d+1)})^M d\xi\right) \\ &\leq n^{\beta d}\left(C_d a^M (C_d/a)^{d/(d+1)} + \int_{(C_d/a)^{1/(d+1)}}^\infty (C_d r^{-(d+1)})^M r^{d-1}\mathrm{d}r\right).\end{aligned}$$

We take $\rho = r/C_d^{1/(d+1)}$ (hence $C_d r^{-(d+1)} = \rho^{-(d+1)}$) to obtain

$$\begin{aligned}\int_{|\theta|\geq n^\beta}|\tilde{q}_t^{n,[M,\infty)}(\theta)|d\theta &\leq C_d' n^{\beta d}\left(a^{M-d/(d+1)} + \int_{1/a^{1/(d+1)}}^\infty (\rho^{-(d+1)})^M(\rho C_d^{1/(d+1)})^{d-1}d\rho\right) \\ &\leq C_d' n^{\beta d}\left(a^{M-1} + \int_{1/a^{1/(d+1)}}^\infty \rho^{-(d+1)(M-1)-2}d\rho\right) \\ &\leq C_d' n^{\beta d}\big(a^{M-1} + (1/a^{1/(d+1)})^{-(d+1)(M-1)-1}\big) \\ &\leq C_d' n^{\beta d}a^{M-1}\end{aligned}$$

if $M \geq 2$. Hence we have established (iii). □

Lemma C.6. *Let $c_3 \in (0,1)$ be a constant. If $M = n^{c_3\alpha\beta/2}$ and $n^{-(1-c_3)\alpha\beta} \leq t \leq T$, then $\mathbb{P}[K_t < M] \leq Ce^{-n^{c_3\alpha\beta/2}}$. Hence, $\int_0^T \mathbb{P}[K_t < M]\mathrm{d}t \leq C_T n^{-(1-c_3)\alpha\beta}$.*

Proof. By a standard tail estimate for the $Poisson(V_1 n^{\alpha\beta}t)$ random variable K_t, since $M \leq V_1 n^{\alpha\beta}t$ we can write

$$\begin{aligned}\mathbb{P}[K_t < M] &\leq e^{-V_1 n^{\alpha\beta}t}\left(\frac{eV_1 n^{\alpha\beta}t}{M}\right)^M \\ &= \exp(-V_1 n^{\alpha\beta}t + M(1+\log V_1 + \log(n^{\alpha\beta}t) - \log M)).\end{aligned}$$

The dominant term in the exponent above is $V_1 n^{\alpha\beta} t$, which is $\geq V_1 n^{c_3\alpha\beta}$, hence

$$\mathbb{P}[K_t < M] \leq C e^{-n^{c_3\alpha\beta/2}}.$$

This establishes the estimate on $\mathbb{P}[K_t < M]$. The estimate on its integral follows easily by splitting the integral over $[0, n^{-(1-c_3)\alpha\beta})$ and $[n^{-(1-c_3)\alpha\beta}, T]$. □

Finally we turn to the proof of our key lemma.

Proof of Lemma C.3.

Recall from (101) the representation of X^n using random walks with step size Y_k: conditioned on R_k, which has density $n^{-\alpha\beta} r^{-(1+\alpha)} \mathbf{1}_{\{r>n^{-\beta}\}} \mathrm{d}r$, $Y_k | R_k = r$ has density $\sqcap_r^{*2}(x)$. Recall also the definition of q_t^n given in (102) and let q be the density of the limiting α-stable process with Laplace exponent ψ defined in (114). We write

$$\begin{aligned} 2\pi |q_t^{n,[M,\infty)}(x) - q_t(x)| &= \left| \int_{\mathbb{R}^d} (e^{t\psi^n(\theta)} - \mathbb{E}[e^{i\theta\cdot(X_t^n - X_0^n)} \mathbf{1}_{\{K_t<M\}}] - e^{t\psi(\theta)}) e^{-i\theta\cdot x} d\theta \right| \\ \leq & \left| \int_{|\theta|<n^\beta} (e^{t(\psi^n(\theta)-\psi(\theta))} - 1) e^{t\psi(\theta)} e^{-i\theta\cdot x} d\theta \right| + \left| \int_{|\theta|<n^\beta} \mathbb{P}[K_t < M] d\theta \right| \\ & + \left| \int_{|\theta|\geq n^\beta} e^{t\psi(\theta)} d\theta \right| + \left| \int_{|\theta|\geq n^\beta} \tilde{q}_t^{n,[M,\infty)}(\theta) d\theta \right|. \end{aligned}$$

Lemma C.5 implies that the first and fourth terms are bounded above by

$$C_{d,T} n^{-\beta(2-\alpha)d/(2(d+1))}, \; C_d n^{\beta d} a^{M-1},$$

respectively, where we also use $t \geq n^{-c_2\beta(2-\alpha)/(2(d+1))}$. The second term is bounded above by

$$C_d \mathbb{P}[K_t < M] n^{\beta d}.$$

Lemma C.4(ii) implies that the third term is bounded by

$$\begin{aligned} \int_{|\theta|\geq n^\beta} e^{t\psi(\theta)} d\theta &\leq \int_{|\theta|\geq n^\beta} e^{-ct|\theta|^\alpha} d\theta \leq C_d \int_{n^\beta}^\infty r^{d-1} \exp\big(- c n^{-c_2\beta(2-\alpha)/(2(d+1))} r \big) \mathrm{d}r \\ \leq & \; C_d \int_{n^\beta}^\infty \exp\big(- c n^{-c_2\beta(2-\alpha)/(2(d+1))} r + (d-1)\log r \big) \mathrm{d}r \\ \leq & \; C_d \int_{n^\beta}^\infty \exp\big(- c n^{-c_2\beta(2-\alpha)/(2(d+1))} r/2 \big) \mathrm{d}r \leq C_d \exp\Big(- \frac{c}{2} n^{\beta - c_2\beta(2-\alpha)/(2(d+1))} \Big). \end{aligned}$$

Combining the estimates for these four terms yields the desired result in (i) for the case $M \geq 2$. Using Lemma C.6, the estimate for $L = n^{c_1\alpha\beta/2}$ follows easily (noting that $n^{-(1-c_1)\alpha\beta}$ is always smaller than $n^{-c_2\beta(2-\alpha)/[2(d+1)]}$ whenever $c_2 < 1$).

For (ii), we observe that it was shown near (62) that the process η_t with generator (62) is a symmetric α-stable process, hence $\eta_t \stackrel{d}{=} t^{1/\alpha}\eta_1$. Let f_{η_t} be the density function of η_t. By Proposition 5.28.1 of [Sat99], since $\int_{\mathbb{R}^d} |e^{t\psi(\theta)}||\theta|^m d\theta < \infty$ for all $m > 0$, f_{η_t} is C^m for all $m > 0$. In particular, this means that the first derivatives of f_{η_t} is uniformly bounded, therefore f_{η_1} is uniformly continuous. This means that

$$|f_{\eta_t}(x) - f_{\eta_t}(y)| = t^{-d/\alpha} |f_{\eta_1}(t^{-1/\alpha}x) - f_{\eta_1}(t^{-1/\alpha}y)| \leq C t^{-(d+1)/\alpha} |x - y|.$$

Hence

$$|q_t(x) - q_t(y)| \le Ct^{-(d+1)/\alpha}|x-y|,$$

as desired in (ii). Part (iii) follows easily from (i) and (ii).

Let f_{X_k} denote the density of $X_k = \sum_{i=1}^k Y_i$. Since the density of Y_1 is radially symmetric and decreasing in $|x|$, the same properties hold for f_{X_k}. Let $X_{k,1}$ denote the first coordinate of X_k, then for $x_1 \in [1,\infty)$ and $\lambda > 0$,

$$f_{X_k}(x_1,0,\dots,0)) \le \mathbb{P}[X_{k,1} \ge x_1 - 1] \le e^{-\lambda(x_1-1)}\mathbb{E}\big[e^{(\lambda,0,\dots,0)\cdot X_k}\big] = e^{-\lambda(x_1-1)}\mathbb{E}\big[e^{(\lambda,0,\dots,0)\cdot Y_1}\big]^k.$$

We would like to estimate $\mathbb{E}[e^{(\lambda,0,\dots,0)\cdot Y_1}]$, for which we calculate, using Lemma C.5(ii),

$$\begin{aligned}
\mathbb{E}\big[e^{(\lambda,0,\dots,0)\cdot Y_1}\big] - 1 &= n^{-\alpha\beta}\int_{n^{-\beta}}^{\infty} \frac{1}{r^{1+\alpha}}\left(\frac{2^d\Gamma(d/2+1)^2}{(r\lambda)^d}J_{d/2}(r\lambda)^2 - 1\right)\mathrm{d}r \\
&= n^{-\alpha\beta}\lambda^{\alpha}\int_{\lambda n^{-\beta}}^{\infty} \frac{1}{\rho^{1+\alpha}}\left(\frac{2^d\Gamma(d/2+1)^2}{\rho^d}J_{d/2}(\rho)^2 - 1\right)d\rho.
\end{aligned}$$

From [AS72], p.362, 9.1.69, Bessel functions are related to generalised hypergeometric functions in the following way

$$\Gamma\Big(\frac{d}{2}+1\Big)J_{d/2}(x)(x/2)^{-d/2} = {}_0F_1\Big(\frac{d}{2}+1; -x^2/4\Big) := 1 + \sum_{n=1}^{\infty}\frac{1}{(\frac{d}{2}+1)\dots(\frac{d}{2}+n)}\frac{(-x^2/4)^n}{n!}.$$

Hence

$$\left|\Big(\Gamma\Big(\frac{d}{2}+1\Big)J_{d/2}(\rho)(\rho/2)^{-d/2}\Big)^2 - 1\right| \le C_d\rho^2$$

for $\rho \in [0,1]$. This implies

$$\begin{aligned}
\mathbb{E}\big[e^{(\lambda,0,\dots,0)\cdot Y_1}\big] - 1 &\le n^{-\alpha\beta}\lambda^{\alpha}\left(\int_0^1 C_d\rho^{1-\alpha}d\rho + 2\int_1^{\infty}\rho^{-(1+\alpha)}d\rho\right) \\
&\le C_\lambda n^{-\alpha\beta},
\end{aligned}$$

where we also use $\big|\frac{2^{d/2}\Gamma(d/2+1)}{\rho^{d/2}}J_{d/2}(\rho)\big| = |\mathbb{E}[e^{i\rho\cdot Z_1}]| \le 1$ in the first inequality. Hence

$$\mathbb{E}\big[e^{(\lambda,0,\dots,0)\cdot Y_1}\big] \le 1 + C_\lambda n^{-\alpha\beta} \le e^{C_\lambda n^{-\alpha\beta}}.$$

which means

$$f_{X_k}((x_1,0,\dots,0)) \le e^{-\lambda(x_1-1)}e^{C_\lambda n^{-\alpha\beta}k}.$$

Plugging the above into the random walk representation yields

$$q_t^{n,[1,\infty)}(x) \le \mathbb{E}_{K_t}\big[e^{-\lambda(x_1-1)}e^{C_\lambda n^{-\alpha\beta}K_t}\mathbf{1}_{\{K_t\ge 1\}}\big] \le e^{-\lambda(x_1-1)}\exp\big(V_1 n^{\alpha\beta}t(e^{C_\lambda n^{-\alpha\beta}} - 1)\big)$$

since $K_t \sim Poisson(V_1 n^{\alpha\beta}t)$. Since $n^{\alpha\beta}(e^{C_\lambda n^{-\alpha\beta}} - 1)) \to C_\lambda$ as $n \to \infty$, we have for $t \le T$ and $|x| \ge 1$,

$$q_t^{n,[1,\infty)}(x) \le C_{\lambda,T}e^{-\lambda(|x|-1)},$$

as desired in part (iv).

For part (v), we obtain,

$$\begin{aligned}
&\zeta_t^{n,[L,\infty)}(x;y,\epsilon) - \zeta_t^{n,[L,\infty)}(x;z,\epsilon) \\
&\quad = \int (\zeta_0^n(x';y,\epsilon) - \zeta_0^n(x';z,\epsilon)) q_t^{n,[L,\infty)}(x-x')\mathrm{d}x' \\
&\quad = \int (\zeta_0^n(x'-y;0,\epsilon) - \zeta_0^n(x'-z;0,\epsilon)) q_t^{n,[L,\infty)}(x-x')\mathrm{d}x' \\
&\quad = \int \zeta_0^n(x';0,\epsilon)(q_t^{n,[L,\infty)}(x-y-x') - q_t^{n,[L,\infty)}(x-z-x'))\mathrm{d}x'. \qquad (117)
\end{aligned}$$

For $t \in [n^{-c_2\beta(2-\alpha)/(2(d+1))}, T]$ and $|y-z| \le 1$, we have

$$\begin{aligned}
&\sup_x |q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)| e^{\lambda|x|} \\
&\quad \le \sup_{x:|x-z|<2} \big|q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\big| e^{\lambda|x|} \\
&\qquad\qquad + \sup_{x:|x-z|\ge 2} \big|q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\big| e^{\lambda|x|} \\
&\quad \le C_{\lambda,d,T}\big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})e^{\lambda|z|} \\
&\qquad\qquad + \sup_{x:|x-z|\ge 2} \min(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))}, e^{-2\lambda|x-y|} + e^{-2\lambda|x-z|})e^{\lambda|x|}\big],
\end{aligned}$$

where we use (iii) for the first term, and (iii) and (iv) (applied with 2λ) for the second. Hence,

$$\begin{aligned}
&\sup_x \big|q_t^{n,[L,\infty)}(y-x) - q_t^{n,[L,\infty)}(z-x)\big| e^{\lambda|x|} \\
&\quad \le C_{\lambda,d,T}\big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})e^{\lambda|z|} \\
&\qquad\qquad + \sup_x (t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})^{1/2}(e^{-2\lambda|x-y|} + e^{-2\lambda|x-z|})^{1/2}e^{\lambda|x|}\big] \\
&\quad \le C_{\lambda,d,T}\big[(t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))} + (t^{-(d+1)/\alpha}|y-z| + n^{-\beta(2-\alpha)d/(2(d+1))})^{1/2}\big]e^{\lambda|z|} \\
&\quad \le C_{\lambda,d,T}(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))})e^{\lambda|z|}.
\end{aligned}$$

Plugging this estimate into (117) yields

$$\begin{aligned}
&\sup_x \big|\zeta_t^{n,[L,\infty)}(x;y,\epsilon) - \zeta_t^{n,[L,\infty)}(x;z,\epsilon)\big| e^{\lambda|x|} \\
&\quad \le \sup_x \int \zeta_0^n(x';0,\epsilon)\big|q_t^{n,[L,\infty)}(x-x'-y) - q_t^{n,[L,\infty)}(x-x'-z)\big| e^{\lambda|x-x'|}e^{\lambda(|x|-|x-x'|)}\mathrm{d}x' \\
&\quad \le C_{\lambda,d,T}\big(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))}\big)e^{\lambda|z|}\int \zeta_0^n(x';0,\epsilon)e^{\lambda|x'|}\mathrm{d}x' \\
&\quad \le C_{\lambda,d,T}e^{\lambda\epsilon}\big(t^{-(d+1)/(2\alpha)}|y-z|^{1/2} + n^{-\beta(2-\alpha)d/(4(d+1))}\big)e^{\lambda|z|},
\end{aligned}$$

as desired. Note that we used the assumption that the support of $\zeta_0^n(\cdot;0,\epsilon)$ is contained in $B(0,\epsilon)$ to bound $e^{\lambda|x'|}$ by $e^{\lambda\epsilon}$. Note also that this calculation holds even if $\epsilon = \epsilon_n$ depends on n. □

Lemma C.7. *There exists $c_5 > 0$ such that for all $t > 0$,*

$$\sup_x \zeta_t^{n,[L,\infty)}(x;z,\epsilon) \le C_d(t^{-d/\alpha} + e^{-n^{c_5}}),$$

where ϵ can depend on n.

Proof. Let $\tilde{\zeta}_0^n(\theta) = \int_{\mathbb{R}^d} e^{i\theta\cdot x}\zeta_0(x;z,\epsilon)\mathrm{d}x$, then $|\tilde{\zeta}_0^n(\theta)| \leq 1$ regardless of ϵ. Let $\tilde{\zeta}_t^{n,[L,\infty)}(\theta) = \tilde{q}_t^{n,[L,\infty)}(\theta)\tilde{\zeta}_0^n(\theta)$, where we recall that $\tilde{q}_t^{n,[L,\infty)}(\theta) = \mathbb{E}[e^{i\theta\cdot(X_t^n - X_0^n)}\mathbf{1}_{\{K_t \geq L\}}]$. Then

$$\begin{aligned}
\zeta_t^{n,[L,\infty)}(x;z,\epsilon) &= \frac{1}{2\pi}\int_{\mathbb{R}^d} \tilde{\zeta}_t^{n,[L,\infty)}(\theta)e^{-i\theta\cdot x}d\theta \\
&\leq \left|\frac{1}{2\pi}\int_{|\theta|<n^\beta} \tilde{q}_t^{n,[L,\infty)}(\theta)\tilde{\zeta}_0^n(\theta)e^{-i\theta\cdot x}d\theta\right| + \left|\frac{1}{2\pi}\int_{|\theta|\geq n^\beta} \tilde{q}_t^{n,[L,\infty)}(\theta)\tilde{\zeta}_0^n(\theta)e^{-i\theta\cdot x}d\theta\right| \\
&\leq \frac{1}{2\pi}\int_{|\theta|<n^\beta} |e^{t\psi^n(\theta)} - \tilde{q}_t^{n,[0,L)}(\theta)|d\theta + \frac{1}{2\pi}\int_{|\theta|\geq n^\beta} |\tilde{q}_t^{n,[L,\infty)}(\theta)|d\theta.
\end{aligned}$$

Since $|\tilde{q}_t^{n,[0,L)}(\theta)| = |\mathbb{E}[e^{i\theta\cdot(X_t^n - X_0^n)}\mathbf{1}_{\{K_t<L\}}]| \leq \mathbb{P}[K_t < L]$, we apply Lemmas C.4($ii$), C.6 and C.5($iii$) to each term above to obtain

$$\zeta_t^{n,[L,\infty)}(x;z,\epsilon) \leq C_d\left(\int_{\mathbb{R}^d} e^{-c_4 t|\theta|^\alpha}d\theta + n^{\beta d}e^{-n^{(c_1/2)\alpha\beta}} + n^{\beta d}a^{L-1}\right)$$

for some $c_4 > 0$ and $a \in (0,1)$. Let $f(t) = \int_{\mathbb{R}^d} e^{-c_4 t|\theta|^\alpha}d\theta$, then $f(t) = t^{-d/\alpha}f(1)$. Hence,

$$\zeta_t^{n,[L,\infty)}(x;z,\epsilon) \leq C_d\left(t^{-d/\alpha}\int_{\mathbb{R}^d} e^{-c_4|\theta|^\alpha}d\theta + e^{-n^{c_5}}\right),$$

for some $c_5 > 0$. This implies the desired result. □